MONASH UNIVERSITY
THESIS ACCEPTED IN SATISFACTION OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
ON................... 25 March 2003 ....

....

Sec. Research Graduate School Committee

[redacted]

CONCEPTUAL DEVELOPMENT IN MECHANICS

by

Charles Taylor, B. Sc., Dip. Ed., B.A.

A thesis submitted in
fulfilment of the
requirements for the
degree of

Doctor of Philosophy

Monash University

2002

TABLE OF CONTENTS

LIST OF FIGURES

BOXES

LIST OF TABLES

MONASH UNIVERSITY

ABSTRACT

CONCEPTUAL DEVELOPMENT
IN MECHANICS

By Charles Taylor

Students learn the concepts involved in Mechanics by first learning prototypical examples. They then learn to distinguish cases which are less prototypical. This is shown to be the case for the concepts of *acceleration* and *force*. The way these concepts are deployed in using Newton's laws is highly dependent on just how close the prototypical situations are to the problems being faced, accounting for many of the difficulties that students face in learning Newtonian Physics.

The students who participated in this study were doing their final year at an independent senior secondary school in Melbourne Australia. Most of them were from non-English speaking backgrounds.

The responses of these students to questionnaires were subjected to factor- and cluster-analyses to clarify the way they understood the concepts *force* and *acceleration*. The test item responses of subgroups of these students, who differed in their conceptualisation of the concept *force*, were compared using ANOVA, to determine the extent to which the conceptualisation of force was associated with the ability to solve problems involving Newton's laws. This statistical approach was complemented by a detailed transcription and interpretation of group problem solving sessions.

In order to obtain a clear understanding of the lifeworld concept of 'force' which these students would have encountered outside the Physics classroom three English language corpora were consulted: these were the CHILDES, COBUILD, and British National Corpora. This enabled the gathering and classification of the different usages of different meanings of the word 'force,' and the determination of the frequency and order of development of other related words.

Approved by the Ethics committee, Monash University, as Project 97/303.

DECLARATION

I declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any University or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material published or written by another person, except where due reference is made in the text of the thesis.

Signed [redacted]

ACKNOWLEDGMENTS

I thank my wife, Nicolette Taylor, for her forbearance during the writing of this thesis and her careful reading of it, and my supervisor, Dr Paul Gardner, for his guidance, his encouragement and his patience. I also thank Dr Deponkar Sengupta for his aid with examining the definition of force in Bengali, and three successive principals of Taylors College, Melbourne – Robert Read, Rod Evans, and Fay Patterson for their support of my studies.

GLOSSARY

ACT-R. A model of cognition developed by John R. Anderson (Anderson, 1983; J. R. Anderson, 1996). It models problem solving by means of networks which use production rules associated with nodes in declarative memory activated by the context of the problem situation.

BNC. The British National Corpus is a very large electronic collection of texts selected to be representative across a wide range of registers – both spoken and written – for example: broadcasts, casual conversations, technical works, books, newspapers etc. It can be rapidly searched for examples of the usage of particular words.

FMCE. The *Force and Motion Conceptual Evaluation* is an instrument devised by Thornton and Sokoloff (1998) to probe the understanding of Newton's laws of motion.

FCI. The *Force Concept Inventory* is an instrument devised by Hestenes, Wells, and Swackhammer (1992) to test students for misconceptions in their understanding of Mechanics. It has been very widely used, both by educational researchers and by physics instructors.

NSM. A Natural Semantic Metalanguage, according to Wierzbicka (1996), is a small subset of words from any language which can be used to describe the meaning of any other word in that language. It is called Natural because the words that are used form part of the naturally occurring vocabulary of a language, as opposed to artificial symbols such as +MALE, which occur in some other theories of semantics.

SBF. The Structure Building Framework is a theory developed by Gernsbacher (1990) that describes the process of comprehension of speech (and written text) as being based on three stages: firstly, the laying of a foundation; secondly, building coherent information onto that foundation to form a meaning structure; thirdly, shifting to a new foundation when incoming information is incoherent with the currently built structure.

Chapter One

Introduction

The understanding of concepts is basic to all forms of communication between people. Especially is this so in the field of science education. Here, the ideas to be communicated are often far from the world of common-sense experience, and the opportunities for misunderstanding are rife.

The literature of science education, or indeed, the recall of one's own experiences in the science classroom, will convince anyone that the quality of communication in science varies widely. Some students learn better than others. Some teachers teach better than others. Some lessons work better than others. Some topics are easier than others.

Presumably there are reasons for these differences. A careful investigation may shed light on these reasons, and perhaps yield useful information enabling us to communicate effectively more often, and miscommunicate less often.

Physics as a 'difficult' subject

Research on science education was changed radically during the mid-seventies by the use of qualitative methodologies, especially interviews asking for explanations (White, 1991 p. 1). Such interviews in the field of physics showed clearly that whatever then current physics assessment was measuring it was not simply understanding of physics. Students who had completed honours degrees in Physics at reputable universities were exhibiting some of the same errors which junior school students made (Peters, 1982).

Here was confirmation, if confirmation was needed, that Physics was indeed a difficult subject. And within the subject Physics, it is the topic of mechanics that has often been singled out as the most difficult area (Gunstone & Watts, 1985).

Thus in investigating conceptual development it is this area which is of central concern. It seems that it is here that misunderstanding is most frequent, and conceptual change is most difficult.

How do physics concepts develop?

The central argument of this thesis is that technical concepts like *acceleration* and *force* are structured and develop in the same way as lifeworld concepts like *cat* or *red*.[1]

This involves clarifying just what is meant by a concept, detailing the way in which lifeworld concepts develop and showing that these technical concepts develop in the same way. The first two of these tasks are accomplished in a literature review, and the third constitutes the bulk of the investigation reported in this thesis.

Evidence from introspection, problem solving and discourse

In investigating the development of concepts several sorts of data proved valuable.

The first of these was data based on introspection. A group of final year students at an independent senior secondary school in Melbourne, Australia, were asked to fill out three instruments designed to elicit their assessment of

- how well a concept fitted a given situation,
- how confident they felt about particular answers to particular problems,
- various aspects of the teaching approaches and learning activities.

Students entered their names on the first two instruments, while the third was answered anonymously to encourage them to write openly.

The data from the first two questionnaires proved useful in uncovering prototype effects (see Chapter Three), and in uncovering the variation in individual students' understanding of concepts.

[1] Within this thesis, in addition to their normal uses, italics and single quotes are used with a special meaning. A single word in italics is used to refer to a concept with that name. Single quotes surrounding a word indicate that the word itself is being referred to. Double quotes are used to indicate quotations. Thus a student may say "Force is mass times acceleration, right," using the word 'acceleration' to refer to the concept *accelration*

The third questionnaire was useful in eliciting some of the motivational aspects of concept learning, an area that is often overlooked.

The second type of data involved students' pen-and-paper answers to a variety of standard types of physics problems completed at intervals throughout the year. These were answered during topic tests just after finishing a topic in class, and during the Mid-year, September and Final examinations.

Some of these tests involved multiple choice items. Particularly important here was the *Force Concept Inventory* developed by Hestenes, Wells and Swackhammer (1992). This test has been widely used and the reported frequency of particular answers could be compared with those of students in the group under investigation. This test consists of sets of multiple choice items designed to elicit misconceptions.

Many questions required numerical answers, while others required written explanations. A few required diagrams to be drawn or graphs to be sketched. These were useful in uncovering long term changes in conceptual understanding. They were also useful in providing evidence of variation in conceptual understanding between students (and within individual students) at any given time.

The third type of data consisted of audio tape recordings of student discussions as they solved, in groups, questions from the *Force Concept Inventory*. Analysing transcripts of these discussions enables one to see in some detail the processes by which the concepts of velocity, acceleration and force were deployed and combined in coming to a scientific understanding of applications of Newton's laws.

A fourth type of data was drawn from the samples of language use accumulated in various corpora: especially the CHILDES (MacWhinney, 2000a, 2000b), COBUILD (2000) and British National Corpora (BNC, 1994). Analysis of these enabled the investigation of the development of some of these concepts outside the classroom, and a description of the non-technical contexts in which people encounter the concept of *force*.

Organisation of the thesis

Following this brief introductory chapter the body of the thesis begins in Chapter Two with a review of the literature dealing with concepts, conceptual change and with misconceptions in physics. This aims to highlight areas where there is general agreement (for example, that students display consistent varieties of misconceptions), areas where there are coherent, although often conflicting schools of thought (varieties of constructivism, for instance), and areas where there has not yet emerged any consensus. Additionally, there is a critique of terminology.

After this the more general area of conceptual development is looked at. Here, work is reviewed in those disciplines that are now interacting in the new field of cognitive science. In particular, work in analytic and hermeneutic philosophy on conceptual analysis and understanding, work in computer science on concept representation, work in psychology on problem solving, memory and mental models, and work in linguistics on semantics, discourse analysis, and structure building are all examined for clues as to how concepts in general are developed and deployed.

Chapter Three looks at the issue of the research methods to be used in this thesis, their limitations, and reasons for use.

Chapter Four analyses the usage of the term 'concept' in science education research, comparing and contrasting this usage with the theoretical accounts already discussed.

The fifth chapter describes students' concepts of acceleration. The sixth chapter is devoted to descriptions of their concepts of force.

Chapters Seven to Nine present a discussion of how the structure of these concepts' meanings influence the usage of Newton's laws.

The concluding chapter discusses the general picture of conceptual development and outlines some implications for teaching and educational research.

Concepts, misconceptions and conceptual development in mechanics

Overview of the chapter

This chapter, consisting mostly of literature review, will explore

- research approaches in science education
- concepts and meaning
- research into the concepts of *acceleration* and *force*
- research into student understanding of Newton's laws

Until the 1970s science education researchers saw the main aim of teachers as being to transmit their own knowledge, or parts of it, to their students. Student difficulties with understanding physics were conceptualised in terms of this viewpoint.

Certainly this view of teaching as transmission of knowledge is appealing. If we know, for instance, where a hammer is and a child doesn't, we can say to the child "Get the hammer off the workbench in the shed." The child now knows where the hammer is: it seems natural to say that the knowledge has been transmitted.

Of course, in some circumstances, the knowledge may not have been transmitted. There may have been too much noise for the child to hear, or the speaker may have mumbled, or the child may be deaf or lazy and so on. Educational research based on the transmission model of learning may deal with such issues: environments conducive (or not) to learning, better modes of presentation, learning difficulties, motivation, individual differences and so on.

However, it does not deal with the activity in the mind of the child when interpreting instruction. It does not take into account that the child must, based on its own current knowledge, construct a meaning for what it hears. Where this meaning is straightforwardly recoverable from the communication we may well speak of knowledge having been transmitted, but the reality is always more complex. It is only when communication breaks down (perhaps the child returns with a mallet from the table in the neighbour's garage) that we may become aware of the complexities that lie beneath the apparent ease with which we transmit information.

This activity in the mind has long been a topic of importance in the humanities. So for instance, in a paper published posthumously in 1836 Wilhelm von Humboldt wrote: "Nothing can be present in the mind that has not originated from one's own activity. … Speaking is never comparable to the transmission of mere matter. In the person comprehending as well as in the speaker, the subject matter must be developed by the individual's own innate power." (von Humboldt, 1985/1836)

Humboldt's emphasis on the activity involved in understanding was one of the key contributions to the field of hermeneutics, the study of interpretation, which has made understanding a key concern (Palmer, 1969). The quotation from Humboldt, taken from an historical anthology of hermeneutic writings, clearly enunciates key issues which have had to be addressed in science education research since the 1970s: the activity of learners in building their knowledge, and the inadequacy of the transmissive view of teaching. However, although hermeneutics is clearly relevant to science education research (Eger, 1992, 1993a, 1993b, 1995) its insights have not been utilised. Perhaps this is due to the separation of the two cultures criticised by Snow (1964). Perhaps it is due to people accepting the argument that sets up hermeneutic understanding as characteristic of the humanities, and as opposed to the explanatory aims of the sciences (Gadamer, 1989; Stanford, 1998). Certainly, hermeneutics is not the only area of the humanities that has not been integrated into the work of science education researchers (see below page 37 for some relevant work in linguistics).

From the perspective of hermeneutics three significant points are highlighted. Let us examine each in turn.

Firstly, interpretation is the continuous and ongoing process of gaining understanding which is characteristic of human beings as human beings (Dreyfus, 1991 pp. 184-214; Heidegger, 1962 pp. 182-195). We will always interpret what we are taught as something that we can make sense of. One can envisage, for instance, that presented with Newton's second law, $F = m\,a$, a student may interpret it as simply a maths exercise as lacking in significance as any arbitrary formula. Or the student may interpret it like a conversion of units: as saying that force is the same as acceleration, just m times bigger. Or, again it might be interpreted as saying that for any force F there is an equilibrating force "$m\,a$", on the model of a "$N = m\,g$" for a book on a table. Or, yet again, it may be interpreted as saying that as well as a force F, there will also be a force "$m\,a$" in the same direction. It may be interpreted as saying that force causes acceleration, or that acceleration causes a force. Evidence for various interpretations of Newton's second law will be discussed below (see Chapter Eight). The mathematical formalism of physics, the axioms and definitions, the algorithms and so on, do not and can not exist in the mind of a student as a separate entity unentangled with interpretations. Thus it is pointless to try to teach the formalism of physics correctly in an abstract way, leaving interpretation until the theoretical apparatus is "in place". Whether teachers like it or not, students will form an interpretation of what is taught, as it is taught: one that is meaningful for themselves; one that is as likely as not a misinterpretation.

Secondly, understanding is an iterative process whereby pre-understandings of a whole form the dominant contribution to the understanding of its parts, and the understandings of the whole are dynamically altered by understandings gained of the parts (Eger, 1992). Thus students' "folk" understandings of physics theory may not be shifted by experience (eg. practical work) alone. Students may see what they expect to see, not what "actually happened". Or students' understanding may be shifted in unexpected directions even when they do see what actually happened (Tasker & Freyburg, 1985). No matter what the outcome of instruction, the key fact remains that student pre-understandings must form the basis for later understandings: there can be no clearing away of misconceptions before proceeding to teach the correct physics. It may be, for instance, that these misconceptions are productive steps in the process of comprehension. The bizarre (from the viewpoint of Newtonian physics) theories put forward by learners may be no more harmful in the long term than the even more bizarre images produced by the mnemonic techniques that have been used by generations of students since antiquity (Yates, 1966). Just as the mnemonics serve to make a meaningful sequence out of an arbitrary list, so misconceptions may serve to impose on the arcane formulae of mechanics an interpretation that is meaningful to the learner. Students are, so to speak, required to live in their houses while they renovate, and to scaffold the emerging structure using only materials that are to hand.

Thirdly, understanding is possible only for social, language-speaking beings like people: it cannot be fully captured in a formalism like the predicate calculus which can be programmed into machines (Dreyfus, 1993). Thus an analysis of student difficulties in terms of explicit but incorrect "beliefs", such as "motion implies force" is necessarily inadequate (McDermott, Rosenquist, & van Zee, 1987). Recognising their inadequacy does not preclude using such formalisms as models to capture some aspect of student learning, but it does emphasise that there will always be other (possibly less precise) but more valid ways of looking at the phenomenon. One can go further: in spite of recognising their inadequacy, one must utilise formalisms. They are our starting point, our pre-understandings, and like our students we must use our pre-understandings in an iterative process to arrive at later understandings.

Recapping: if science educators had been sensitised to these three issues then they would have known, firstly, that children learning science, like all people, continually seek understanding, and, secondly, that children reinterpret what they see and hear. Thirdly, science educators would know that we in turn need to interpret, not just measure, children's understandings. Bearing these three points in mind, then, and with the benefit of hindsight, the existence of "children's science" seems predictable, and its discovery inevitable once qualitative research began to be undertaken.

Children's science

In the event, it was not until the late 1970s that investigators began to uncover what has been called "children's science." Earlier qualitative work by Piaget and his collaborators (described in Ginsburg & Opper, 1969) had not done so – perhaps because of Piaget's concentration on using the results of his investigations, which included transcripts of interviews with young children, as evidence to support his formal apparatus of stages, rather than analysing their content.

In the 1970s a number of researchers like Viennot (1979) and Driver (1983) began to investigate the difficulty people generally have in understanding science, and in particular physics (Osborne & Freyberg, 1985). Students find it difficult to build up some sorts of knowledge. They fall back on common sense where they cannot understand.

Since then it has become well known that students often have idiosyncratic interpretations of what they have been taught in school science classes (e.g. Driver, Squires, Rushworth, & Wood-Robinson, 1995). When they can, students build up their new knowledge on the basis of what they already know, and often this results in misunderstanding. This has sometimes been referred to as Children's Science.

Constructivism

The most important 'theoretical' accounts of Children's Science have been labelled 'constructivism'. There is, however, no single account of the field of constructivism that is accepted by all those who would call themselves constructivists. Nevertheless, there are survey articles (e.g. Driver & Erickson, 1983; Duit, 1995; Gilbert & Watts, 1983; Gunstone, 1988; White, 1991) that provide good introductions to the field. There are also book length treatments (e.g. Fosnot, 1996; Steffe & Gale, 1995; White, 1988a), and critiques of constructivism have also been written (e.g. Fox, 2001; Matthews, 1995; Osborne, 1996; Solomon, 1994).

After going through these the reader finds that the one agreed upon point for constructivists is that learning is not simply a matter of transmission of knowledge from teacher to student. That, on the contrary, students must themselves construct their own knowledge from their experiences (including dialogue with others, as well as physical interactions with the world). As a consequence of that knowledge being built on the foundation of what is already known by the student, it may very well not be the same knowledge that the teacher had intended the student to attain.

Apart from this single point, then, constructivists can and do diverge in their accounts. Thus the scare quotes around 'theoretical' above are significant. Constructivism is neither a well-developed theory in itself, nor are constructivists a coherent grouping of researchers basing their work on a larger scale paradigmatic theory, nor indeed is it either easily distinguished from or integrated with other well developed theories. It is true that individuals within constructivism may try to integrate their work with other bodies of theory: say the theories encountered in cognitive science (White, 1988a). The point here is that constructivists as a group do not integrate their work with any central theory of their own or with those of other thinkers. At least, not if by the word theory one means something like the theories associated with names like Copernicus, Newton, or Darwin: that is to say a core of beliefs deployed to organise and explain a large body of knowledge about the world. It is important to be clear that there is no question here of demanding the same level of explanatory adequacy for educational theory as for physics. It is obvious that, like the other human sciences, educational theory is not a simple matter of causes and effects. Causes in education are multiple and confounded. In any case, much of educational theory is not a matter of assigning causes (explanation) but of understanding the essentially unique educational events which occur in each student's and each teacher's life. Nevertheless, this is not to say that educational theory is of necessity vague or confused, nor that it is simply common sense.

One of the criticisms levelled at constructivism by Matthews and Solomon is that this agreed upon single point is a matter of common sense: that there is no 'theory of constructivism'. Supporters of constructivism argue that this is true, but that this common sense was not so common twenty years ago.

On the other hand, Anderson trying to characterise what goes by the word theory today, says that it 'describes a particular practice of scholarship incorporating leading figures, scriptural texts, characteristic claims, conventional methods, typical performances and practicing members.' (J. Anderson, 1996, p. 8). So perhaps it is possible to identify the theory of constructivism on this basis?

Certainly it is possible to identify many of Anderson's features in constructivism. The leading figures would include Baird, Driver, Duit, Galili, Goldberg, Gunstone, Hewson, McDermott, Minstrell, Osborne, Viennot and White: all of whom are frequently cited in the constructivist literature. As well, there are those who organise large-scale conferences, like Novak.

Agreement with a short quotation from Ausubel is perhaps the best touchstone for identifying a constructivist: 'The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly.' (Ausubel, 1968). Perhaps this would qualify as a 'scriptural text' for some constructivists, in Anderson's terms?

The types of claims that are characteristically made by constructivist researchers are to do with descriptions of ways in which students make sense of science: Driver et al. (1995) provides a handy compendium of such claims. These types of claims are markedly different from the older style educational researchers' claims that were typically to do with statistically significant effects of 'treatments' on samples of students compared with control groups. White describes the gradual changeover in styles of educational research in an Australian context (White, 1991).

The conventional methods of constructivists include the elicitation of alternative beliefs by a variety of means: White and Gunstone provide a useful guide here. Popular methods include: concept mapping; transcribed interviews where students are questioned about their beliefs; class discussions where students discuss their beliefs before teaching commences; the presentation of demonstrations where students are asked to predict what will happen, and explain why they think it will happen, make their observations and then (usually) explain why it didn't happen (White & Gunstone, 1992).

This last, P.O.E. (demonstrations where students must Predict, Observe, Explain) probably shades off into what Anderson refers to as performances.

To round out Anderson's list there are the practising members: the multitude of researchers who have written the thousands of articles that Pfundt and Duit have included in their bibliography of the area (1994). Finally, there are those who attend the academic courses, conference presentations, seminars and workshops run by this multitude of researchers.

Although one cannot identify a unified theory (in the traditional sense), constructivism nevertheless has a certain unity of the type described by Anderson. Constructivism could perhaps best be characterised by Kuhn's term 'pre-paradigmatic', intended to describe the situation where 'the practitioners of a science are split into a number of competing schools, each claiming competence for the same subject matter but approaching it in quite different ways' (Kuhn, 1974). There are certainly a number of competing approaches and disagreements amongst those who fall under the constructivist umbrella.

In so far as the work reported here is intended as a contribution to constructivist theory, it is important to address at the outset the possible limitations and weaknesses of this approach. These exist both internally (disagreements between constructivists) and externally (lack of articulation with other bodies of knowledge). As usual, bald generalisations such as this are actually a matter of degree. In spite of the attempts by writers such as Duit (1995), to get constructivism to cohere, and by writers such as White (1988a) who have made some headway in connecting constructivism with cognitive psychology, I will be arguing that internal disagreements and lack of external connections are areas of weakness in constructivism as a theory. In what follows, first the internal difficulties will be discussed, then the external ones. Finally, an attempt to achieve some sort of synthesis will be made.

Internally, key disagreements exist in four major areas in the conceptualisation of constructivism: the locus of 'construction'; the theoretical entities to be used to describe constructions (mainly reflected in disputes about terminology); epistemological assumptions about the status of constructions; and the relative roles of affect and rational cognition in constructions.

The locus of construction

A number of writers locate constructions in the mind of the individual learner. Exemplifying this tradition, influenced by Piaget and cognitive psychology, is the work of White. He sums up the approach succinctly: 'Each of us builds a world.' (White, 1988a, p. 8) Others in this tradition include Gunstone and Watts (1985) and Posner, Strike, Hewson, and Gertzog (1982). These researchers concentrate on metacognition, conceptual conflict and reasoning.

Others, following Vygotsky, insist that knowledge is inherently social before it is individual, and that knowledge is constructed in the process of interaction between individuals (Vygotsky, 1986). Here are writers like Wertsch (quoted in Confrey, 1995 p. 187):

> Any function in the child's cultural development appears twice, or on two planes. First it appears on the social plane and then on the psychological plane. First it appears between people as an interpsychological category, and then within the child as an intrapsychological category.

A key concept of this approach is the "zone of proximal development", or ZPD, the region of knowledge construction where the student is able to operate with the assistance, or "scaffolding", of a teacher. Researchers in this tradition (like Fawns & Sadler, 1996; Roychoudry & Roth, 1996) investigate the role of small group discussions, and teacher student dialogue.

Still others, social constructivists, claim that knowledge is inherently social, embedded in our language and our culture. The key tradition here is twofold, on the one hand the later Wittgenstein and the sociologists of science such as Bloor, Barnes, and Knorr-Cetina, and on the other the social theory phenomenologists like Schutz, and Berger and Luckmann. Woolgar has written a sympathetic account (1988) and Cole a critical assessment (Cole, 1992).

Indirectly, via historians and philosophers of science such as Kuhn and Lakatos who were themselves influenced by the social constructivists, these traditions have had an input into the most widely accepted account of conceptual change within constructivism. Posner, Strike, Hewson and Gertzog based their description of the stages of conceptual change (dissatisfaction with the current conception, and the intelligibility, plausibility and fruitfulness of an alternative) on Kuhn's and Lakatos's accounts of scientific revolutions within the community (de Jong, 1996; Posner et al., 1982).

While the social constructivists themselves have had little direct influence on science education (rarely being cited, for example) a key tenet of this approach, that scientific knowledge has no special status with respect to any other type of knowledge, has developed independently in science education.

Taking this principle seriously means that students' understandings of phenomena should be examined as seriously when teaching as the accepted scientific accounts. The term 'misconception' was therefore considered by some to be a misnomer, implying only a mistake that had to be corrected rather than a viewpoint that had to be addressed. The terminology being used to describe knowledge constructions therefore became an issue (Gunstone, 1989). Osborne and Freyberg used the term 'children's science' (1985), and Driver, Guesne and Tiberghien used terms like 'alternative conception', 'intuitive notion' and so on (1985).

Issues of terminology

However, choice of terminology is not simply a matter of attitude. It is true that one reason for a constructivist researcher to choose particular terms is that this will reflect his or her attitude toward the worth of the mental entities constructed by students: so that choosing 'misconception' might imply rejection, while 'children's science' would imply a form of acceptance. This, however, is not necessarily what the researcher wishes to convey by the use of the term 'misconception.' Duit, for example, makes a different distinction: 'misconception' implies a mental entity resulting from previous teaching, while 'preconception' is used for mental entities present from other sources (Brown & Clement, 1989). Dykstra (1992) takes yet another position, closest to the one advocated here: "conceptions are fundamental beliefs about how the world works", while concepts are at a smaller "grain-size". Carey (1991) uses the same distinction, but prefers to speak of 'beliefs' rather than 'conceptions:'

> By concept, belief and theory I mean mentally represented structures. Concepts are units of mental representation roughly the grain of single lexical items, such as object, matter, and weight. Beliefs are mentally represented propositions taken by the believer to be true, such as Air is not made of matter. Concepts are the constituents of beliefs; that is propositions are represented by structures of concepts.

Gauld in a brief survey claimed that researchers' other attitudes toward these mental entities accounted for many other words and phrases in common use (Gauld, 1987). For example researchers who wished to emphasise the degree to which these mental entities were part of a structured whole used words like 'frameworks', or 'systems' while those who wished to emphasise their independence used words like 'concept', 'idea' or 'notion'. Researchers who wished to emphasise the degree to which the mental entities influence behaviour used words like 'beliefs' or 'commitments', while those who wished to downplay this aspect would presumably use words like 'thoughts' (Gauld specifies only the first end of this scale).

In an earlier paper, Gilbert and Watts (1983, p.69) had also noted the widespread terminological disagreements in studies of student understanding of science. They discussed the variety of "descriptors of the outcomes of such studies", and attempted to impose some order upon them. They proposed a distinction based on the researcher's viewpoint: a 'conception' would be the subjects' individual "theorising and hypothesising" as reflected in their responses to the researcher's probes; 'categories' would be "groupings of responses" within a study, gathering together those individuals' conceptions "construed as having similar meanings"; a 'framework' would be one step further removed from the data base of responses, generalising across 'categories.'

Their suggestions have not been taken up by the research community. In fact, most constructivist writers follow a similar practice to Gunstone (1988, p. 81) who, after discussing the above sorts of terminological choices and their implications, selects his own preferred cover-term, 'idea/belief,' to stand for "what seems to be one broad issue."

While this all-inclusive label procedure is preferable to sterile debate on terminology, it can serve to paper over genuine differences in approach across research traditions. And, without doubt, it does effectively mask the essential difference between the semantics of concepts (eg. force; mass), propositions (eg. Newton's laws) and higher order entities (eg. chains of reasoning; mental models of situations). It is not uncommon for writers to move from one level to another without noticing. Gilbert and Watts (1983), for example, simply do not distinguish concepts and propositions at a theoretical level. They are quite explicit: "'Universal statements' ... or concepts, as we would now call them..." With other researchers, directing attention away from differences between these semantic levels leads to surface analyses in terms of research methods used instead of in terms of topics being investigated. For example, Gunstone (1988) groups together interviews about instances (clearly investigations of the denotations of concepts, as the task is simply to determine whether an event falls into a category) with interviews about events (where the task involves explanation, hence the production of propositions, and the deployment of mental models). On the one hand, this is clearly reasonable, as both are interviews; on the other hand the differences needed in interpretation of the results will not be addressed. The point being made here is not that the use of a general cover-term is wrong: on the contrary, it is necessary. What is being argued is that there are clear lines of demarcation that are not being utilised or even attended to for lack of a sufficiently clear terminology within which results could be communicated.

A second issue involving terminology is related to the way we interpret student responses. The starting point of most studies about preconceptions is to point out some surprising responses or comments shared by a significant number of children or students (Viennot, 1985).

An example is that of a problem that asks students to mark in forces acting on a ball thrown into the air. Typically students mark an upward force acting on the ball on the way up, zero force at the top, and a downward force acting on the way down.

The researcher then tries to account for the responses. However, this is usually done in the researcher's terms. We deploy our technical notions of motion and force, for example and characterise the student conception as "motion implies force".

As Viennot (1985) points out:

> ...in spontaneous reasoning students are usually not conscious of the 'notion' they use and may call it, sometimes indifferently, 'force', 'impetus', 'energy', 'momentum', and so on.

This raises the issue of the metalanguage to be employed in order to discuss the student responses. What is needed is a vocabulary which is sufficient to describe the meanings used both by scientists and by students, but which does not force the thought processes of one to be described in terms used by the other. A possible method of achieving this is the Natural Semantic Metalanguage described by Wierzbicka (1996), which is discussed later in this chapter.

The side-issue of epistemology

If it were not possible to find a common metalanguage in which to describe student conceptions then we would be faced with an impasse something like that of scepticism: there would always be a doubt as to whether we understood what another person was saying. As Viennot points out above, the fact that a student uses the same technical terms as us does not allow us to infer that he or she means the same thing. However, if one extends this reasoning to all words then we can never really know what anyone is saying! Something like the Natural Semantic Metalanguage is needed as a way out of this sceptical dead-end.

However, not everybody would agree that this scepticism is a dead end. For example, Ernst von Glasersfeld has come to some fairly sceptical conclusions. He has also proposed a quite different sort of terminological distinction (1989; 1992; 1995) to the ones discussed above. He claims that constructivist research can be divided into two types: "trivial constructivism" and "radical constructivism". Both types accept that the learner constructs knowledge. In von Glasersfeld's system the distinction between them is made upon the basis of the epistemological status the researchers assign to the mental entities, or constructs, which they observe. Work which explicitly recognises that such constructs are adaptive (that is, more or less viable rather than true or false) is "radical", the rest is "trivial". Von Glasersfeld holds a sceptical epistemological position, that we have no access to the "real world". Perhaps only a mathematician who deals almost exclusively in stipulative constructs could operate within this framework. While it may simplify mathematical discussions to bracket off the everyday world, it surely makes discussion of physics more obscure. Certainly, although this distinction between radical and trivial constructivism has been widely noted, it does not seem to have been widely adopted by science educators.

Possibly this is because much of the most interesting work done in investigating and describing science learners' conceptions would fall into what von Glasersfeld wants to call the trivial category, making the label at best misleading (a similar point is made by Duit 1993, p. 20). That is, few investigations of student understanding of, say, force and motion trouble to discuss epistemological issues. At best, epistemology is a side issue for physics education researchers. This situation is comparable with that of a native speaker learning grammar. Knowledge of grammar could be useful for a native speaker, but the speaker's key concern is with what is being said. Just so, it could be convenient for researchers to have a well-articulated theory about what knowledge and belief are, but their key concern is with what students know and believe.

And, of course, physics education research is regrettably not even a side-issue for epistemologists (surveys: Kornblith, 1994; Pollock, 1986) so little direct help can be expected from them. Indeed, it could be argued that the label trivial would best apply to von Glasersfeld's epistemological position that viability is the essential property of beliefs. In the long term, natural selection will ensure that this is the case. In the short term, it seems simply false: it could be argued that road toll statistics, for instance, show that many drunks who believed that they were safe to drive, were not.

Von Glasersfeld's arguments are sceptical, in that they attack the notion that *truth* is a useful concept. But, as well as von Glasersfeld, there are others who have come to philosophically sceptical conclusions: people who have questioned whether the notion of *the world* is a useful concept, arguing that worlds are not discovered, but invented.

One such is Strike (quoted in Duit, 1995) who like von Glasersfeld, characterises constructivism by two principles. Firstly, Strike argues that the mind is active in constructing knowledge. Secondly, he argues that concepts are invented rather than discovered. This second distinction can then be interpreted along the lines of idealism: in this interpretation, since all our concepts for dealing with the world are invented, our world is invented and does not exist apart from our conceptions of it. This interpretation does not necessarily follow, however, as it is based upon an ambiguity in what we mean by discovered.

If one accepts Strike's principles, then that concepts are invented follows from his first principle, and in that sense no concept is discovered. However, which concepts out of those that we construct are the useful concepts must of course be discovered, and this must be found in our dealings with the world. We cannot simply decide to invent a useful concept, and have it so. The product mv is useful, and is given a name, momentum, and incorporated as a concept into the structures of physics. The ratio m/v is no more complex in (indeed, less so, as v here must be the scalar speed rather than the vector velocity), no less a possible concept and no less invented: but it proves useless, and is consequently not incorporated in physics. In this sense, then, out of all the possible concepts, those that are useful must be discovered.

A similar point is made by Fox (2001):

> When constructivists argue that we 'construct the world' or that 'the world is a product of minds' we need to resist the temptation to infer that our constructions need only be products of the will, or that the world beyond and independent of mind, is whatever we desire it to be. Indeed, much of our learning consists in coming to terms with the constraints of our own physical and biological make-up as well as the physical and biological constraints of the wider environment. To sum up, conceptual relativity is a breakthrough; it allows us, for example, to realise that the same reality can be represented in many ways. But it need not force us into an implausible subjective or relativistic epistemology. We need to accept that our knowledge is fallible, rather than certain, but who these days denies this? We also need to maintain some form of feedback from the non-human world, in order to avoid falling into an individual or social form solipsism (and, incidentally, in order to survive).

Thus adopting a constructivist position does not entail epistemological positions like idealism, solipsism, or relativism: the resistances and affordances of the world constrain and guide our knowledge construction. However, it is equally true that constructivism does not entail a realist position: constructivists like White (1988a, pp. 1-13) have been happy to espouse relativism. The point being made, once again, is that epistemological issues are not in any direct sense involved – one way or the other – in constructivism.

The conceptual change model

While epistemological issues are, as argued above, largely irrelevant to constructivism, the topic of conceptual change is central. Above, it was argued that we construct those concepts which we discover are useful in our dealings with the world. The question then arises, as to when and under what circumstances we construct new concepts. The most widely accepted account of conceptual change within constructivism is that of Posner, Strike, Hewson, and Gertzog (1982). They argued that learners will change their conceptual structures only under certain conditions which could be found by looking at the history of science. The difficulty of changing conceptions could be attributed then to failure to satisfy these conditions.

Posner et al. speculated that the process whereby students learnt of new physics concepts would be in important respects like the process by which new concepts spread through the scientific community. They therefore looked to historians and philosophers of science, particularly to Kuhn, Lakatos, and Toulmin, for clues to the key elements in students' development of scientific concepts, and they specifically looked to develop a theory which emphasised the rational nature of student concept development: "Our central commitment in this study is that learning is a rational activity" (Posner et al., 1982, p. 212). (By contrast, in this thesis, it is argued that far from being like the historical development of scientific concepts, that students' development of scientific concepts is like their development of non-scientific concepts, and clues to the process are sought from anthropologists, such as Rosch, psycholinguists like Gernsbacher, and semanticists like Wierzbizcka. The process of concept development, it is argued here, is the result of largely unconscious, pre-rational processes.)

In Posner et al.'s formulation of the process by which learners change their conceptual structures ("accommodate" in Piaget's terminology) they require that four conditions be met: the learners must be dissatisfied with their current understanding, and find the alternative intelligible, plausible and fruitful. The evidence put forward by Posner et al. for the usefulness of their criteria was initially obtained from analysis of interviews and think-aloud problem solving sessions with students who were learning Special Relativity. They found differences between students in the intelligibility of relativistic arguments, and also found that students did not find concepts like length shortening plausible. However, they did not find evidence of dissatisfaction with non-relativistic concepts like absolute time amongst the students who were studied, and they also presented no evidence for the role of fruitfulness amongst the students learning new concepts.

While Posner et al.'s four criteria are often cited, another idea has been less widely taken up: that of a 'conceptual ecology.' This phrase was intended to describe the stage of a learner that corresponded with the pre-paradigm stage in Kuhn's description of scientific revolutions. In the pre-paradigm stage different scientific theories compete and there is no clearly preferable approach adopted by the scientific community. In Posner's conceptual ecology, different conceptions are assumed to be available to students, and these conceptions are seen to be in competition. A number of writers have found this a useful way of conceptualising the development of a scientific concept (e.g. Taber, 2001), and a related idea is used in this thesis.

However, Posner et al.'s idea of a conceptual ecology differs significantly from the position put forward in this thesis in two ways. Firstly, in the Posner et al. approach, the two different conceptions are seen to be in competition for across-the-board usage. Whichever gains the ascendancy will be used henceforth in all contexts, and inconsistency in usage is ascribed to the to-ing and fro-ing which occurs as first one, then the other conception gains a temporary lead before a final decision is made. By contrast, it will be argued here that more than one conception is available to students at any given time, and inconsistency in usage is triggered by context specific factors. Secondly, in their approach, the decision as to which conception will be adopted is made according to a set of criteria, modelled on Kuhn's discussion of scientific revolutions. A new conception will be adopted, they claim, provided that students are dissatisfied with their current conception, and find the new one intelligible, plausible and fruitful. (An interesting application of this theory of conceptual change is due to de Jong (1996) who taught his students to use these criteria explicitly and then transcribed their discussions as they wrestled with explaining phenomena.) In their approach, decisions are made consciously and rationally using generally applicable across-the-board criteria.

By contrast, this thesis argues that decisions are typically made unconsciously due to context specific factors. Further criticisms will be levelled at Posner et al.'s conceptual change theory in Chapter Six and Chapter Seven where I will be arguing that it does not adequately describe the cognitive processes involved in learning a new conception. Before moving on to look at the sorts of teaching strategies which have been inspired by constructivism and conceptual change theory there are two further issues to be discussed which are related to the conceptual change theory: interpreting the evidence for conceptual change, and issues of the role of emotion as opposed to cognition in conceptual change. These will be discussed in turn.

Problems in the evidence for (lack of) conceptual change

As discussed above, much of the impetus for constructivist interpretations of science learning was provided by evidence of misconceptions amongst students learning science: even after extensive instruction, the same types of misconceptions appeared (Peters, 1982). Widely used instruments such as the Force Concept Inventory, hereafter *FCI*, (Hestenes et al., 1992) provided further evidence of the widespread nature of misconceptions (Hake, 1998). However, it is important that we re-examine the assumptions upon which the difficulty in correctly answering particular types of problems, such as those in the *FCI*, is used as evidence that conceptual change has not occurred.

In particular, it is argued here that a response to a question which shows that, for example, a non-Newtonian misconception is held, does not necessarily indicate that there has been no conceptual change, nor that Newtonian concepts are not held.

The mechanism that accounts for this apparently counter-intuitive proposal is well known in psycholinguistics: where a word has multiple meanings we automatically and unconsciously select the appropriate one for the context. Analogously, when one holds multiple conceptions in physics one will also automatically and unconsciously select what seems the appropriate one for the context.

Just as the word "bat" when heard immediately conjures up multiple meanings ("flying fox", "stick for hitting ball") which are then automatically and unconsciously filtered according to the context to give the appropriate meaning (Gernsbacher, 1990), so I am arguing a problem situation triggers a number of conceptions, which are normally then filtered according to the context until only a single conception remains, and this is then used to solve the problem. Correspondingly, just as individual concepts attached to a word have different levels of activation (Gernsbacher, 1990), so the different conceptions that can be used in a physics problem also have different levels of activation.

Hence it is possible that conceptual change has been achieved by a student, but that the requirement that a single answer be given to a problem obscures this. Where the correct conception is selected it will seem that conceptual change has occurred, and where a misconception is selected it will seem that conceptual change has not occurred.

In brief, I am arguing that asking for a single answer to a problem forces the filtering out of all but a single conception. There is evidence that students entertain multiple conceptions (Bar, Zinn, & Rubin, 1997; Galili & Bar, 1997; Palmer, 2001; Palmer, 1997; Schau, Mattern, Zeilik, Teague, & Weber, 2001; Tao & Gunstone, 1999; Tytler, 1998; Watson, Prieto, & Dillon, 1997), and furthermore there is evidence that students are aware of multiple conceptions when dealing with particular problems in the *FCI* (Taylor & Gardner, 1999). In this paper Taylor and Gardner asked students to rate each of the five multiple-choice responses for each question on the *FCI* on a scale from 0 (definitely incorrect) to 10 (definitely correct). Students had no difficulty in supplying ratings for the various choices offered as the answers to the *FCI* questions. Arguably, the levels of activation achieved by each answer when all are considered by the student were the source for these ratings for the different choices in the questions of the *FCI*.

What has been argued here is that a commonly used method for examining student concepts is flawed. While the method of investigating student understanding by examining the ability of students to *use* a concept to solve a problem is an advance over the technique of asking students to *recite* a definition, it is not sufficient. That a student can correctly use a concept to solve a particular problem reveals nothing about alternative conceptions that may have arisen in transiently in the process of problem-solving.

Furthermore, investigating the understanding of concepts by getting students to use the concepts to solve problems does not reveal anything about the internal structure of the concept in the learner's mind. The assumption made in most science education research is that concepts have no internal structure. One either knows and understands a concept or one does not. This has been reflected in the way that the issue of misconceptions has itself been conceptualised, that is, as a problem of conceptual *change*. The problem, it has been assumed, is that students believe an incorrect conception (for example, that motion implies force), and that they need to *change* this to the correct conception (Newton's laws).

This has caused two problems. The first is it has led research into dead ends: there are large numbers of papers reporting on (failed) attempts to *change* conceptions. Invariably it is found that changed understanding is a matter of degree — of higher mean scores on this or that test (for example studies using the FCI used in this research Hake, 1998), but not a matter of a complete changeover of conceptions. Students do not simply switch from non-Newtonian to Newtonian: rather, they use different conceptions in different contexts. This has led to the consistently reported research result that conceptual change is *difficult*. I argue that conceptual change (in this sense of *replacing* one concept by another incommensurable concept) is *impossible*. In the same way that we do not change our understanding of the word "bat" by replacing its meaning "flying fox" with "stick to hit ball", but rather retain both meanings and deploy them in the appropriate contexts, so we deal with physics conceptions. In addition, I argue that conceptual change, in the sense of seeing that the use of one conception (e.g. Newtonian) is better than another (e.g. force implies motion) in a particular context is comparatively *easy* (see Chapter Six).

A second problem arises from the failure to see that conceptions are simultaneously available. This is the misinterpretation of the often noted context dependency of conceptions (Finegold & Gorsky, 1991). The standard interpretation has been that particular conceptions are tied to particular contexts. This begs the question of how a conception could have been attached to a context which the student had never encountered before the test question brought it to mind. It also makes it hard to see how one can ever succeed in generalising. By contrast, the position put forward here is that in each context several conceptions are available, and that the one chosen is simply the one which achieves the highest activation at the time at which a decision has to be made. Conceptions, then, are completely general but particular contexts bias us toward using one or another. (Analogously, in the context of a sentence beginning "The cricketer..." one would be biased toward using one meaning of "bat", but this meaning is not tied to that context.)

The roles of emotion and cognition

The critique of the conceptual change model which has been offered above has been centred on what one could call the cognitive aspects of conceptual change. However, others, such as Garrard (1986) criticise this conceptual change model on different grounds, arguing that this sort of theorising overestimates the importance of analytic, logical, rational thought processes and does "not address the whole thinking, feeling, and socially interacting student."

Gunstone, in a lecture, has emphasised the role of affect in learning with an anecdote of a student who stayed loyal to the "clashing currents" model of electricity throughout an entire course: he believed this model, he said, because his father was an electrician, and that is what his father had told him.

And White, too, argues that affect is important in the learning of concepts: the remembered thrill of fairground rides must be part of the concept of circular motion, according to his model of what a concept is.

A number of writers have emphasised the importance of the way students feel about their studies in general. Those that feel as if learning is out of their control, a matter of talent, luck in tests, and so on, have less success than those who feel that they are in control (Marton & Booth, 1997).

There is a good deal of empirical work that has been done in investigating the role of affect in learning, but most of this has been outside the science education field. For example the role of mood, trauma, and "shockingness" in recall have all been investigated (Bower, 1992). Mood dependent recall is the recall of memories which are congruent with one's current mood: depressed people, for instance, more frequently recall distressing events from their past than happy ones, happy people the reverse (pp. 22-23). Post traumatic stress syndrome is characterised by the persistent recall and preoccupation with an emotionally loaded happening. People who are in this situation have difficulty in learning, their attentional resources being siphoned off by their obsessive concern with the traumatic event (pp. 14-15).

People learning arbitrary lists, it is well known, recall best the first and last items from the list: investigators who asked their subjects to memorise sequences of photographs of 15 emotionally neutral objects got exactly this effect. Then they switched one of the middle neutral photographs for a "shocking" human nude. In spite of its middle position in the order of presentation the vast majority of subjects recalled this picture: more than the number who recalled the first or final items. Interestingly, the items immediately before and after the shocking item were recalled less well (Bower, 1992, pp. 18-20).

General accounts of the role of affect and learning in science education have tended to be hortatory rather than investigative (eg. Stone & Glascott, 1997), although some descriptive work has been done on the teaching of potentially emotional topics such as radiation (Alsop & Watts, 2000). However, these papers have focussed on the motivational aspects of making the subject Physics attractive, rather than on the role of affect in the learning of concepts.

Thus we have these scattered investigations, some undeveloped proposals and some anecdotes but the need for coherently incorporating Garrard's "whole thinking, feeling, and socially interacting student" into constructivism still exists. Currently a number of writers are beginning to explore issues of social constructivism, and situated cognition. It remains to be seen whether they will achieve this integration of the whole person into a constructivist theory.

Summary of the conceptual change model

The conceptual change model of Posner et al. developed as a psychological analogue to sociological accounts of scientific change and has been highly influential. It introduced the idea that conceptual change required four conditions: learners must be dissatisfied with their current conceptions, and find the new conception to be intelligible, plausible and fruitful. It has been criticised because of its bias toward cognitive as opposed to emotional factors, but is disputed in this thesis because of its assumptions about the nature of conceptions.

The practice of constructivism

A point that has been made by a number of writers is that, if it is granted that all learning is by its nature constructivist, then the phrase 'constructivist learning' is at best a tautology and at worst misleading. Where learning has occurred, whether in a traditional lecture, a practical session in a laboratory, a display at a museum or anywhere else, the learners have constructed their knowledge.

The phrase 'constructivist teaching' is less open to this line of criticism. Not all teachers take advantage of the knowledge that learning is constructed by the learner. Those who do so may well employ methods of teaching which are seen as likely to be effective according to the tenets of constructivism. Teachers may endeavour to make students dissatisfied with their current conceptions by the process of eliciting student conceptions, encouraging students to discuss their ideas with each other so that a variety of conceptions can be compared, and then confronting their students with counter-intuitive experiments, and so on. In so far as such teachers differ from others who rely on simply telling students what they need to know, or who rely on discovery learning in practical classes, then one needs a label to describe them, and 'constructivist teacher' serves as well as any. In fact, with its emphasis on the activity of the learner, the teaching methods often advocated by constructivists are similar to what is often called 'progressive education' (Dewey, 1963; Lawrence, 1970). It may be that for some the two ideas become conflated.

However, it is important to bear in mind that constructivism is essentially a descriptive (as opposed to prescriptive) account of learning. The understanding that knowledge is constructed by the learner may suggest that certain methods are superior to others in the classroom, but the proof of the pudding is in the eating: where a method of teaching results in learning, then knowledge construction has taken place. That is to say, 'constructivist learning' is, simply, learning. That is, the personal constructs of students are as important in traditional classrooms (ie. those where the teacher and students hold 'transmissive' views of learning) as they are in self-consciously 'constructivist' ones.

Metaphorically, the teacher is a knowledge transmitter only in the manner of the 1950s science fictional matter transmitter. In a well-known movie only the instructions on how to reassemble the matter were actually sent, and the matter itself was reconstructed from available resources at the receiving end. But then someone was reassembled at the receiver on the basis of the DNA of an intruding fly...

What is disheartening to the physics educator is just how often students construct such conceptual 'monsters'— and just how hard it seems to be to eradicate them.

Misconceptions research

Research into such misconceptions can be classified in a number of ways. Pfundt and Duit, in their well known bibliography (1994), classify the papers into nine groups, according to their content:

1. General research in this area
2. Research into the relations between everyday conceptions and scientific conceptions
3. Comparisons of the history of science with individual conceptual development
4. Research into the relation between language and conceptions
5. Discussions of methods of investigation
6. Information about students' conceptions, further subdivided by topic. (This is the largest group.)
7. Investigations into teaching which takes students' conceptions into account, again subdivided by topic.
8. Investigations of teachers' conceptions, subdivided by topic.
9. Conceptions and teacher training.

Perhaps it is easier to collapse these into three strands: analytic (their groups 1 to 5), descriptive (their groups 6 and 8), and eliminative (7 and 9). These categories of research are not completely clear-cut, nor are they complete, nor exclusive, but they do provide a useful scaffold for discussion. The first, analytic, strand of research has been discussed above.

From research in the second, descriptive, strand we have a fairly clear notion of what sorts of misconceptions are around, where they are likely to be found, and which misconceptions are likely to co-occur (eg. Clement, 1982; Driver, 1995; Driver et al., 1985; Viennot, 1979).

From the third, 'eliminative', strand we also have some indications of the sorts of educational 'treatments' which might serve to reduce the incidence of new misconceptions, and to aid in the elimination of misconceptions which children bring with them to their classes (reviewed in Scott, Asoko, & Driver, 1992). Two main groups of approaches are centred around 'conceptual conflict' and 'conceptual extension' (p. 312). A third grouping centres on getting students to develop mental models (Halloun, 1996, 1998; Hestenes, 1992, 1996).

The conceptual conflict grouping includes those who focus on events that conflict with student preconceptions (eg. Nussbaum & Novick, 1992), those who focus on conflicts between student preconceptions (eg. Stavy & Berkovitz, 1980), and those who focus on conflicts between student preconceptions and scientific conceptions (eg. Champagne, Gunstone, & Klopfer, 1985).

The conceptual extension grouping includes those who use a 'bridging strategy' to aid in the generalisation of concepts (eg. Brown & Clement, 1989), and those who use a 'substitution' strategy to aid in the differentiation of concepts (eg. Grayson, 1994).

While some progress has been made, there has been no discovery of a 'magic bullet' that would enable educators to rid all students of their misconceptions, nor should we expect such a discovery. Where, as in New Zealand, constructivist theories have been used to aid in official curriculum prescriptions, they have been subject to the sort of polemics which often accompany curriculum reform (Matthews, 1995).

That similar problems occur in the USA can be gathered from the following introduction from Green and Gredler (2002)

> Educational movements, complete with recommendations for major shifts in teaching practices, periodically emerge in American education. These movements typically arise as a reaction to existing practice, are ill-defined and unsupported by research, and gain widespread currency as a result of their intuitive appeal. Often they consist of efforts to translate a complex conceptual framework into classroom activities. One example in American education is the project method, a distortion of John Dewey's progressive education. The current educational movement with these characteristics is known as constructivism. It emerged, in part, in reaction to the "overselling" of the computer as a metaphor for learning (Bredo, 1994), and the perceived transmission-of-knowledge focus of information-processing theory (Marshall, 1996). The movement currently is prominently featured in academic and practitioner journals and books (e.g., Brooks & Brooks, 1993; *Educational Leadership*, 57{3}; *Educational Research*, 23{4}; *Journal of Special Education*, 28{3}; *Learning Disabilities: Research and Practice*, 11{3}; Richardson, 1997), and it has played an influential role in policy formation (e.g., National Council of Teachers of Mathematics, 1989). However, with only a few exceptions (e.g., Brown & Campione, 1994; Palincsar & Brown, 1984), empirical research on constructivist classrooms has yet to be conducted. *(Note that the references are part of the original text and are not included in the bibliography at the end of this thesis.)*

Lacking the clear-cut successes which would enable ill-informed or misconceived critiques to be shrugged off has meant that constructivists have had to engage in point by point rebuttals (see, for instance, the account in Bell, 1995).

Given that so much of constructivist teaching has been concerned with the aim of providing students with conceptual understanding, it is important to be clear about exactly what is meant by this. What, one might ask, is a concept? This important topic is taken up in the next section.

Concepts

A concept has generally been understood as an ability to categorise things (see e.g. Bolton, 1977; Howard, 1987). So knowing the concept dog enables one to categorise the things one encounters into dogs and non-dogs, for example.

The traditional account of categorisation was formulated by Aristotle in *The Categories* and in Book VI of *The Topics* (Aristotle, 1899). A definition of a category requires one to specify the genus and the differentia. Thus a human was a rational (difference) animal (genus). By specifying differentia more and more closely one could form sub-categories to any level of detail which one wished. A concept was defined by giving a list of individually necessary and jointly sufficient differentia (in effect, conditions that had to be satisfied). Knowing the concept was thought of as knowing the definition of the category.

The so-called "method of definition" was used by early psychologists in their investigations of concepts. Subjects would be presented with a concept and asked for its definition, or presented with a group of attributes and asked what concept they defined, or given a group of specifics and asked to name the genera, and so on. These investigations were criticised on the grounds of their being exclusively verbal, liable to elicit rote responses of school learned definitions (*Scheinbegriffe*) as opposed to genuinely thought through concepts, and dealing only with the completed concept and not its genesis (Sakharov, 1994/1930, p.74).

To address some of these criticisms an alternative method of investigation was developed. It aimed to present subjects with the task of abstracting common features from presented figures or objects. In one version subjects were presented with two sets of shapes. One shape was common to both sets, and the subject had to identify it: in theory this indicated that the features common to both had been abstracted. In another investigation, which sought to find which features would be abstracted, young subjects were shown, for example, a red triangle and asked to select the same thing from a set that contained only red circles and white triangles. (It was found that subjects up to the age of approximately five would pick a red circle, relying on the feature of colour, while older subjects would pick a white triangle, having abstracted the feature of shape.) These and other related methods of investigation were criticised because they did not involve language, whereas concepts in real life were heavily involved with the use of language. In addition, both of the exclusively verbal and the exclusively thing based methods were unrelated to goals: yet in everyday life words for concepts were developed in response to perceived needs (Sakharov, 1994/1930, pp. 76-9).

Vygotsky and his collaborators developed an experimental task to investigate language formation that aimed to overcome the above limitations. Subjects were presented with a game board containing pieces of different shapes, colours, sizes, and so on. On the bottom of each piece was written a short nonsense word. The experimenter turned over one piece to show the label, say *bek*, and informed the subject that the pieces on the board were from a game in another country where this piece was called a *bek*. The subject was then given the task of putting all the *beks* to one side without looking at the labels. After each failed attempt the experimenter would turn over a misclassified piece to expose its label, until the subject had successfully completed the task. The task successfully combined words, objects and goals and allowed the experimenter to observe the gradual acquisition of the artificial concept *bek*. The results of these experiments, and comparison with earlier experiments enabled Vygotsky to argue against the notion that concepts could be formed by simple processes of association of commonly co-occurring features: the interaction between language and goals was an essential ingredient in concept formation (1994/1931, p. 203). He also noted three stages in the way subjects responded to the task: in the first stage, syncretic grouping, the blocks are grouped randomly or by contiguity, in the second, thinking in complexes, blocks are sorted by combinations of shape and colour etc., and in the third stage thinking in concepts, the subject is able to logically solve the problem (see Bolton, 1977, pp.68-72). The distinction between the two later stages was that in the second the concept could be used and also dealt with abstractly and explicitly, whereas in the first stage it could only be used (Vygotsky, 1994/1931, p. 251).

The idea of stages in the development of conceptualisation was not unique to Vygotsky. Of all educational researchers, Piaget is perhaps most closely identified with the idea of developmental stages. Piaget was endlessly inventive with experimental tasks and details will not be given here (see eg. Ginsburg & Opper, 1969 for details). With respect to categorisation Piaget claimed that there were three stages. In the first, from about 2 ½ to 5 years of age, a child can categorise objects, but constantly changes the features by which they are classified: first they are sorted according to one feature then by another. From about age 5 to 7 years the child is capable of non-figural collections: coloured shapes are sorted into appropriate classes of say squares and circles, and these then in turn subdivided into say blue and red. However, according to Piaget, they have not achieved an understanding of class inclusion. (So, for example, when asked if there are more chickens or animals in a particular group containing four birds and two horses, the child replies more chickens. Of course, Piaget's interpretation of this result is not necessarily correct – children are likely to understand the word 'animal' to mean something with fur and four legs. If this is the case, then the child's answer is straightforwardly correct: there are four chickens but only two four footed furry creatures.) In Piaget's system, from the age of 7 onwards the child has entered the stage of concrete operations and becomes capable of understanding class inclusion. However, even at this stage, children tend to create classes from abstract classes: mice and ducks are grouped together because they are "fairly small animals" for example (see Bolton, 1977, pp. 68-72).

Bruner built on the work of both Vygotsky and Piaget in his investigations of concept development. In early work (Bruner, Goodnow, & Austin, 1956) subjects were successively presented with cards, which showed figures in different numbers, shapes and colours, and had to decide whether or not the card was a *mib*, say. Their only evidence was the experimenter's response of correct or incorrect to each successive judgement. Their strategies for determining the defining features of the concept *mib* were investigated. Some chose a focusing strategy, where the initial card was used as a standard and others judged by their similarity to it. Others chose a hypothesis (*mibs* are red, say) and discarded it for a new hypothesis on receiving a disconfirming instance. Bruner's investigation of strategies was a new element in the understanding of conceptual development. In later work he (Bruner, Olver, & Greenfield, 1966) used a method where subjects were presented with either a word or a picture of one common object, say a banana, and then another one, say a peach. They were then asked to say what was common and what was different for the two items. Then a third item, say a potato, was added and again the subject was asked what they had in common and how they differed. This process was repeated a number of times. Bruner summed up the results in a list of five main ways in which subjects classified things:

1. by perceptible qualities (e.g. size or colour)
2. by functional equivalence (e.g. we eat them)
3. by affect (e.g. I like them both)
4. by ready made categories (e.g. fruit, vegetables)

5. by fiat (e.g. "they are the same")

Bruner's results in this experimental situation show that his subjects classified things by using criteria which were not mutually exclusive. For example, given a mango and an orange, one of which was liked and one disliked, both which could be eaten, with similar colours, but different sizes, a subject could classify them as similar or different depending on the criterion used at the time.

One response to this variability in classification is to look for the meaning of concept in ways other than those to do with classification. One of the few writers to consider the nature of concepts in science education is White (1988, pp. 22-48) whose discussion of concepts is not related directly to categorisation, but rather to association. White argues that students associate concepts with different memory elements, and that the particular associations determine the meaning of the concepts.

In particular, White proposes that concepts (like density) should be analysed as arbitrarily grouped associations of memory elements with a 'label'. Thus a child's concept of force and an adult's will differ according to their experiences and their resulting associations. One person's concept of 'force' might involve an association of a set of propositions like 'a force is a push or a pull' together with episodic memories of 'pulling on springs and elastic', images like 'arrows for vectors' and intellectual skills like 'can add two vectors' (p. 54). Another's might lack some of these, but have sufficient others in common for communication to be achieved. On this account, concepts would be more like elaborate encyclopaedia entries than terse dictionary definitions (Haiman, 1980).

In some ways this is an appealing idea. It gives a concrete theory of what a concept is. It helps, too, to explain the difficulty which young students have in understanding the adult teacher's concepts: since the associated elements are bound to be fewer for younger students, communication will be more difficult. It also provides some insights into the difficulties associated with curriculum implementation: teachers associate different elements with the labels in the course outline: it is only after some time that a consensus as to what is and what is not on the course can develop.

In spite of its appeal, there are some serious problems with such a theory of what a concept is. These problems are basically centred on the lack of concept structure which is implied, firstly by a single mechanism (association via 'labels') for concept formation, and, secondly by a single measure of similarity (set intersection) which is used (White, 1988 pp. 44 and 46). As will be discussed below, purely associationist theories are unable to account for many easily observed properties of concepts, including their relations with other concepts (eg. synonymy or antonymy), and the internal structure of concepts (especially the importance of prototypes). Therefore, it seems that an account of concepts, such as White's, that does not deal directly with the relationship of concepts to categorisation is inadequate.

From the above accounts of the work done in psychological investigations into the development of concepts by Vygotsky, Piaget and Bruner, and despite occasional challenges such as White's, it is plain that Aristotle's account of categorisation has proved remarkably robust. It inspired useful hypotheses and investigations, and could be used to arrive at meaningful interpretations of the results of these investigations. It survived until well into the twentieth century. However, it entails at least three points that have proved to be problematic (Lakoff, 1987 pp. 12-57).

Firstly, the traditional account assumes that there are some things that are essential to the category and others which are accidental. Thus humans are rational animals in essence, and featherless bipeds by accident. That is to say, although featherless bipeds happens to coincide with the class of rational animals in the world as it is, so that either description would pick out the same set of creatures, the first is a 'true' definition; the second does not get at the essence of what it is to be human. (One cannot create a human, for example by plucking a duck.) Thus the traditional account assumes that there are 'natural kinds' (eg. gold, water) defined by essential qualities which are independent of humans. If something possesses these essences it is in the category, otherwise it is outside the category.

Wittgenstein showed that a concept is not necessarily a matter of shared essences (Wittgenstein, 1953, pp. 31e-35e). He pointed out that for the concept 'game' there were no features common to all exemplars. Instead he described the concept of 'game' as being characterised by what he called family resemblances. No game needs to have any particular set of features, as long as it has some sufficient subset of gamelike features (for example: competition, amusement, skill, chance) it will still be called a game. At the very least, then, our representation of the meaning of a concept must be expanded to cover such cases.

Secondly, the traditional account entails that all members of a category possess equivalent status in the category: an emu is as much a bird as a robin; a whale is as much a mammal as a dog.

Rosch and her colleagues gathered extensive empirical research that demonstrated that this is simply not so (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Rosch started by replicating earlier work by Berlin and Kay (1969) in which certain colours were shown to be cross-culturally selected as 'good' examples ('focal colours') for colour categories. Her earliest evidence for prototypes (Rosch, 1978; Rosch & Mervis, 1975; Rosch et al., 1976) came from studies in which people were asked to pick out from an array of colour chips examples of a particular colour, for example, red. It was found that everybody picked out central instances but that people made different decisions about whether or not to include colours near the boundaries: where red left off, and orange took over, for example.

She and her co-workers showed that for any given concept there were 'prototypes', which played the same role of 'good example' as the 'focal colours', did for colours. Thus, when asked for an example of a 'bird' people will give an example of a 'robin' far more often than 'emu' or 'penguin'.

Furthermore, it was discovered that certain concepts had multiple prototypes. For example there are languages where the colour terminology is restricted to a very few terms, such as the Dani language of New Guinea, which uses just two terms 'mola' (referring to the warm colours) and 'mili' (the cool colours) to divide the whole of colour-space. In this language, different informants asked to identify the best examples of *mola* sometimes identify the prototypical *red* and sometimes the prototypical *white* of English (see Kay & McDaniel, 1978, pp. 616-617).

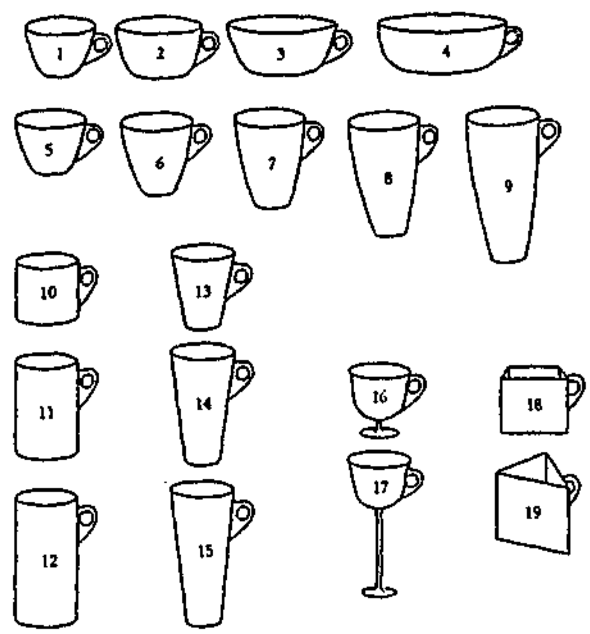


Figure 1: Varieties of 'cup' from Labov (1973), reprinted in Goddard (1998).

Similarly, in Labov's (1973) investigations of the meaning of 'cup', subjects were presented with arrays of drawings where the height, width and other features of a container were systematically varied (see Figure 1). Again, people agreed on the prototypical cup (labelled 1), but disagreed on the instances near the boundaries with bowl (4), or mug (11) and so on.

A number of different experimental methods lead to the same conclusion, that certain exemplars are more central to a concept than others. These include (Lakoff, 1987, p.41):

- responses to requests for examples (as above).
- direct rating of the typicality of examples of a concept using a Likert scale.
- reaction time measurements. (Shorter times are taken to decide on the truth of statements of the form 'An X is a C' if X is a prototypical example of the category C.)
- assymetry in similarity ratings: Y is rated more similar to X than X is to Y, when X is more prototypical than Y.

Thirdly, the traditional account of concepts accords no priority to categories in the middle of a taxonomic hierarchy. Given a hierarchical sequence like 'Animal: dog: retriever' or 'Furniture: chair: rocker' the traditional account makes no prediction of any priority for a particular level. Yet Rosch and others claim that the middle term in the hierarchy is more basic than others. Amongst other effects, basic level terms like 'dog' or 'chair' are (Lakoff, 1987, p. 46):

- identified most rapidly
- the first to enter a child's vocabulary
- represented by the shortest words
- used in neutral contexts (i.e. it would be natural to be told 'There is a dog on the porch', but 'There is a mammal on the porch' would indicate something unusual in the situation).
- the level at which most of our knowledge is represented.

With respect to the latter point Lakoff (1987, p. 47) writes

> The fact that knowledge is mainly organized at the basic level is determined in the following way: When subjects are asked to list attributes of categories, they list very few attributes of category members at the superordinate level (furniture, vehicle, mammal); they list most of what they know at the basic level (chair, cat, dog); and at the subordinate level (rocking chair, sports car, retriever) there is virtually no increase in knowledge over the basic level.

Taken together with identification of a concept with the ability to categorise, these three points have a number of implications for physics education. Take the example of kinematic and dynamic quantities: the student is faced with a class of 'measurement' entities (eg. time, position, displacement, distance, direction, velocity, speed, acceleration, deceleration, momentum, force, and others defined in terms of these but not usually given names: rate of change of acceleration, occasionally called 'jerk'; rate of increase of speed, which is not usually distinguished from 'acceleration' proper). Each of these entities can be applied to any given physical object, so there can be no question of classing physical objects: these representing velocity, those representing acceleration, as can be done with objects representing concepts like 'cat', or 'bird'. This level of abstraction, while certainly not unusual – abstraction is obviously involved in classifying actions such as 'lying' (Coleman & Kay, 1981), for example – can nevertheless be expected to make interpretation of definitions more liable to error. In this context, the three points referred to earlier have the following consequences.

Firstly, the fact that a definition is given to students, say 'acceleration is rate of change of velocity', does not mean that students will necessarily assume that something with a 'family resemblance' (velocity for instance), will fall outside the category.

Secondly, certain exemplars will be prototypical: I will argue later that 'increasing speed' is the prototype for 'acceleration', and that 'imposition of one's will' is the prototype for an everyday sense of the word 'force', while a kick or punch is a prototype for the physics-sense of 'force', for example (see pages 37 and Chapter Six respectively).

Thirdly, there is the empirical issue of identifying the psychologically basic level, superordinate and subordinate concepts. (For example, while a physicist might adopt the hierarchy 'Any acceleration: uniform acceleration: constant velocity: constant zero velocity', where each term is a special case of the term before, students are more likely to have a far less differentiated hierarchy 'Moving: not moving' and as a result will very likely fail to group together the latter two of the physicist's hierarchy when dealing with Newton's laws.)

Overall then it can be seen that there is a large body of evidence which runs counter to the assumption in science education research that concepts have no internal structure. That this assumption is made is beyond doubt: a survey of the literature of science education research in English, involving extensive reading and electronic searching of databases indicates that the prototype structure of concepts in general has not been taken into account. (I have only found two mentions, one in Brown, 1992 where it is not elaborated upon; and the second in di Sessa & Sherin, 1998 which mentioned the idea only to dismiss it out of hand).

By contrast, in the body of this thesis there will be a good deal of discussion about the structure of student concepts in mechanics, and this will illustrate further the relevance of the work discussed in this section. In this discussion, much of the interest will centre on the meanings attached to concepts. It is, therefore, important to be clear about just what is meant by 'meaning'. This will be discussed in the next section.

Meaning

The study of meaning is semantics, which has a rich literature in its own right, and which cannot be surveyed here in detail. The discussion in this section is focussed only on developing five points which are relevant for the discussion in the body of this thesis. These points are not controversial, and treatments of them can be found in any of a number of places (for example: Aitchison, 1994; Hurford & Heasley, 1983; Leech, 1981; Lyons, 1981).

The first point which needs to be clarified is the distinction between 'sense' and 'reference': the reference of a word or expression is that thing or class of things which it picks out; the sense is metaphorically speaking the mental roadmap which enables one to do the picking out. The classical example of this distinction is 'the morning star', 'the evening star' and the physical object, the planet Venus. The two expressions have different senses, but the same reference. And of course there are many words with sense but no reference: dragon, unicorn, and so on: we know how to identify such creatures, but they aren't there to be identified. And again, there are many words which have a well-defined sense, but no possibility of reference: *therefore*, *if*, *and* and *but* for example.

Secondly one must distinguish between the terms 'vagueness', 'ambiguity' and 'polysemy'. A meaning is vague in so far as there is some indeterminacy in its reference. The word 'bald', for example, is vague because there is no precise number of hairs on one's head which prevents one from being classified 'bald'. Ambiguity refers to words like 'bank' that have several unrelated meanings: 'side of a river' or 'financial institution' in this case. Polysemy refers to a word like foot, with several related meanings: for example a person's foot, the foot of a mountain, and the foot of a bill.

Thirdly one must distinguish word meaning from phrase meaning. The meaning of a phrase 'Dog bites man' for example is not simply the sum of the meanings of its component words, as one can see by re-ordering it.

Fourthly, one must distinguish between a proposition, a sentence, and an utterance. The proposition represents the logical gist of a statement and can be represented by various sentences (a sentence being a particular sequence of words) such as 'the cat is on the mat', 'there is a mat directly beneath the cat', 'un chat est sur le tapis'. An utterance is the use of a particular sentence on a particular occasion: 'Go directly to jail' has a different meaning in a law court and in a game of Monopoly.

Fifthly, and finally, one needs to be aware of the relations which the concepts represented by words and phrases can have with one another: these include synonymy, antonymy, hyponomy, meronymy (or partonymy) and endonymy, for example. Synonymy (or paraphrase in the case of sentences) is apparent when one can substitute one term for another without altering the meaning of an utterance. Antonymy, or opposition of meaning, comes in different varieties: opposites (as *black* and *white*), different ends of the same scale (as *hot* and *cold*), incompatibility (as *Monday* and *Tuesday*) and differing viewpoints (as *buy* and *sell*). Hyponymy is a relation of sense inclusion: a hyponym (eg. *opposite* or *dog*) is more specific than its superordinate (*antonym* or *animal*). Meronymy is the relation of part to whole: arm, leg, and head to body, for example. Endonymy occurs where concepts all refer in their meaning to a given concept: thus *spaniel* or *dachshund* are not only hyponyms but are also endonyms of *dog*. But many other words other than breeds refer to dogs in their meaning: words like 'bark,' 'muzzle,' and "leash' also are what one might call "dog-words," that is to say, endonyms of *dog*.

This fifth point emphasises the inadequacy of the idea that our knowledge is a matter of 'associative' links between memory elements. We keep clear in our minds many sorts of links between concepts without confusion between them: we do not confuse the meaning *spaniel*, a type of *dog*, with the meanings *opposite to dog*, *same as dog*, *part of dog* and so on. Yet if all links were simply associational this sort of confusion should be possible.

The words that we use in physics teaching are commonly used in non-technical senses in everyday life. Some examples are 'force', 'energy', 'momentum', and 'power'. Students need to determine from their experience in the science classroom the extent to which the technical sense is related to the everyday: in particular they will need to decide if terms are ambiguous or polysemous. For example, is 'force' as used in everyday life possessed of a completely different meaning to that it has in physics (ambiguous), or is it the basically same core meaning with variations determined by the context (polysemous)?

Where technical terms are rarely used in everyday life (eg. 'vector', 'centripetal') there is still the question of how (if at all) the student arrives at an understanding of its meaning, and the extent to which its relations to other concepts is known.

As well, there are terms, like logical connectives, that are used in physics just as they are elsewhere (eg. 'thus', 'therefore' or 'implies'), but which nevertheless cause problems to the students (Gardner, 1974, 1977). These terms are central to logical argument, and in so far as understanding of physics requires following a logical argument (for instance, in deriving the direction of acceleration in circular motion) central to physics instruction. Thus the difficulty of these terms, and the logic they represent needs to be taken into account.

In order to deal with these issues it was necessary to adopt some tools used and developed by linguists. These will now be discussed.

Constructivism and concepts: the need for a Natural Semantic Metalanguage

From the previous sections one can see that there are a large number of issues to do with words, concepts and meanings which will impact upon students' understanding of science. In order to clarify these issues there is a need for a tool that will enable us to describe the meanings of concepts labelled by words in a clear and unambiguous way. Such a tool is provided by the Natural Semantic Metalanguage (NSM) which will be described below. However, as well as this practical need there is a subtler and more deeply embedded way in which a Natural Semantic Metalanguage is required by the theory of constructivism.

As mentioned above, a standard position in constructivist accounts of education is that the key to what a student can learn is what they have already learnt: 'The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly.' (Ausubel, 1968). There is, however, a theoretical consequence of this point. It follows, if one accepts this principle, that one can trace the process of learning in reverse: that which the "learner already knows" is built upon something in its turn. Given that we know what we have learnt, and that learning is always built upon what we know we face an infinite regress.

The standard response to this sort of prospect since at least the time of Plato has been to appeal to innate knowledge: there are some things that we simply know, without having first to learn them. Of course, there is controversy as to just what these ultimate backstops to our knowledge are.

An interesting approach to this problem has been taken by Wierzbicka and her colleagues who argue that whatever these innate ideas are, they must be common to all people. Hence they should be expressible in all languages. By looking for words that are used with exactly the same meaning in all languages Wierzbicka and her colleagues have been able to uncover a common core, of around sixty words (see Table 1). Wierzbicka and her colleagues assert that this set of words is adequate to define any other word (Goddard, 1998; Wierzbicka, 1985, 1996). They claim that they comprise a Natural Semantic Metalanguage, usually abbreviated to NSM, a phrase that requires some unpacking.

Table 1 Proposed NSM semantic primitives (after Wierzbicka 1996)

Substantives:	I, YOU, SOMEONE, PEOPLE/PERSON; SOMETHING/THING
Mental predicates:	THINK, KNOW, WANT, FEEL, SEE, HEAR
Speech:	SAY, WORD
Actions, events, and movement:	DO, HAPPEN, MOVE
Existence:	THERE Is
Life:	LIVE, DIE
Determiners:	THIS, THE SAME, OTHER
Quantifiers:	ONE, TWO, SOME, ALL, MANY/MUCH
Evaluators:	GOOD, BAD
Descriptors:	BIG, SMALL
Time:	WHEN/TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME
Space:	WHERE/PLACE, HERE, ABOVE, BELOW; FAR, NEAR; SIDE, INSIDE
Interclausal linkers:	BECAUSE, IF
Clause operators:	NOT, MAYBE
Metapredicate:	CAN
Intensifier, Augmentor:	VERY, MORE
Taxonomy, partonomy:	KIND OF, PART OF
Similarity:	LIKE

When we talk about a language we need to distinguish between the language as topic and the language being used to describe that language: the second is termed a 'Metalanguage'. So for example, in a textbook used for learning Japanese in Australian schools, the language being described would be Japanese and the Metalanguage would be English. Wierzbicka argues that the meaning of words in *any* language can be defined by using the NSM, hence it comprises a universal Metalanguage. In so far as Semantics is the study of meaning, and these fundamental words are found in naturally occurring languages this set of words (loosely speaking: in English some are short phrases) is termed the Natural Semantic Metalanguage (NSM).

Obviously, an empirical project to identify the ultimate constituents of meaning is a major undertaking. It is a linguist's equivalent of the biologist's human genome project, or the physicist's search for the ultimate constituents of matter, and it is being pursued with far fewer resources. In so far as the particular list of words to be employed is an empirical hypothesis, the NSM is subject to falsification: the words currently thought to be in it may prove in fact not to be innate but definable in terms of other words, the NSM words may not, in fact, be used in all languages, and they may not be sufficient to define all other words. Nevertheless, the NSM has proved to be a flexible and useful tool for constructing definitions, and will be used in this thesis as a tool to define the senses which are already known to students from their lifeworld experience, and which form the basis upon which the technical concepts of physics are developed.

The Natural Semantic Metalanguage is one tool adopted from the work of linguists that will be used in this thesis. Another area related to linguistics that will be important for this thesis is discourse analysis.

Discourse Analysis

Discourse can be looked at in two aspects, externally and internally. Externally, one can examine the actual speech events. Internally, one can consider the processes which go on in the minds of the participants. These two aspects will be looked at in turn.

Studies of the external aspects of student discussions have a lengthy history. As long ago as the early 1970s Barnes and his co-workers had established the importance of spoken interaction between students in learning (Barnes, 1975; Barnes, Britton, Rosen, & L.A.T.E., 1971; Barnes & Todd, 1977). This work looked at the way small group discussions in classrooms were structured, and showed that student discussions were sometimes more sophisticated than their teachers expected. Whether or not this work was an influence on them, constructivists have been leaders in incorporating class discussions, small group discussions and individual interviews into the science classroom (eg. Alexopoulou & Driver, 1996, 1997).

Outside of the classroom, there has also been a great deal of interest in the structure of informal discourse. For example, a five-part pattern to spoken narrative talk in groups has been identified by Labov and Waletzky (quoted in Chafe, 1995, p. 128). They assert that when people tell a story to a group, their narration falls into five stages. Labov and Waletzky called these orientation, complication, climax, denouement, and coda. Patterns that show distinct stages have also been reported in recent work on the functional structure of science text books (Unsworth, 2001): these were orientation, implication sequences, and closure. This work will be used to cast light on the structure of student discussions in Chapters Five and Six of this thesis.

The internal processes which go on inside the minds of participants in class discussions are not accessible as easily as a taped and transcribed discussion. It is in this area that the painstaking work of people like Gernsbacher can provide insights and reinforcement to constructivist science researchers.

Gernsbacher (1990) argued that the process of reading comprehension was a matter of structure building. From a series of ingeniously designed experiments (involving the measurement of reaction times to vocabulary test items during the reading of text) she was able to argue that the reader was actively constructing a mental model of the situation described in the text.

The familiar U-shaped curve for recall (where the first and last items presented are recalled better than items in the middle) could be explained on the basis that the first item was used as the foundation upon which the mental model had to be built, and the last item was the most recently actively attached to the mental construct.

By measuring reaction times before and after 'episode boundaries' she was able to argue that at the end of an episode one structure is completed and another was begun.

At the opposite end of the discourse analysis spectrum one finds theorists like Chafe (1980; 1995; 1976) who has no interest in the psycholinguistic techniques of Gernsbacher, but who instead analyses the use of language in naturally occurring discourse. He notes when people pause and where there is stress, notes the distribution of the use of 'the' as opposed to 'a', notes the use of deictic (words like 'here', 'I', 'now' whose meaning depends on context), and the use of viewpoint words (words like 'come' versus 'go', 'buy' versus 'sell'), and the times when people use descriptive phrases like 'the guy with the ugly face who we saw in the city', or names like 'Jeff', or pronouns like 'he'.

From this evidence he arrives at conclusions which are compatible with those of Gernsbacher.

Chafe argues that as we talk we are constantly building a mental structure. However, as we build it cognitive resource limitations restrict us to adding on just one linked item of information at a time. This new information is represented in speech by a short portion of an utterance. This short portion is the 'tone group,' a group of words spoken together with a single accented content word and usually bounded on either side by a brief pause (many others have contributed to the idea of the tone group, and a recent review of work in this area can be found in von Heusinger, 1999). The tone group is identifiable in discourse both by its intonational contour and by its informational content. Its intonational contour is characterised by the pitch of the speaker's voice quickly rising to a higher level and then more gradually dropping back to a lower pitch. Its informational structure is characterised by the speaker starting with given information and proceeding to a new unit of information. As these tone groups are heard the listener focuses attention on the part of the mental structure corresponding to the given information and integrates the new information.

This process is reflected in speech by the use of particular types of referring expressions, or anaphora, which are chosen on the basis of whether the information is given, or new. Once the speaker judges that the listener knows what one is talking about (that is to say, has built a mental model) some things can be easily identified within the discussion (i.e. they are already part of the structure, or 'given') and these can be referred to with pronouns or with the definite article. Other items, which are being added to the structure, will be introduced to the hearer by the speaker using the indefinite article. On the other hand, when one wishes to recall a referent from long term storage a name or descriptive phrase is needed. This can be illustrated by an invented example:

> A man walked into that long established department store on Bourke St, well known for its Christmas window displays, because of a sudden downpour. He stayed inside it until the rain stopped. Although it sold men's accessories he didn't buy an umbrella.

When referents are introduced ("a man", "a sudden downpour" or "an umbrella") an indefinite article is used to mark their status as new. Subsequent references are either pronouns ("he," or "it") or take the definite article ("the rain") to indicate that they are given. Things expected to be known to the listener, but needing to be retrieved from long term memory, are referred to by names ("Bourke St") or descriptive phrases ("that long established department store on Bourke St, well known for its Christmas window displays").

Chafe (1995) accumulates evidence based on the study of the distribution of anaphora within spoken and written language to argue that there are three levels of activation for mental structures: consciously attended to (active), accessible (semi-active), and stored in long term memory. His emphasis on the importance of conscious attention to structure building reinforces the importance accorded to metacognition in some constructivist writings (White, 1988b).

Chafe's picture of the building of mental structures during the process of comprehension is one of the human mind feverishly building, attaching, abandoning, reattaching, retrieving and filing away various mental structures at a rapid rate. It is in some respects similar to the Heideggerean picture of the mind constantly and insatiably interpreting its world (Heidegger, 1962, pp.182-195).

However, this is a very different picture to what one would expect from the reading of constructivist literature. There the emphasis is on the need for long term exposure to ideas before conceptual change can occur.

It will be argued in this thesis that this apparent contradiction is a result of the lack of distinction between the various layers of knowledge which is typical in constructivist literature. Where the proper distinctions are made there is no contradiction. All three approaches converge on a picture of the activity of the mind of the learner; the speed of this activity is simply dependent on its scale.

Knowledge Representation

Both Gernsbacher and Chafe, working within different traditions, conclude that the process of comprehension involves the building of mental structures. The question therefore arises as to how one should represent these mental structures.

Given that how we picture the representation of knowledge in the mind of the learner determines the questions researchers pose and the interpretations given to the answers found, it is important to examine the views which have been put forward in the literature. Here one encounters a variety of proposals: some from psychologists, some from philosophers, educationalists, artificial intelligence workers and others. Of these, the early behaviourist notion that all our knowledge was simply an accumulation associations of stimuli with responses (described in Gardner, 1987, pp. 15-19) will be briefly described first as it claims that there is no need for any structured knowledge representation. In the next section the evidence that memory does indeed require that knowledge representation be structured will be reviewed. After this, the main modern theories of knowledge representation will be discussed.

Behaviourist Models

Behaviourist proposals essentially bypass the problem of knowledge representation. The more radical behaviourists argued that there was no need to posit any intermediary knowledge representation. Learning was simply a matter of setting up a direct association between the correct response and the stimulus. For such theorists, for example, test paper questions and student answers to them are seen as merely examples of stimuli and responses. However, even though the importance of the idea of association as a component of knowledge is widely accepted (eg. White, 1988, pp. 41-48), and some attempts to map word associations amongst words used in Newtonian mechanics have been made (Gunstone, 1979, 1980), and even though the idea of associations is closely linked to that of spreading activation in semantic networks, nevertheless behaviourist theory is unsatisfactory. The fundamental problem with behaviourism is that it cannot account for complex phenomena because of its over-reliance on the single mechanism of stimulus and response as an explanation for all phenomena.

Memory

The behaviourists argued that the memory was a repository of associations of stimuli-and-responses. This, it will be argued, is implausible, based firstly on the evidence of the importance of meaning in learning (as investigated by Bartlett, and as reflected in the practical tradition of the art of memory) and secondly on the evidence for multiple systems of memory. The arguments adduced in this section are perhaps an exercise in beating a dead horse: the demolition of behaviourist doctrines began in the 1950s, spearheaded by writers like Chomsky, Miller and Simon, and their work is well known (Gardner, 1987, pp. 10-14).

The first scientific research on memory was carried out in the late nineteenth century, by Ebbinghaus (Baddeley, 1997). Because this was devoted to the rote-learning of nonsense syllables it fitted well with the early twentieth century doctrines of behaviourism. Even so, work done in the twenties and thirties such as that of F. C. Bartlett (1932) on memory for meaningful material (for example narratives as opposed to lists of nonsense syllables) soon uncovered phenomena which were not easily accounted for by behaviourist theories.

Firstly, Bartlett observed that when learning unfamiliar material (amongst other things, Bartlett used Amerindian folk tales) subjects typically reorganised it mentally into a format which held meaning for them, and retained this format during subsequent recalls over many years, often distorting the material in the process. The distortions were such that unfamiliar material could be assimilated to what Bartlett called a 'schema'.

(The idea of schema has since been elaborated upon by later theorists (see eg. Howard, 1987), and is particularly relevant to science education in so far as Stavy and Tirosh (1996; 2000) have provided evidence of the widespread influence of schemata, which they call intuitive rules. According to Stavy and Tirosh, students naturally work with an intuitive rule: more of one quantity will lead to more of another quantity, a schema which they call "More A – More B." Thus, for example, it will be argued in Chapter Nine, an important difficulty that students must overcome, before they can understand Newton's second law, is that they must somehow overcome this natural tendency to interpret the equation $F = ma$ as an example of the schema "More A – More B." Stavy and Tirosh (2000, pp. 42-63) also show that across many different subject areas, students appeal to a schema which suggests that whenever quantity A is the same in two instances, then quantity B will be the same also.)

Furthermore, in addition to calling attention to the phenomenon of schema, Bartlett also showed that recall of meaningful material was a matter of reconstruction rather than (stimulus-and-response-style-) reproduction, with much of the (re)construction occurring at the time of first exposure to the material (1932 p. 93). This is directly relevant to the topic of this thesis: the mislearning of unfamiliar material like Newton's laws may be due to the way students construe what is said into terms with which they are familiar.

Secondly, Bartlett noted that certain salient details became fixed in successive reproductions of the story, progressively taking an earlier place in the retelling. He obtained similar results with subjects who were asked to draw reproductions of abstract symbols they had only seen for a short time. Certain details tended to become the readiest means for identifying the symbols (1932, p. 107). This raises the question of whether there are particular details in, say, the topic of the mechanics which are preferentially retained, and which serve to organise the memory for what is learnt.

Bartlett's work was important for the scientific study of memory because it raised questions with which the then dominant behaviourist paradigm in psychology could not cope. But his work was not the only relevant stream of knowledge outside this paradigm.

While Bartlett originated the empirical investigation of meaningful learning, there has of course been the far older practical tradition of advice on remembering. This tradition, which has been in existence since at least 500 BC, is detailed in Yates (1966) and Carruthers (1990), and is exemplified by Lorayne (1986) and numerous other books with titles like *Learn How to Study*. The amount of time students put into their study of topics like mechanics, their preferred methods for the learning of this topic and how conscious students are of alternative methods and many other such questions about the conscious knowledge of and deployment of different methods of study by students is often termed 'metacognition' (White, 1988a). Some writers (eg. de Jong, 1996) have suggested that directing students to consciously and explicitly address the conditions for conceptual change may have an impact upon their learning of physics concepts, for example.

A final point to make on the topic of memory is that it is not a single homogeneous system. Modern accounts of memory (eg. Baddeley, 1997; Gruneberg & Morris, 1994; Morris & Gruneberg, 1992; Norman, 1976; Parkin, 1993) detail the differences between working and long term memory, and the distinctions which are made within long term memory. These distinctions include those between episodic, autobiographical, semantic, procedural, implicit and explicit memory. Adding to this complexity, separate stores have been hypothesised for memory for faces, for words denoting abstract and concrete concepts, for proper names and so on. It is hard to see how the single mechanism of stimulus and response could account for the multiple systems of memory which have been discovered.

Work in cognitive science

Abandoning behaviourist theory, and its attempt to account for learning without positing any mental constructs beyond stimulus-and-response associations, modern cognitive science has attempted the task of describing knowledge structures. A number of methods have been developed for doing this, including semantic networks (eg. Sowa, 1984), production rules (eg. Sell, 1985, 33 ff.), propositional representations (eg. Hayes, 1979) and connectionist models (eg. Dayhoff, 1990). These will be discussed in turn.

Node

Semantic Networks

A semantic network is a formal method for describing knowledge which gives rise to notations which are similar to what are often called concept maps: see Figure 2.

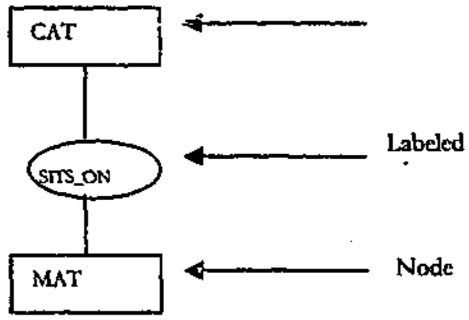


Figure 2: Simple semantic network.

Semantic networks (externally) represent the way we (internally) represent knowledge by nodes representing things connected by labelled lines representing verbs and other connectors. This is implemented in a formalism involving bipartite directed graphs: that is to say, by labelled nodes connected by labelled links (also called edges). In so far as it is usual to label the nodes of semantic networks with nominals (eg. CAT, MAT) and the edges (i.e. the joining lines) with verbs (SITS_ON, IS_A) one can read off the knowledge represented by starting from one node, reading its label, proceeding along an edge, reading its label, and arriving at another node and reading its label in turn: CAT SITS_ON MAT, for example. Although the basic idea is clear enough, a variety of technical difficulties make any implementation of semantic networks on computers by artificial intelligence researchers formidably complex. The necessity of including prepositions and articles as part of the edge labels in SITS_ON or IS_A gives some idea of the sorts of difficulties that are encountered. Transitive verbs like 'is' or 'sits' are easily represented by an edge connecting a 'doer' or agent with a 'done to' or patient. But what of intransitives like 'sleeps': only one node is attached to the edge to give CAT SLEEPS. Is the node an agent or a patient? Ditransitives like 'gives' create more trouble: three things need linking by edges: giver, 'givee' and given. And so on. Sowa (1984) gives a particularly thorough and clear introduction to semantic network formalism. Way (1991) shows how semantic networks can capture some of the ability of people to use their knowledge metaphorically.

Most writers are content to leave the details of the formalism to the experts and use the idea of a semantic network itself as a fruitful metaphor or image for our interlocked knowledge of the world. For example Chafe (1980) offers the metaphor of a searchlight beam playing across the semantic network to illustrate the idea of 'priming', the process by which the activation of a particular concept activates associated concepts. (So that for example reading or hearing the word 'bat' activates the words 'vampire' and 'cricket' because these are associated with bats in different ways.) The idea is that 'nearby' concepts are 'illuminated' at the same time as the 'target' concept by the 'spotlight' of attention. (Priming, and spreading activation models in general, will be discussed further below.) As a metaphor, the 'semantic network' has proved useful. However, as a formalism that aims to capture and make explicit human knowledge in machine-readable form, semantic networks are subject to the same (hermeneutic) strictures as other formalisms.

Production Rules

While semantic networks tend to be esoteric, production rules, which are used extensively in the type of computer programs known as 'expert systems', are individually very easy to understand. Production rules consist of sets of preconditions and actions to be taken in response (Sell, 1985). In many ways they are a generalisation of behaviourist stimulus and response theories. Expert systems use them to achieve their aim of representing the often tacit knowledge of experts in a given field of knowledge, for example diagnosis of diseases, oil prospecting and so on. An expert system might for example contain the production rule: 'If the patient has spots AND the patient has a fever AND the patient is itchy THEN hypothesise measles.' As well as in expert systems, production rules have been used notably by the early Chomsky (1965) in characterising the deep structure of sentences (Rules are defined, such as S → NP VP; NP → det N; VP →V PP; PP → prep NP: prep → 'on'; det → 'the'; V → 'is'; N → 'cat'; N → 'mat'. Repeated application of these production rules leads from S to, eventually, 'The cat is on the mat.') Such representations of expert knowledge and of grammar have proven to be useful in certain limited domains, but are widely acknowledged to be 'brittle'. That is to say, a system of production rules designed to deal with diseases will give completely erroneous responses if it is presented with symptoms of a broken limb, or if there are complicating factors such as sunburn present. Likewise, Chomsky-style grammars are regularly used in the design of compilers for computer programming 'languages': however, the slightest error, such as a comma in the wrong place can cause massive and incomprehensible (to the average human) errors. Human reasoning is characterised by its robustness and only gradually loses accuracy as the situation gets further from that which is the human expert's field of expertise. There seems to be little reason, therefore, to believe that production rules are adequate, by themselves, to give a representation of a student's knowledge of physics.

Propositional Representations

Propositional representations, by contrast, seem to be the most popular way of describing student knowledge. One reads for example that students believe that motion implies force (Galili & Bar, 1992), that circular motion causes an outward directed force (Galili & Bar, 1992; Gardner, 1984; Gardner, 1982; Viennot, 1979), and so on. However, such useful representations of some small portion of a student's knowledge by a proposition (i.e. the logical content expressed by a sentence) must be clearly distinguished from the thesis that all knowledge is held in propositional form. That is to say: from the thesis that whatever is known can be characterised by a database of facts of greater or lesser generality. ('Cats meow'; 'Dogs bark'; 'Cats and dogs can be pets'; 'Cats and dogs are mammals'; '$F = m\,a$'...) Objections to such a thesis are plentiful: it does not account for such knowledge as the ability to recognise a face, ride a bicycle, speak a language or understand a joke for example. Indeed, part of the problem with physics instruction has been that it has been content with providing students with propositional knowledge such as that represented by '$F = ma$', without ensuring students' full understanding. Dreyfus (1993) spent much of his book following up the consequences of the notion that propositional representation of knowledge is inadequate, using arguments gleaned from Wittgenstein and Heidegger amongst others. This has not stopped computer scientists from attempting (as yet unsuccessfully) to achieve a representation of our physical knowledge of the world in this format (see for example Forbus, 1988; Hayes, 1979).

Although the idea that what we know can be fully characterised by a database of propositions is unsatisfactory, it does correspond fairly well with the 'folk theories' of 'transmissive' learning: bits of knowledge going from teacher to student. As many people tend to hold this view uncritically, educationalists should keep it in mind too. Biggs (1987) details the extent to which such a surface approach (rote learning of details) is present in the student population, and the extent to which such an approach depends on student motivation and course assessment. Again, the issue of the significance of affective structures for students arises.

As well, the precision with which models of knowledge can be constructed using a propositional representation, and the mathematical formalisms and semantics which they enable one to use (Bach, 1989), mean that a propositional representation of knowledge can be a very useful idealisation for some purposes, or even a suggestive metaphor. Thus problems with implementations of propositional databases have led to researchers' emphasis on non-monotonic reasoning (whereby propositions can be withdrawn from the database when they prove to be false, enabling such things as defaults - people assume that if something is a bird then it can fly, unless there is evidence to the contrary), partitioned representations (allowing one to agree that Sherlock Holmes was a detective, and also to agree that Sherlock Holmes never existed), and so on (Dinsmore, 1991). These developments are highly relevant to physics education where students assume that objects will slow down unless there is evidence to the contrary, and where they may agree that $F = ma$ in the context of textbook physics problems, but then agree that motion implies force in examples drawn from everyday life.

Connectionist Models

While propositional representations of knowledge are at least partially inspired by everyday ideas, connectionist models of knowledge representation were originally inspired by neurologists' investigations into the workings of the human brain. In the brain, as elsewhere in the nervous system, individual neurons are either activated or suppressed by other connected neurons, and in turn activate or suppress others again. The level of activation of neurons can be measured by the rate of firing (Dayhoff, 1990, pp. 1-36). While it is true that connectionist models were originally inspired by neurology, they have changed a great deal from their simple beginnings. Connectionist models, for example, do not have any equivalence to the effect hormones have on neurons, nor has the mechanism of 'back propagation' (a method of adjusting levels of activation in response to feedback) as used in connectionist models been observed in the brain. (Dinsmore, 1991, p.25).

Nevertheless, by partially mimicking the responses of neurons, computers can be programmed to perform useful tasks, and even to 'learn'. However, the type of learning involved requires hundreds of thousands of highly organised presentations of the material to be learnt, and is quite different from what we would normally think of as learning (Dinsmore, 1991, p.26). In spite of these differences, the ideas of 'activation' and 'suppression' have proven useful in describing language comprehension (Gernsbacher, 1990) which is a central part of human learning. As well, connectionist models perform surprisingly well at tasks like pattern recognition, characteristically human abilities that have defied other approaches.

In one way this is because of the holistic approach of connectionism, which differs from all the other approaches in so far as they all particularise knowledge. In a connectionist network, 'knowledge' essentially resides in the strengths of the interconnections between the 'neurons', and so is effectively spread throughout the network. This means, however, that there is nothing in the network that one can 'point' to and identify as a particular piece of knowledge: all things 'known' by the network are mixed indiscriminately together. For the process of categorisation this causes no problems: the network receives a certain input and recognises John as opposed to Mary, or an act of seeing as opposed to an act of eating. But this does cause problems when the network is faced with even the simplest compositional representations like 'John sees Mary'. With everything mixed in together, it is hard to make sure that things are linked together in the correct way: thus connectionism suffers from much the same sort of problems as, say, White's theory of concepts and learning, discussed above (White, 1988). For this and other reasons it seems likely that connectionism must be supplemented by other methods of knowledge representation (Dinsmore, 1991, pp. 28-30).

ACT-R

A theory of cognition, referred to as ACT-R in its most recent version, which has had some success in integrating these various methods of knowledge representation is due to Anderson (J. R. Anderson, 1996). In this theory it is assumed that a production rules operate upon declarative memory, and the theory emphasises the role of activation based processing in relating these two systems. A brief overview of the theory is provided in a co-authored article (Anderson, Bothell, Lebiere, & Matessa, 1998), some extracts from which are given below (ellipsis added to indicate omissions) :

> Different traces in declarative memory have different levels of activation which determine their rates and probabilities of being processed by production rules.
>
> ...
>
> According to ACT-R, procedural knowledge, such as mathematical problem-solving skills, is represented by production rules which coordinate the retrieval of declarative information ... for purposes of problem solving.
>
> ...
>
> *Activation*
>
> Activation of declarative structures has always been an important concept in the ACT theories. Basically activation determines how available information will be.

Although the level of sophistication in the modelling of behaviour which characterises this theory will not be attempted here, an account of cognition which is like that of the ACT-R theory will be assumed.

Overview

When, later in this thesis, an attempt is made to represent the knowledge students have of two of the central concepts of Newtonian mechanics, *acceleration* and *force*, all of the above forms of representation will be utilised. Semantic networks will be used to characterise dictionary definitions, partitioned representations, or episodes, will be shown to exist in transcripts of student discussions, activation and suppression in connectionist networks serve as a model for Gernsbacher's (1990) Structure Building Framework and this is used as a model for the processes involved in utilising concepts in problem solving, and production rules serve as models for the triggering of particular conceptions by particular contexts, as is argued in the ACT-R theory of cognition.

However, before proceeding with these investigations, the literature related to the two concepts *acceleration* and *force* will first be examined.

Acceleration

Although there are a number of articles about strategies for the teaching of acceleration (eg. Bunker, 1991; Flores & Turner, 2001; Kraft & Motz, 1995; Newburgh, 1998) there are only a few that claim to investigate students' concepts of acceleration.

The first of these was written by Trowbridge and McDermott (1981). In it the authors confidently assert that

> The criterion for assessing understanding of a kinematical concept is the ability to apply it successfully in interpreting simple motions of real objects.

While this may be one method for assessing students' level of understanding, it is difficult to accept that it is the (only) way to do so. Furthermore, the ability to *use* a concept in a particular context is not equivalent to the *understanding* of that concept. That the students had an understanding of acceleration is indicated by their success in an initial task "which presented no challenge to the students interviewed" as a "primitive notion of acceleration as speeding up seemed to be adequate." This is as close to a characterisation of the student concept of acceleration as the authors get. While their data could have been used to analyse student usages of the word 'acceleration' with the aim of characterising its meaning in the same way that lexicographers examine usages of a word when trying to define its shades of meanings, they did not in fact use it in this way. What they did in fact achieve in their paper was a survey of the various procedures by which students try to answer the problems that were posed by the authors. In the most interesting of these, students were asked to observe the motions of two balls rolling down inclined planes and to use their observations to determine whether or not the accelerations of the two balls were numerically equal. The authors were primarily interested in the extent to which students utilised the notion of measuring acceleration by direct use of the ratio $\frac{\Delta v}{\Delta t}$. Other methods, some of which were appropriate, were not accepted as demonstrating full understanding. For example substitution into a kinematic formula such as $s = ut + \frac{1}{2} a t^2$, utilisation of dynamics formulae such as acceleration down an inclined plane = $g \sin \theta$, or observations that the same velocity change occurred over different distances were all disallowed. Similarly, students who used strategies that would work in some contexts but not in the ones presented were assessed as not understanding the concept. Indeed students who argued that the ball which travelled the greater distance had the higher acceleration (which would be correct if both balls had started from rest and travelled for the same amount of time) were assessed as exhibiting "confusion between position and acceleration." Others, using different procedures, which also could have worked in simpler contexts, were said to exhibit "confusion between velocity and acceleration." Such claims are difficult to accept: if students truly confused, say, the first pair of these concepts they would have had to believe that two objects standing still in different positions had different accelerations. It is surely more likely that the students were simply applying intuitive rules (Stavy & Tirosh, 1996, 2000) such as "more of A – more of B," in lieu of any better way of coping with the demands of the task.

The authors hedge their statements somewhat:

> In referring to a confusion, we are using this word in a restricted sense. It is not intended to indicate mistaking of one fully developed concept for another but rather to characterize thinking in terms of nondifferentiated protoconcepts.

However, this hardly helps to avoid the absurd conclusions which follow from the idea of student inability to distinguish between position and acceleration. Whether or not these are "nondifferentiated protoconcepts," students who could not distinguish them would still be unable to tell the difference between objects with different positions and objects with different accelerations. Furthermore, it is hard to see what benefit is gained by the invention of a new category of mental entity, the "protoconcept." How, one might ask does this differ from a concept?

In spite of these criticisms, the Trowbridge and McDermott paper did uncover a number of interesting phenomena, and led to more work in this area.

A later paper by Jones (1983), referred to the paper by Trowbridge and McDermott described above, and aimed to extend its coverage of age groups (university level students) to "identify the concepts of speed, velocity and acceleration held by students in the 11-16 years old range."

This was done by presenting students with cards illustrating various events and asking them, as part of a sequence of questions "Has it got any acceleration? Why do you say that?"

The method applied here was clearly related to the task of finding the denotation of the concept *acceleration*, that is to say, the class of events which students classified as being examples of *acceleration*. (By contrast, the paper by Trowbridge and McDermott had been devoted to characterising the procedures by which students solved certain problems involving acceleration, a focus only tangentially related to the concept itself.) However, the limited range of evidence obtained in the research upon which Jones' article was based meant that its conclusions were quite rightly limited in scope:

> From the interviews it became obvious that the majority of students realized that acceleration involved the idea of a change in speed but the relationship between the two quantities was not the scientific definition (i.e., acceleration = change in speed/change in time).

One could quibble with the last part of this statement, which does not adequately characterise the scientific definition $a = \frac{dv}{dt}$, but the basic thrust of this conclusion is clearly correct. It is, however, very vague. One wants to know: in what way is the idea of a change of speed "involved" in the concept *acceleration*? There is clearly room for further work.

However, the only further work in this area that I have been able to locate was carried out by a number of researchers working in the phenomenographic subfield of educational research, who applied their approach to investigating the concept of *acceleration* (Dall'Alba et al., 1993). In this investigation a number of students were asked to complete the task shown in Figure 3 (Figure 1 in the original article). Students responses (which were encouraged by "non-directive questions, such as 'Could you explain that further?'") were recorded, transcribed and subjected to "rigorous phenomenographic analysis." This involved one member of the team of authors carefully reading the transcripts and suggesting a classification of the ways students understood acceleration in the problem, followed by an iterative process of fine tuning of the classifications by the team as a whole.

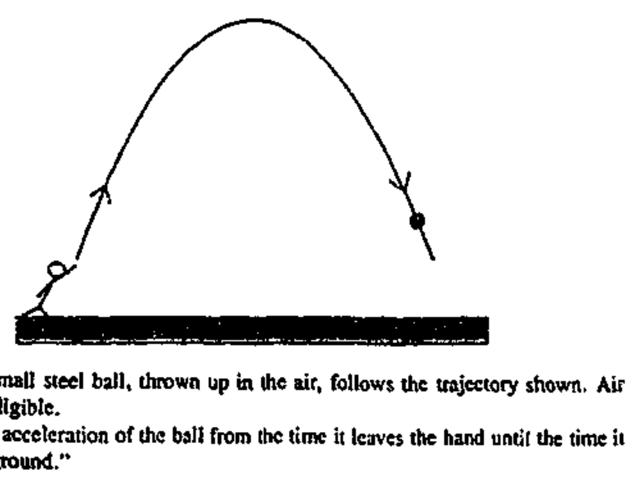


Figure 3: Diagram used to prompt discussion of acceleration (Dall'Alba, Walsh et al. 1993).

It was assumed that any particular student's understanding of this concept fell under a single classification. Indeed the authors emphasise that classification was not made "by matching particular statements with specific categories" but rather by "taking the transcript as a whole." This decision, of course, meant that any variant usages of 'acceleration' that occurred within the transcript would necessarily have to be ironed out in the process of analysis.

Like the study of Trowbridge and McDermott (1981), the authors (Dall'Alba et al., 1993, p. 627) already had a notion of what the student understanding of *acceleration* should be, and were interested in seeing how closely the students approached this ideal:

> ...the necessary elements of an adequate understanding of the acceleration of the ball include instantaneous rate of change of vertical velocity and the causal relation with gravity.

The resulting classifications were arranged from best to worst: the best included both these points, the two next best levels included one or the other but not both, and then there followed three other classifications where various confusions were evident. In these three remaining categories students clearly ascribed force-like attributes to *acceleration*. However, given the context, where gravitational acceleration and the force due to the gravitational field are both often indiscriminately called 'gravity' this should not be surprising.

While this work calls attention to some ways in which confusions can arise in the qualitative understanding of projectile motion, as an investigation into the meaning of acceleration it is less successful. Firstly, the requirement that acceleration should be understood as the rate of change of vertical velocity is unduly specific. This is merely one way of calculating a measure of the acceleration, not its meaning. Secondly, the requirement that students understand the "causal relation" of acceleration with gravity introduces the more complex ideas of Newton's second law, $F = ma$, into the definition of the simpler concept *acceleration.* Apart from the inherent problem in defining a simpler term by the use of more complex ideas, there is the objection that scientists and non-scientists alike had used the term acceleration well before Newton's time (the *OED* gives its earliest examples from the 1530s). This is not to argue that it is not better for students to understand the causal relation between the gravitational force and the acceleration of a projectile: clearly it is good for students to do so. However, this should not be confused with the issue of student understanding of *acceleration* per se. The understanding of acceleration is possible independently of the understanding of *force.*

Force

The concept *force* has been examined both within the field of science education and within the field of linguistics. Both investigations commenced in the late 1970s independently of each other, and there has been little, if any, contact between the two fields. This is evident when one examines the citations in the literature – neither group cites the other. The science education research will be discussed first, and then the linguistic research.

Science education and the concept of force

Warren (1979) was amongst the first to report on the serious conceptual misunderstandings which were evident in students' accounts of the forces acting in different situations. Some situations in which Warren demonstrated that there were incorrect understandings included the case of motion along a curved path, and the case of surface tension. In the case of motion along a curved path, an example being the motion of a pendulum bob, students did not understand that the direction of the net force must point at an angle to the path followed by the object: the most common error was to draw the force in the direction of motion. In the case of surface tension many students (and many textbooks!) showed a net inward force acting on molecules at the surface of a liquid: if this were the case then the surface of the liquid would, of course, accelerate inward. While Warren gathered convincing evidence that students lacked understanding of the concept of *Newtonian force*, he did not, however, investigate the concept of *force* which they held.

Such an investigation could be carried out in two different ways. One way would be to investigate the reference (or denotation) of the concept, another would be to investigate its sense (see p. 36 above). While there has been little or no interaction between the fields of science education and semantics, a number of researchers worked on studies which had results which could be interpreted in this way.

One way of investigating the denotation of *force* is to provide students with depictions of a range of situations and ask if any forces were present (eg. Gilbert, Watts, & Osborne, 1985). A variation on this is to ask them to mark forces on diagrams of various situations. When investigations that included these kinds of tasks were carried out it was soon found that situations in which an object was not moving were assessed by many students as not showing forces, despite the constant presence of normal reaction forces by surfaces that support objects resting upon them (see eg. Minstrell, 1982). Furthermore, in situations where things were moving, a force was believed to be present and acting in the direction of the motion. These results have together been dubbed the 'motion implies force' misconception (see the review of research in Gunstone & Watts, 1985).

In addition to the work using these methods there has been a great deal of work done in the way students use the concept of *force* in solving problems. As much of this is relevant, however, to the understanding of Newton's laws rather than to the understanding of *force* per se, it will be discussed in a separate section below.

Another way of investigating the denotation of the students' notion of *force* is to ask whether particular types of forces accepted as such by physicists are also thought of as forces by students. An investigation into "the conception of force and motion of students aged between 10 and 15 years" (Twigger et al., 1994) reported the following data:

Table 2: What counts as a force (based on Table 1, Twigger et al., 1994)

Responses	Percentage frequency of response (N=36)
Force identified as a push	69%
Force identified as a pull	64%
Weight identified as a force	61%
Friction identified as a force	31%

As all of these would fall under the denotation of *Newtonian force*, and as it has already been noted that the normal reaction force which supports objects at rest on a surface is often not considered to be a force by students, these results together clearly indicate that the denotation of many students' concept of *force* differs from that used in physics.

Investigations which give insight into the sense, rather than the reference, of student conceptions of *force* have also been carried out. Hart (2002) asked students, working in pairs, to consider all the forces that they experienced on their way to school. Although the purpose of this was to "make connections between the students' experiences and the content of the unit that was to follow", it proved to be a more valuable exercise than expected. This was because it provided a corpus of usages of 'force' from which the students' senses of the word became apparent. (The same sort of collection of usages is done by lexicographers in defining the variant meanings of words for a dictionary entry.) Some of the responses were classified by Hart as metaphorical usages:

> "My mother forced me to get out of bed."
> "I forced my eyes open."

Others were straightforwardly Newtonian forces, such as a bat striking a ball, or pushing down the lever of a toaster (although the latter sparked an argument between partners as one of them did not regard such a gentle action as a force).

Further insights into students' meanings of force were prompted when, to emphasise the interactional nature of forces this terminology of an agent and a receiver of a force was introduced to the students. Students were then asked to identify the agent and the receiver of a number of forces, and to identify the effect of the force on the receiver. Hart writes:

> However, there were also many surprises. For some of the forces, students did not identify the effect of the force on the relevant receiver. Thus they said that the effect of the force from your hand on the tap was that the water came out (whereas the effect on the receiver – the tap – is that the tap begins to move); or that the effect from the hand on the steering wheel was that the car changed direction (rather than the steering wheel begins to move); or that the effect of the force from the bike on the ground was that the bike moved.

Puzzling results like those above, become easier to understand if one refers to the literature on *force* within linguistics.

Talmy's account of causation in terms of force dynamics

While it has become increasingly evident to science education researchers that grasping the Newtonian concept of force is not only essential, but also problematic for students, research carried out in linguistics has argued for the importance of the concept of physical force in understanding the patterning of language. According to Talmy (1976; 1985), physical force interactions are fundamental, and these are metaphorically extended to cases of psychological, social interactions.

Talmy (1976; 1985) pointed out a regularity in languages which appeared over and over again at various levels: not only in conjunctions, prepositions and other closed class elements, but also in open class lexical items (e.g. verbs, nouns), grammatical structures (e.g. in English, the modal system), and at higher levels too, in the rhetoric of persuasion for example, or our conceptualisation of interpersonal behaviour. In all of these areas it was plausible to interpret the patterning of language as being based on experience with physical forces.

In the first instance Talmy provides evidence for a pattern based on what he describes as our experience of force in space: a four way division which Talmy organised into two columns and two rows, see Table 3. The patterns he observed have been widely accepted (Johnson, 1987). Two roles were postulated: the agonist (the one with a goal) and the antagonist (the one who opposes the agonist). Two types of goal were postulated: action and inaction. Various possibilities can occur according to which is strongest: agonist or antagonist. The application of Talmy's scheme is illustrated with his sample sentences in the table below: the agonist is in bold, the antagonist in italic.

Table 3: Sample sentences from Talmy (1985).

	Agonist stronger	**Antagonist stronger**
Goal is action	**The ball** kept rolling despite *the stiff grass.*	**The log** kept lying on the incline because of *the ridge* there.
Goal is inaction	**The shed** kept standing despite *the gale wind* blowing.	**The ball** kept rolling because of *the wind* blowing on it.

In the sentences in the top row of Table 3 the idea is that the subject of the sentence is being talked about as if it had a goal: to keep moving. In the left hand sentence the ball achieves this goal in spite of the difficulties caused by the stiff grass – the ball is thought of as having more force than the stiff grass. In the right hand sentence the log is being talked about as if it had the goal of moving, but the difficulty caused by the ridge stops this from happening – the log is thought of as having less force than the ridge. A similar analysis is proposed for the sentences in the lower row, although in these, the goal is inaction. These four sentences are intended by Talmy as examples which explicate the senses of the phrase "to keep doing." Talmy's basic idea is that we develop our understanding of abstract notions by metaphorically extending our experience with the effects of force on objects in the natural world. In this experience, according to Talmy, the object with the greater force wins: it achieves its goal.

Two short quotations from Talmy give the flavour of his claims.

> As language treats the concept, an entity is taken to exert a force by virtue of having an intrinsic tendency toward manifesting it-- the force may be constant or temporary, but it is in any case not extrinsic. …
>
> A further concept in association with opposed forces is their relative strengths. As language treats this, the entity that is able to manifest its tendency at the expense of its opposer is the stronger. ... Finally, according to their relative strengths, the opposing force entities yield a resultant, an overt occurrence. As language schematizes it, this resultant is one either of action or of inaction, and it is assessed solely for the Agonist, the entity whose circumstance is at issue. (Talmy, 1985 p. 71)

In spite of his belief that these patterns are based on physical experience, Talmy is careful to hedge each claim by the phrase "as language schematizes it". This caution is not evident in some who take his work as the basis for their own speculations. For instance, Johnson (1987) clearly intends to identify Talmy's force with the force of physics:

> …force has a vector quality, a directionality... the force is exerted in one or more directions. As the baseball flies through the air, it traces a path that we can describe by a force vector, or series of vectors, leading from the pitcher to the catcher.

What Talmy describes as force "as schematized by language" bears a resemblance to the typical errors made by students.

1. Force is a property of an entity.

2. The resultant action depends only on the stronger force (so that resultant motion must be an indication of a force in the same direction).

What is written by Talmy offers a possible account of why students get misconceptions: they assimilate what they learn in physics to the linguistic schematization of force. (What is written by Johnson, on the other hand, exemplifies misconceptions rather than accounts for them.)

In so far as physical reality is not governed by forces which behave in the manner Talmy describes, it is of course impossible that we should have structured our language using such forces as the basis. What is it then, if not physical laws, which is present around the world so as to lead to this cross-linguistically consistent patterning of language (Talmy, 1985)? A possibility which will be offered here is that it is the internal processes of enhancement and suppression posited in the Structure Building Framework, or *SBF*, developed by Gernsbacher (1990) which provide this model. These have the requisite behaviour:

- particular conceptions are the source of enhancement or suppression of other conceptions.
- the conception which is ultimately active is alone present: other senses are simply no longer present to consciousness.

In so far as we are consciously aware of our thought processes, as we are in the more strategic forms of thought like problem solving, the process of considering and comparing factors could perhaps serve as the basis for our understanding of the social world around us, and thence to understanding the physical world.

Bearing this possibility in mind, it is interesting to see that Wierzbicka (1998) proposes for a NSM definition of the verb 'to force' one that is resolutely in terms of people, not things:.

Person X forced person Y to do Z (e.g. to apologise)

X wanted Y to do Z (a)
X knew that Y didn't want to do this (b)
X thought that if X did something to Y then Y would have to do Z (c)
because of this X did something to Y (d)
because of this Y had to do Z (e)
Y wouldn't have done z if X had not done this to Y (f)
when Y was doing Z, Y thought I don't want to do this. (g)

This definition clearly includes Talmy's agonist (X) and antagonist (Y), and the lines (f) and (g) clearly indicate the idea of resistance which was so important in the historical development of the concept of inertia (Franklin, 1976), as well as the concept of goal in line (a). I will argue that these are the elements in the lifeworld conception of force which can serve as the foundation upon which students build their technical understanding of Newtonian force.

This definition also resolves some of the puzzling results obtained by Hart (2002), discussed above on page 60. It becomes plain that "My mother forced me to get out of bed," far from being metaphorical, is a completely literal use of the word 'force.' It also explains the curious inability of students to identify the effect of the force from one's hand on a tap as the turning of the tap handle. The goal of exerting a force on a tap is to get water to flow, and it is this goal which is salient in the concept of *force*, as defined above, not the rotation of the tap handle. Similar explanations apply to the other two situations: the force of the hand on the steering wheel, and the force of the bike wheel on the ground. In each case the students identified the goal of the action as the effect, rather than the immediate effect on the steering wheel or the ground.

Summary

The results of research in linguistics and in science education into the meanings of the concepts *force* and *acceleration* are mutually illuminating.

Research into Newton's laws

As noted above, much of the science education research in mechanics that is aimed at elucidating conceptual understanding has not dealt with concepts of *force* and *acceleration* per se, but rather with their use in particular contexts. Particularly important contexts for these concepts occur when problems arise which require the use of Newton's three laws.

Because of the importance of these contexts for the use of these concepts the literature related to them will now be examined. The research into student understanding of Newton's laws will be examined beginning with Newton's first and second laws, and then discussing Newton's third law.

Newton's first and second laws

As was the case with the topic of acceleration much of the literature related to Newton's first (Williams, 2000) and second laws (see eg. Bryan & others, 1988; Domann, 1982; Geiger, 1968; Hood, 1992; Kikoin, 1979; Kurtze, 1994) is related to alternative demonstrations and methods for teaching them.

Newton's first law states that when no net force acts on an object at rest, it remains at rest, and that when no net force acts on an object which is moving it continues to move in a straight line at a constant speed.

Newton's second law, which is usually presented to students in the form of an equation, $F = ma$, states that the net force on an object is proportional to and in the same direction as the object's acceleration.

The key research findings which are relevant to Newton's first and second laws have already been mentioned: in spite of explicit teaching to the contrary many students cling to the view that a constant force is needed to keep an object moving (Gunstone & Watts, 1985). This clearly contradicts both laws. According to Newton's first law no net force is needed to keep an object moving. According to Newton's second law a constant net force applied to an object will cause it to accelerate, not move with a constant speed.

Newton's third law

People have noted confusions in the interpretation of Newton's third law by students for many years. Warren (1979) noted that students sometimes confused (fictional) centrifugal and (actual) centripetal forces with action/reaction pairs. Gardner (1981) argued that centrifugal forces should be avoided in elementary teaching of physics, for this amongst other reasons, and similar proposals continue to be made from time to time (de Jong, 1988; Lan, 2002; Smith, 1992).

Some questions intended to elucidate difficulties with Newton's third law were included in a survey by Watts and Zylbersztajn (1981), which was intended to investigate both the distribution of children's ideas about forces, and also the extent to which their teachers were aware of these ideas. Further work was done by Maloney (1984) who argued that most students were using one of five simple rules to answer questions involving Newton's third law. Terry and Jones (1986) used a set of seven questions to examine the understanding of Newton's third law by 16-year-old students. Brown (1989) reported on three studies investigating aspects of students' understanding of Newton's third law. Brown interpreted the results of the first two studies, interviews with pre-physics high school students, as indicating that before instruction students viewed forces as properties of objects. This view clearly causes students difficulty in their understanding of Newton's third law. This is because such a view would entail that students would have to interpret Newton's third law as stating that what one might call the 'force-property' of two objects was equal in size and opposite in direction. Such an interpretation would make Newton's third law essentially a mystery: 'After all,' students might think, 'why should the force-property in an interaction be the same when, say, the mass- or colour-properties were not?' In a later paper Brown (1992) investigated the dimension of 'activity' versus 'passivity' and its effect on student understanding of situations involving action/reaction pairs where one object is resting on another (as in diagram *d* in Figure 4). He makes the point that 'it may be necessary to show *how* certain examples are like prototypical cases (Rosch, 1973; Rosch & Mervis, 1975) rather than simply telling students which examples illustrate the application of a concept' (p. 18, emphasis and references in the original). The remainder of the paper is devoted to an exploration of bridging analogies as a method for motivating the extension of the category of situations where Newton's third law should be used. This work was elaborated in Brown (1994).

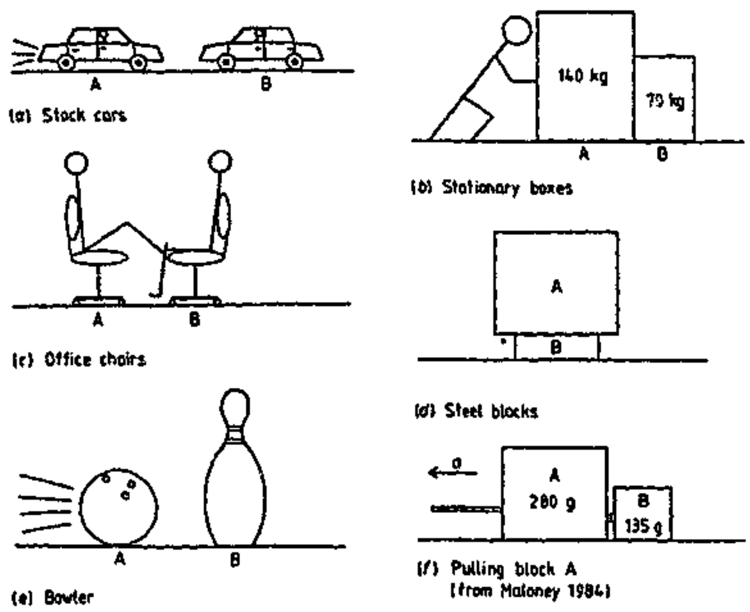


Figure 4: Diagrams from Brown (1989)

The data from the investigations mentioned above will be reanalysed in the light of prototype theory in Chapter Five of this thesis.

Another investigation by Bao and others (2000) of students' use of Newton's third law attempts to mathematically model the contributions of the dimensions of mass, velocity, activity, and acceleration but – like the mathematical formula obtained by Labov (1973) for the weighting of the features which determine whether an item is said to be a cup – this serves simply to describe the data they obtained from the responses of a particular group. It will be argued later that the key point is not the particular weightings for different features, but rather the fact that there are features of differing importance which together determine whether the idea of Newton's third law is activated in the minds of students.

The remaining work that has been published, related to Newton's third law, falls into two main categories: work by Gauld (1993; 1998) related to the historical development of the ideas of action and reaction, and suggestions for improving the teaching of Newton's third law by means of change of terminology (Hellingman, 1989, 1992; Styer, 2001). From the former one can gather that the prototypical situation about which scientists were concerned was that of the head-on elastic collision between two pendulum-bobs. However, the concerns and experiences of past scientists are quite different from those of modern day students, and while historical analogies are sometimes suggestive they may also be misleading (see e.g. Lythcott, 1985). From the second group of writings one is alerted to the differences between the everyday and Newtonian senses of the word force. However, suggestions such as that of Hellingman that the word 'force' be redefined are unlikely to be taken up, and if they were would likely compound the terminological confusion.

Chapter summary

This necessarily lengthy chapter has reviewed the key findings relevant to the research which is reported in the following chapters. The examination of the literatures of science education, psychology and linguistics has suggested that it is possible to fruitfully combine their conclusions and methods in order to come to a better understanding of the difficulties which students face in learning mechanics.

The following chapters report upon a series of studies which are aimed at elucidating the nature of the conceptual difficulties which challenge students in the study of mechanics. They are directed first to the preliminary step of clarifying the nature of the concept of *concept* in science education research, and then to characterising and investigating the internal structure of the concepts of *acceleration* and *force* which students bring with them to their classes in Physics. Finally, they will deal with the way these concepts are deployed in problem solving using Newton's laws.

This thesis has as its central argument that prototype structure is the key to conceptual development in mechanics.

This will be shown to be the case not only at the level of concepts, but also at the level of application of theories to problem situations. It will be argued that the ability of students to generalise Newton's laws requires that they recognise, hence categorise, situations according to their relevance to Newton's laws.

This account of conceptual development is inconsistent with the commonly accepted conceptual change theory put forward by Posner et al. and it will be argued that this theory is therefore incorrect.

Methodology

Before proceeding to the body of this thesis, and in order to help justify its conclusions, it may be worthwhile to provide an overview of the research methods, both quantitative and qualitative, which are drawn upon. This overview is perhaps all the more indicated in so far as these methods originate in different research perspectives: Science Education, Psychology, and Linguistics.

This variety of methods (and the variety of data to which they are applied) is in part a natural corollary to the conclusion arrived at in the literature review: that one can fruitfully combine these traditions in order to understand the difficulties which students face in learning mechanics. Considered in this light, the need for a mix of methods to investigate conceptual development in physics is due to the fact that this area of research does not have a neat fit with any one particular discipline. Any number of problems from applied technology could serve as an analogue: for example, just as solving the problems of speech synthesis draws upon the disciplines of Electronics, Acoustics and Linguistics, so here the solution to a problem in physics pedagogy draws upon a variety of related disciplines.

However, the mix of methods also reflects a larger hermeneutic ideal: that in so far as one seeks understanding one should use a variety of approaches. If they all point to the same conclusion, then one's confidence in that conclusion is correspondingly greater. Phrasing this differently, and in terms of the philosophy of science, a theory which ties together different areas and shows an underlying unity in apparently disparate phenomena is surely to be preferred to one which covers only the phenomena for which it was designed. Thus the mere fact of the variety of data and research methods utilised can act, in part, as a warrant for the conclusions of a research investigation.

Theorists of qualitative research (Morse & Richards, 2002, p. 76) would classify this thesis as utilising a triangulated research design:

> Triangulation refers to the gaining of multiple perspectives through completed studies that have been conducted on the same topic and that directly address each other's findings. To be considered triangulated, studies must "meet" – that is, one must encounter another in order to challenge it (for clarification), illuminate it

(add to it conceptually or theoretically), or verify it (provide the same conclusions). Goffman coined the term in 1989, drawing on the metaphor of the surveyor's practice of making sightings from two known points to a third.

Of course, a mixture of research methods without a clear theoretical justification applied to bits and pieces of selective data proves nothing in itself. Morse and Richards (2002) also note on the same page that although a "text search of grant applications in many countries would return a high count for the odd word *triangulation* – so would a vote among researchers for the most misused term." Thus, it is the function of this chapter to justify the use of those research methods which were in fact utilised – to explain their use in addressing the question of conceptual development in mechanics, and to address their limitations. Because of the level of detail required, however, the question of how the results clarify, illuminate or verify this thesis' conclusions must be left to the individual studies reported in later chapters.

Outline of research methods used

The aim of this thesis is to show that concepts in mechanics develop in the same way as everyday concepts.

This involves

- defining these concepts
- elucidating the prototype structure of these concepts
- examining the use of these concepts

Thus one strand of this thesis is oriented to explicating the meaning of the words 'concept', 'acceleration' and 'force.' For this purpose lexicographic methods were used.

A second strand of the thesis is oriented to examining the prototype structure of words. For this purpose survey instruments were developed in order to gather data, and statistical analysis used on the results to ensure that they reflected real differences rather than random fluctuations due to variations in sampling. The statistical methods of the behavioural sciences are appropriate here, and are also appropriate for looking at the test scores of students reported in Chapter Six.

A third strand of the thesis is directed toward understanding the use of these concepts by students. Here some of the tools of Physics Education Research are used, including particular test instruments:

- the Force Concept Inventory (Hestenes et al., 1992). This is a very widely used multiple choice test (Hake, 1998, reviews 62 studies using it).
- The Force and Motion Conceptual Evaluation (Thornton & Sokoloff, 1998).

Using these standard tools was appropriate as they have been thoroughly validated (see for a detailed examination of this process Saul, 1998, pp. 131-166) and as it enabled comparisons to be made between the answers of the students who were studied in this investigation with the answers from a wider population. Nevertheless, these are multiple choice tests and there are issues associated with their use that need to be addressed later in this chapter.

A fourth strand of the investigations reported in this thesis is the use of the tools of discourse analysis: in addition to examining student answers to multiple choice tests to find evidence for the use of concepts, student discussions of the answers to these problems were audio-taped and analysed.

Overall, then, four main types of research methods were used: lexicographical, statistical, educational, and discourse-analytic. These will now be discussed in turn.

Lexicographic methods

According to Sinclair (1985, p. 81), in order of decreasing popularity, the main sources of lexicographic evidence are:

1. Other dictionaries
2. Introspection
3. Observation of language in use.

Dictionary-use

The first of these methods is, then, the straightforward approach to finding the meaning of a word by consulting a dictionary. Again, according to Sinclair (1985, p. 81):

> The great value of received description is that the information is already organized. Language change is not so rapid that descriptions go out of date quickly, and from decade to decade we must assume that most existing work is valid and accurate.

This method is used in Study Three of Chapter Six. As is detailed there, although it has the advantage of useful organisation and ease of access, this method suffers the limitations of

- inconsistency – different dictionaries give different definitions
- obscurity – less complex words are explicated by more complex words
- circularity – words are defined ultimately in terms of themselves
- inappropriate focus – the aims of the dictionary makers are not those of science educators.

Language corpora

To overcome the limitations inherent in dictionary definitions one needs to put oneself in the place of the writer of such a definition, and go to the primary data: the recorded usages, and one's own introspective awareness of the sense of words in use. Language corpora were used in Chapters Four (for 'concept'), Five ('acceleration') and Six ('force').

According to Biber, Conrad and Reppen (1998, p. 4):

> The essential characteristics of corpus-based analysis are
>
> - it is empirical, analyzing the actual patterns of use in naturally occurring texts;
>
> - it utilizes a large and principled collection of natural texts, known as a "corpus," as the basis for analysis;
>
> - it makes extensive use of computers for analysis, using both automatic and interactive techniques
>
> - it depends on both quantitative and qualitative analytical techniques

These characteristics make corpora useful in this study. The principled selection of texts (see Kennedy, 1998 pp. 13-87 for details) ensures that the relative frequency with which students encounter various lifeworld uses of, for example, senses of the word 'force' can be reasonably approximated. The very large sizes of the corpora make it more likely (although, of course, not certain) that all relevant usages are found. Computers can rapidly sift through the massive database to provide KWIC (Key Word in Context) concordances which provide the standard method for determining the different senses of words by dictionary compilers (see Biber et al., 1998 pp. 26-28; and Kennedy, 1998 pp. 251-258; see also Wierzbicka, 1985 for detailed discussions on the methodology of definition; and Winchester, 1998 for a fascinating account of pre-computer methods of obtaining such concordances), and they can rapidly produce a random sample to reduce the vast quantities of material to a manageable size without introducing bias. This facility was used, for example, to obtain 250 random samples of general usage of the word 'concept' for analysis in Chapter Four, and 1000 random samples of the use of the noun 'force' for analysis in

Chapter Six. (It was unnecessary to reduce the sample size for 'acceleration' as this term was used less frequently, and all 137 recorded usages in the corpus could be examined.)

However, computer technology can only do so much, and while the KWIC concordance of 1000 usages of the noun 'force' used in Chapter Six can be examined quantitatively to ascertain the relative frequencies of usages of different senses, it is first necessary to decide just what senses of the word are represented, and which usages represent which senses – decisions which must be made qualitatively, indeed introspectively. The standard method for presenting such qualitative decisions, as used in serious dictionaries is to quote supporting instances of usage for each sense, thereby enabling the reader to judge the correctness of the analysis. This is supplemented in this thesis by detailed analysis, argumentation and justification for particular decisions. This sort of detail is important because, as Wierzbicka (1985, p. 19) puts it:

> … to understand the structure of the concept means to describe fully and accurately the *idea* (not just the visual image) of a typical representative of the kind: that is to say, the prototype. And to describe it fully and accurately we have to discover the internal logic of the concept. This is best done not through interviews, not through laboratory experiments, and not through reports of casual, superficial impressions or intuitions (either of 'informants' or the analyst himself), but through *methodical introspection* and *thinking*. (Emphases in the original.)

Even the comparatively simple quantitative decision involved in determining how many words were being used in something identified as a concept (see Chapter Four) requires qualitative judgement: no computer program is as yet able to make this sort of decision. However, in this case, the standard method of quoting supporting instances of each case is sufficient: for people the decision process involved in counting words is mechanical.

Statistical analysis

This thesis uses descriptive and inferential statistics and exploratory data analysis. These quantitative methods will be discussed in turn.

The use of descriptive statistics (such as relative frequency of word usages, or ages at which children start to use particular words) is relatively straightforward. Obtaining the figures on word usage from a corpus is a matter of using the tools which form part of the software packages which vary from corpus to corpus: for example the *freq* command in CHILDES (MacWhinney, 2000a, 2000b), or saving Frequency Tables from the program SARA for the British National Corpus (BNC, 1994). Details of the relevant software tools can be obtained from the references listed, and details of the corpora themselves and the data obtained from them will be found in the relevant chapters of this thesis.

Inferential statistics are also important to the argument of this thesis.

The two main tools of inferential statistics used in this thesis are the *t*-test for difference between means (used in Chapter Six), and Analysis of Variance (ANOVA, used in Chapters Five and Six). The first of these is appropriately used to determine whether the difference between the mean scores of two groups can reasonably be attributed to chance, and the second extends this to three or more groups (and gives equivalent results to the *t*-test when applied to two groups).

These tests give the probability p that a difference between means is due to chance. It is customary to reject the null-hypothesis (that there is no difference between means) if p falls below 0.05.

However, it must be born in mind that when many tests are performed this level will result in 5 false positives for every 100 tests. It is worth noting, too, that there is nothing magical about 0.05 – its choice is simply a convention – a p value of 0.051 and one of 0.049 fall on either side of this conventional limit, but there is in truth no great difference between them.

It has been noted that inferential statistics are being used less often in Science Education Research (White, 1997) as qualitative studies have become more popular.

This may be due partially to the difficulty in interpreting their results: the bare fact that a difference is unlikely to be due to chance is not a particularly informative result. One is far more interested in explaining what a difference is due to. For this one needs an explanatory model: typically this has been that some experimental intervention (application of fertiliser on this crop but not that crop in agricultural studies; or use of a new teaching technique on this class as opposed to a standard teaching technique with that class in educational studies) has caused the difference. One can then argue that since the difference is unlikely to be due to chance, it should be explained by the experimental intervention. However, if one does not have a causal model, the mere fact of the presence of a statistically significant difference between group means is not very helpful.

Furthermore, statistical significance and practical significance are quite different matters. Particularly for large groups, very small effects of experimental treatments can be detected. A difference of 0.4 marks out of 100 between two groups' mean test results may well be highly statistically significant, but of little or no practical consequence.

Given that there are problems with false positives when many tests are performed (exacerbated by the tendency to not publish negative results), that results require interpretation in the light of an explanatory model, and that statistical significance does not always correspond to practical significance one might well ask why these techniques should be used at all (Carver, 1978).

However, this would be an over-reaction. Even apart from their other uses, statistical tests serve as one way to avoid self-deceit. It is a simple fact of experience that students vary: if the differences between groups are no greater than would be expected due to random assignment of students to groups, then it is as well to be aware of this.

As well, where an explanatory model is available, and where the difference between means is clearly of practical significance, and where the *p* value is significant or highly significant (<0.001 for example), and this difference is consistent with the explanatory model, then this constitutes supportive evidence.

In particular, when only ordinal ratings are required (as is the case in considering prototype effects) "null hypothesis testing is an optimal method for demonstrating sufficient evidence for an ordinal claim." (Frick, 1996, p. 379)

As well as descriptive and inferential statistics, two exploratory methods of quantitative data analysis are also employed in this thesis: these are cluster analysis and factor analysis.

Cluster analysis (Everitt, 1988) is a technique used to produce classifications from initially unclassified data. As such it is clearly appropriate to classifying the components of the concept of *force*, as reflected in responses to a questionnaire, in Study One of Chapter Six. However (Everitt, 1988, p. 604):

> There are many problems associated with using cluster analysis techniques in practice, of which perhaps the most difficult is the assessment of the stability and validity of the clusters found by the numerical technique used. This presents a problem primarily because most (probably all) cluster analysis methods will give clusters even when applied to data sets containing no cluster structure.

Hence, the use of cluster analysis is restricted to the role of corroborative evidence when it is used in this thesis. That is to say, the similarity of the "cluster tree" to the results of earlier discussion is noted as additional evidence for the validity of the analysis, and the resulting groupings are then subjected to further tests. (The details can be found in Study One, Part One of Chapter Six.)

Factor analysis (Kline, 1994; Spearritt, 1988) is used to represent the relationships between a set of variables using a smaller set of inferred variables. It is used in Study One, Part Three of Chapter Six to point out a limitation in the construction of the questionnaire which was used, rather than as a direct contribution to the investigation of the topic of the thesis. The details can be found in the relevant section.

Multiple choice test instruments

As mentioned earlier, multiple choice tests were used in this investigation. These have many advantages and are widely used because of these. In contrast to the intensive labour involved in transcription and interpretation of discourse, they can be scored easily and quickly, yet can be written to test many different facets of a student's understanding. Instruments can be written that test factual recall, higher order reasoning abilities and, as here, typical student misconceptions. Instruments like the ones in this investigation use test-time efficiently, allowing topics to be covered in some depth. As well, because of their extensive use, their level of difficulty is known.

However, multiple choice tests have received much criticism. For example, it is clear that when teachers and students expect assessment to be mainly multiple choice this may lead to 'multiple choice teaching' (Smith, 1991, p. 10 quoted in Bennet, 1993, p. 18). As well, because it is difficult to create items that test higher order skills there is a tendency for many multiple choice tests to consist of items that test factual recall.

There is also the issue of test-wise students (Holland & Wainer, 1993; Mehrens & Lehmann, 1984, pp. 156-157). There may also be a bias against women students when testing is restricted to multiple choice items (Bennet, 1993, p. 19). While these are all important issues, they are not relevant to this investigation: the *FCI* was not used for course assessment purposes, and students were aware that although their tests would be marked, the result would not affect their course marks.

Indeed, partly to emphasise the difference between the *FCI* and standard classroom tests an unusual marking scheme was utilised, which might be called a probability format. The use of this sort of system is discussed in detail in Taylor and Gardner (1999) and a brief justification of its use in this investigation is given below.

Probability formats and marking schemes

It is often the case that students know enough to be able to eliminate from consideration several multiple choice options in a test question, without knowing enough to be certain as to which of the remainder is the correct answer. Marking schemes have been developed which give credit for this partial knowledge.

Developers of such schemes initially assumed that it would be necessary to come up with some way to motivate students to get them to admit to possessing only partial knowledge. It was assumed that:

> In order to be able to measure these degree-of-belief probabilities in an educational environment, it is necessary to have a scoring system designed in a special manner that guarantees that any student, at whatever level of knowledge or skill, can maximise his expected score if and only if he honestly reflects his degree-of-belief probabilities. (Shuford, 1966, p. 126, quoted in Hutchinson, 1991, p.86)

In order to achieve this, mathematical models were developed that awarded higher marks for answers where the subjective probability (the student's own assessment of the probability) of an answer being correct matched the objective probability (defined as the frequency of correct answers to the total number of answers for questions of a given subjective probability) of its being correct (de Finetti, 1972; Hendrickson & Buehler, 1971). Although by using these mathematical models a number of marking schemes were devised which could be shown rewarded students whose subjective probabilities matched objective probabilities (Savage, 1971), these schemes had the disadvantages of being complicated to score, and requiring pre-training in the format for the students. But students, even after some training with such a test format, were not using it appropriately (Koehler, 1974). The schemes were simply too cumbersome. They also confounded test scores with personality dimensions, because certain personalities tend to unwarranted confidence (Hansen, 1971; Jacobs, 1971).

An alternative approach to eliciting partial knowledge was taken by Rowell, Dawson and Madsen (1993). They had students rate each of the possible choices for an answer to a multiple choice question on a Likert Scale. The choices offered the students were: definitely correct, probably correct, maybe correct, probably incorrect or definitely incorrect.

On the one hand, the use of Likert scales in rating answers was not new, since similar proposals had been put forward in education since at least the 1960s, (eg. Michael, 1968; Rippey, 1968; Rippey, 1970). These are reviewed in Hutchinson (1991). There are even earlier versions of this proposal in psychophysical work (eg. Hollingworth, 1913). However, the emphasis of Rowell, Dawson and Madsen was different because they aimed at obtaining insight into student conceptualisations rather than focussing on assessment.

In their work with year 10 students Rowell et al. (1993) provided evidence that using Likert scales to rate multiple choice answers provided additional insights into student (mis)understanding.

However, there is a problem with interpreting student answers on a Likert scale. At least some students answer inconsistently. For an example consider the set of responses to one question illustrated in Rowell et al.'s article (1993, Fig, 4, p. 65). This indicates that out of 26 students, 17 described one answer as definitely correct and 15 described another answer as probably correct. Since the answers were mutually exclusive this indicates that at least 6 (i.e. 17 + 15 - 26) students inconsistently rated one answer to a question "definitely correct" and a different answer to the same question as "probably correct". Obviously, if one of a set of mutually exclusive answers is definitely correct, then the others would be definitely incorrect.

There are, then, problems associated with interpreting verbal statements of probability as used in these Likert scales. That this is the case is confirmed by Brun and Teigen (1988, p. 390), for example, who found that 'different contexts influence the interpretation of probability terms.'

To overcome this problem a simple modification of multiple choice answering was used in the research described in this thesis: each question was worth 10 marks, but these marks could be distributed amongst the answers. The score out of 10 for each question was the number of marks assigned to the correct answer. The instructions provided to students are given in Appendix B. Clearly if one answer was believed definitely correct, then it would receive all 10 marks, and no marks could be assigned to other answers, ensuring consistency in interpretation.

Students had no difficulty in using this scheme, perhaps because of their familiarity with the idea of a "mark out of 10." Taylor and Gardner (1999) show that while this marking scheme provides additional information on student misconceptions for particular questions, it nevertheless gives total test scores equivalent to the scores obtained by students who use the standard method of answering Multiple Choice Tests.

Discourse Analysis

Answering the FCI individually, and using the unusual answering method described above, also provided the basis for the discussions which took place the following day, where students worked in groups to answer the same questions. It was these conversations that were transcribed and analysed using the techniques of discourse analysis.

Level of detail in transcriptions

It is often the practice of educational researchers dealing with transcriptions of student discussions to tidy them up by eliminating pauses, hesitation phenomena ("ums" and "ahs"), repetitions, correcting grammatical mistakes and so on. For many purposes this is necessary: where the focus is on the content of what is said, the incoherencies of spoken discourse are merely distractions. If we as readers are trying to get at the gist of what is being said by the students, these things get in the way, and we are grateful that they have been cleared up by the author before presenting them to us.

However, in the transcripts used in the following chapters all four of these – pauses, hesitations, repetitions and ungrammatical utterances – are recorded. This is done because, amongst other reasons, pauses and hesitations give clues to the structuring of knowledge into 'episodes' (Emmott, 1997) in the minds of students, repetitions give a clue to what students are attending to (Chafe, 1980; Chafe, 1976), and, finally, the exact form of words may give a clue to just how concepts are conceived by students. It is clearly grammatical to say "a force is exerted on a ball" and not grammatical to say "a force is given to a ball," so that, particularly when dealing with discourse by ESL students, one is tempted to silently correct the latter. However, these phrases may be indicative of different views of the meaning of 'force': force as interaction, versus force as transferable property, for example. For all of these reasons then, in this thesis the transcriptions will be more detailed than is usually the case in science education research, while not as detailed as might be encountered in linguistically oriented research where exact phonetic transcriptions of accent, voice quality, laughter, overlap, and so on may well be recorded.

Tone groups

Researchers into discourse state that when speakers communicate they package individual 'pieces' of information into what have been called 'tone groups' (Chafe, 1980, 1995; Halliday, 1985). Because these tone groups will also represent the way students package their knowledge of physics in their discussions, the transcripts in this thesis attempt to identify them, and put each on a separate line.

When one person is speaking at a time it can be straightforward to identify the tone groups by physical means: the pitch reaches a high at the tone centre (the first new information) and then follows a gradually decreasing curve, until the next tone group begins. (An example of the pitch contour of a tone group is shown in Figure 5. It was obtained by digitising the audiotape and processing it with the program Speech Analyzer from the Summer Institute of Linguistics (SIL, 2001). Details of the theory of acoustic phonetics, and the significance of fundamental formant traces can be found in Fry (1979).) Although the pitch curve is not completely smooth the pattern is clear enough. Variation is due to such factors as: 'content' words have their own distributions of stress (or accent) superimposed upon the overall pattern; and, low information words typically receive low stress, so where they occur the pitch drops. By 'low-information words' is meant

- function words like the articles, the conjunctions, and auxiliary verbs, such as 'is' and 'have,' which are required by grammar, but without which the content of a message would still be clear, although expressed perhaps in an unnatural telegraphic style.

- words which refer to things already mentioned: for example pronouns, indications of assent, and polite echoes of what has already been said.

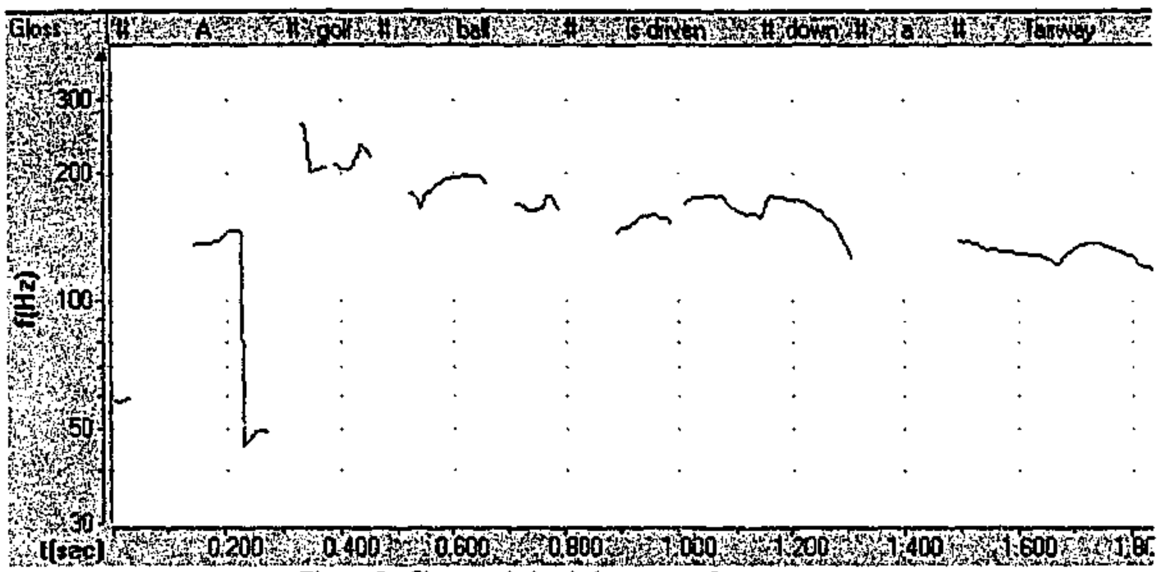


Figure 5: Characteristic pitch contour for a tone group. The words are approximately synchronised with the pitch contour below.

The example in Figure 5 is from the beginning of Group 97 5-3's discussion of question 22 (see Appendix 1), where Male Speaker 1 (MS1) is describing the question before the group discussion proper begins. A clear and regular drop in pitch after the first content word ("golf") is easily seen.

However, once a lively conversation starts to take place it becomes far more difficult to isolate tone groups by physical means (see Figure 6). Furthermore, the process is time consuming.

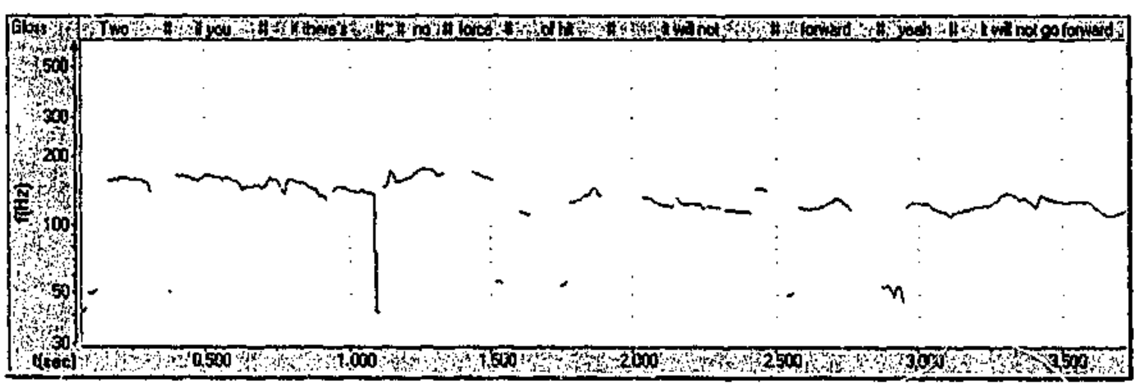


Figure 6: Pitch contours of an exchange between MS1 and MS4.

Fortunately, in the light of these difficulties, it is reasonably easy to analyse the discussions into tone groups by attending to the information structure and grammar.

Slight pauses, intonational patterns, and the introduction of new information are typically good clues as to where tone groups are, and it is therefore on this basis, not acoustically, that the transcriptions were divided into tone groups.

Summing up this section: it must be born in mind that while the transcripts are more detailed than have been typically used in Science Education Research in the past, they nevertheless suffer from some limitations: they are not as detailed as might be required for a study directed to their examination as discourse phenomena *per se*; as well, the tone groups were determined without examining pitch contours.

Conclusion

This thesis addresses a topic which does not fit neatly into a single disciplinary area, and so needs to draw upon methods used in the areas of Educational, Psychological, and Linguistic Research. It features a triangulated research design which uses a variety of data and analytical techniques that offer different approaches to the same research question.

Concepts in physics education research

Introduction

The topic of 'concepts' has been widely discussed in science education research over at least the last thirty years. However, within this field the properties of the 'concepts' themselves have not been investigated. In so far as there has been any consistency in the theoretical underpinnings of the idea of a concept, it has been consistently, traditional, indeed Aristotelian, notions of concepts that have been assumed. However, modern linguistic investigations of concepts have shown that they exhibit prototype structure and graded membership. To investigate the applicability of this work within science education, a number of research papers were examined to determine the actual usage of the word 'concept.' In spite of the Aristotelian assumptions behind the explicit discussion of the terminology of 'concept', the actual usage pattern for the word reflected a prototype structure. Hence, it is argued that the definition of concept used in linguistics should be adopted. Not only does this provide greater precision in discussing student understanding, but it also aids in making sense of the way the word 'concept' has been used in science education research.

The concept of 'concept' in science education research

When discussing, above, the issues of terminology used in science education research it was obvious that there was little agreement as to how 'concept' should be defined. However, definitions are not the only method of access to concepts. Taking seriously the linguistic notion that any concept is related to classification entails that one can also access the concept of 'concept' by looking at what is classified as such.

Dictionary makers need to examine the usage of words in context in order to discover and characterise the range of meanings they may have. In order to do this large collections of text have been created, called language corpora. These are often made available for use by researchers. One corpus which is available in this way is the British National Corpus. Using this it was possible to examine the usage of the word 'concept' in a representative sample of naturally occurring utterances. Using a sample size of 250 utterances, I counted the number of words which were used to describe what was labelled as a concept. From the results, shown in Figure 7, one can see that *concept* is most often used to refer to ideas that can be characterised by a single word. Examples of the utterances analysed can be seen in Table 3.

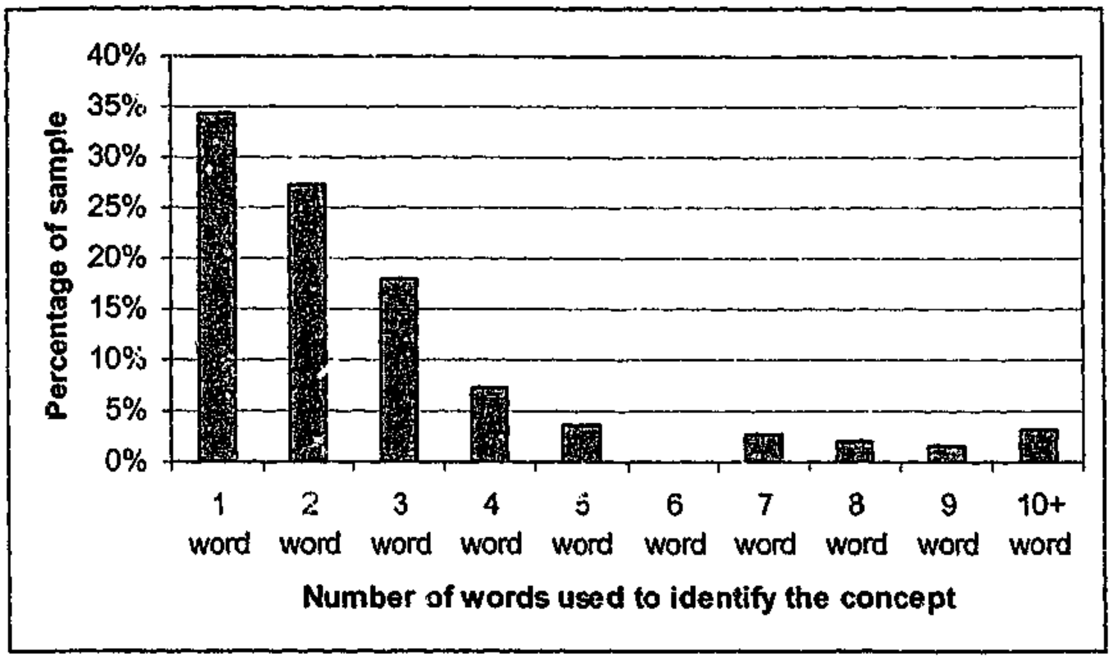


Figure 7: The usage of the word 'concept' in general discourse

Table 4: Illustrative utterances from the British National Corpus showing the usage of the word *concept*.

Number of words	Sample utterances
1 word	His current work on the concept of **flow** with Thom and Deleuze… …stressing the need to define the concept of **work**... …understanding of the elusive concept of "**literature**".
2 words	…concept of **home-ownership.** In "The Hollow Men", the concept of **wandering souls** would link…
3 words	We are implacably opposed to the concept of **payment to climb.** …palaces, **value for money** is a concept which cannot be satisfactorily…
4 words	This appears to make the concept of **locally-available core services** a… The concept of **loss of self-control** has received little close attention…
5 words	however difficult it may be to attach a precise concept to a **"general rise" in prices**. The concept of **a minister having Cabinet rank** - and therefore implicitly…

This data is consistent with the meaning of a single word being the prototype for the concept *concept* in general use. It remains a question as to how the word is used in its technical sense.

Using the Natural Semantic Metalanguage, one can attempt to characterise the meaning of 'concept' in its general use as follows:

X is a concept

> X is something people think. 1
> When people hear a word people think of something. 2
> X is something like this. 3

The idea of such a definition is to use the restricted defining vocabulary of the Natural Semantic Metalanguage to avoid circularity of definition. However, because of the very limited vocabulary, NSM definitions can give the impression of being oversimplified. Nevertheless, if one examines the definition given above one can see that it could be reworded using a larger vocabulary, so as to avoid this impression. The first line of the definition (1) indicates that a concept is not a physical, but a mental entity. Line 2 describes the prototype upon which the definition is based. Line 3 indicates that things are called concepts if they are similar to the mental entities which are called to mind when one hears a word. The use of the word 'people' rather than 'someone' in the definition is necessary to bring out the social nature of word meanings: particular individuals may well have their own idiosyncratic associations which are called up upon hearing a word, but these are not what we would call concepts. Although, therefore, one could rephrase the definition along the lines of "a concept is a mental entity similar to the mental entities brought to mind when people use a word," the conciseness of the rephrased definition, and its more adult tone, are at the expense of using words of greater complexity than necessary, and this also introduces the risk of circularity of definition.

For comparison the Oxford English Dictionary defines concept as follows (ignoring the obsolete usage which it lists as meaning 1):

> **2. a.** *Logic* and *Philos.* The product of the faculty of conception; an idea of a class of objects, a general notion or idea.
> **b.** Hence in weakened use, a general notion or idea, esp. in the context of marketing and design; a 'theme', a set of matching or co-ordinated items, of e.g. furniture, designed to be sold together. Chiefly advertisers' jargon.

The circularity of defining 'concept' in terms of 'conception' is apparent. Defining it using the terms 'idea' or 'notion' at first seems to avoid circularity: until, that is, one consults the definitions of these terms only to discover that they are defined using the term 'concept'.

If one, therefore, accepts the NSM definition as a characterisation of the meaning of the word concept in general usage, the question arises as to whether the usage of this word in science education research is similar or different. To determine this within the genre of science education research involves assembling a corpus of textual material and examining those things within the corpus that are explicitly referred to as concepts. There are two aspects to this: some papers will provide explicit definitions of the word concept, and others will simply illustrate its meaning implicitly.

Explicit definitions

Much research has already been done in identifying what students already know (or what they think they know) and there is a vast literature on student conceptions (for details see e.g. Pfundt & Duit, 1994) . One could question what more of value there is to be done.

Investigations have been conducted into how students make sense of what they are taught in science classes (Driver et al., 1995), their mental models (e.g. Gentner & Stevens, 1983; Johnson-Laird, 1983), and their beliefs (e.g. Dawson & Rowell, 1990; Gauld, 1987; Gunstone et al., 1989; Keil, 1991; Koballa, 1989). Researchers have investigated the process of change in conceptions (e.g. Posner et al., 1982), and methods for getting students to abandon non-scientific conceptions (reviewed in Scott et al., 1992).

What they have not done, by and large, is investigate students' concepts *per se*. This is in spite of the use of the word concept in the titles of many articles (e.g. Dekkers & Thijs, 1998; Domenech, Casasus, & Domenech, 1993; Domin, 1996; Evrard, Huynen, & Borght, 1998; Galili & Kaplan, 1996; Grayson, 1994; Greca & Moreira, 1997; Heinze-Fry & Novak, 1990; Heller & Huffman, 1995; Hestenes & Halloun, 1995; Hestenes et al., 1992; Huffman & Heller, 1995; Langford, 1987; Leonard, Gerace, Dufresne, & Mestre, 1994; Pankratius, 1990; Pines & Leith, 1981; Roth, 1993, 1994; Ruiz-Primo & Shavelson, 1996; Stavy & Berkovitz, 1980; Tornkvist, Petterson, & Transtromer, 1993; Trowbridge & McDermott, 1980, 1981; Wrobel, 1994). Clearly, what is at stake here, when it is claimed that few of these writers have actually investigated 'concepts' is at one level a question of terminology. However, it will be argued that more than this is at stake. The argument that will be made below has two parts. Firstly, it will be argued that the idea of a 'concept' in science education research has been theorised loosely: furthermore, the definitions put forward by different authors do not agree. Secondly, it will be argued that in spite of the various definitions of *concept* provided by different authors, that the word can nevertheless be seen to be used consistently. The idea of prototypes not only enables one to make sense of data that has been collected about colour terms, and words like bird, it is also useful in making sense of the way the word 'concept' has itself been used in science education research. In Chapters Four and Five of this thesis, the idea of prototypes will also be used to provide greater precision in discussing student understanding of physics concepts.

Implicit illustrations of the meaning of 'concept'

Journal papers with the word 'concept' somewhere in the title (including cases where it was part of a word, as in 'misconception') were used in this study, as it seemed likely these papers would explicitly label things as 'concept' within the body of the text. Twenty-five of these papers were examined (Pfundt & Duit, 1994, lists 804 such papers altogether). Of these twenty-five papers, six were from *The International Journal of Science Education,* five from *Science Education,* four from *Journal of Research in Science Teaching,* two each from the *American Journal of Physics* and *Studies in Science Education,* and the remainder from other sources. While this is clearly not a random sample from the population of science education papers, the purpose of the investigation was not to form a statistical characterization of this population, but rather to determine how the word 'concept' could be used. For this purpose, the sample chosen proved serviceable.

For each paper all those things that were judged to be explicitly referred to as concepts were recorded, together with sufficient context to indicate why this judgment was made. Where there is no context indicated, an ellipsis generally indicates a phrase like "the concept of". The results are shown in Table 5.

It is apparent that the vast majority of things referred to as 'concepts' were represented by single words (such as 'animal', 'force' or 'time'), consistent with the position of Carey (1991) and of Dykstra (1992), mentioned above, p. 13. A smaller number were represented by two words (for example 'chemical reaction', 'potential difference'). There were also examples composed of a short phrase (for example "force as an innate or acquired property"), and examples of statements, or propositions, referred to as concepts (e.g. "Weight is determined by the measuring method, including what one's own body feels"). These latter are obviously inconsistent with the position of Carey and Dykstra if one takes the concept *concept* to be classical. However, if one accepts that *concept* itself has a prototype and graded membership then the evidence falls into place.

Examining the use of these different examples one can see that all the listed authors, with the exception of Greeno (in this particular article: 1983), use 'concept' to refer to 'single-word-concepts'. It seems that the prototype for the concept *concept* is a mental construct which can be named with a single word. Other uses are poorer examples: for example, concepts formed by restricting the meaning of one word either by another modifying word before it (an electric current is a particular type of current), or by a modifying phrase following it (as in "a table as a rigid body"). Further elaboration results in short sentences (for example one could have written with little difference in meaning "a table is a rigid body", in the same way as what was written as "Weight is established by Newton's Law of Gravitation" could have been "weight as established by Newton's Law of Gravitation"). Similarly as one progresses from simple words to sentences one sees a progression from simple adjacency as in "the concept gravity" to use of the preposition "of" as in "the concept of force as…" to the use of the preposition "about" as in "the concepts about free fall and gravity". Personally, I find the first two usages perfectly acceptable, but find the third less so, again reflecting the prototype structure of the concept *concept*.

Table 5: Details of all the things explicitly referred to as concepts in a sample of 25 papers.

Details	Reference
…animal, … plant, …living, …force, …friction, …gravity, …electric current, …light, …chemical reaction…	(Ameh & Gunstone, 1985, p.151)
…weight …heaviness or lightness as a property of objects	(Bar, Zinn, Goldmuntz, & Sneider, 1994, p. 149)
Scientific Concepts: Weight is established by Newton's Law of Gravitation …{and other statements} Everyday Concepts: Weight is determined by the measuring method, including what one's own body feels. …{and other statements}	(Table 1 Bar et al., 1994, p. 150)
…the concept of force as an innate or acquired property	(Brown, 1992, p.25)
…a change in their concept of a table as a rigid object	(Brown, 1994, p. 209)

Details	Reference
to the table as a kind of spring	
...space...time	(Brouwer, 1984, p. 602)
...density, heat...temperature	(Brouwer, 1984, p. 603)
The banking concept of science teaching	
...a physical object	(Carey, 1991)
...object, matter, weight	
direct current concepts	(Chambers & Andre, 1997, p. 107)
electrical concepts	
science concepts	(Chambers & Andre, 1997, p. 110)
...mass, acceleration, momentum, charge, energy, potential difference, torque	(Clement, 1982)
...scope... force	(Cobern, 1996, p.579)
... concepts of evolution... scientific concept... everyday concepts... prior concepts... long-held concepts... alien concepts	(Cobern, 1996, p. 583)
"...To see through all things is the same as not to see"...This is a difficult concept for the scientifically inclined...	(Cobern, 1996, p. 585)
...literacy	(Cobern, 1996, p. 586)
...the concept about electric current...	(Cobern, 1996, p. 595)
...force	(Dekkers & Thijs, 1998, p. 31)
...interaction	(Dekkers & Thijs, 1998, p. 43)
...force	(di Sessa & Sherin, 1998, p. 1156)
...living thing	(di Sessa & Sherin, 1998, p. 1159)
...force...mammal	(di Sessa & Sherin, 1998, p. 1161)
...bird... bachelor... force... number	(di Sessa & Sherin, 1998, 1164)
...force... number... animal... velocity... acceleration	(di Sessa & Sherin, 1998, p. 1166)
...the concept of spatially-localized and permanently existing entities...	(di Sessa & Sherin, 1998, p.1173)
...the concept of object permanence...	
...force... acceleration	(di Sessa & Sherin, 1998, p.1175)
...dog... force... number	(di Sessa & Sherin, 1998, p. 1187)
...mass	(Domenech et al., 1993)
...work... energy	(Driver & Erickson, 1983, p. 45)
...volume... density	(Driver & Erickson, 1983, p. 52)
...mass, volume, density... particulate nature of matter... speed	(Driver & Erickson, 1983, p. 53)
...energy	(Duit, 1993, p. 15)
...*mib* {*Artificial concept invented for the purpose of*	(Dunn, 1983, p. 648)

Details	Reference
research.}	
…insect	
…velocity, acceleration… force	(Dykstra Jr., 1992, p. 42)
…position... instantaneous velocity… acceleration… momentum	(Graham & Berry, 1996, p. 75)
…energy… current	(Grayson, 1994, p. 2)
…constant voltage…constant current	(Grayson, 1994, p. 3)
… momentum… force… energy	(Grayson, 1994, p. 5)
…abstract concepts such as conversion of momentum…	(Greeno, 1983, p. 240)
…electromagnetic field… field	(Greca & Moreira, 1997, p. 712)
…atom… molecule	(Griffiths & Preston, 1992, p. 611)
…chemical bonding, chemical reactions, ions…states of matter	(Griffiths & Preston, 1992, p. 612)
…force… motion	(Gunstone, Champagne, & Klopfer, 1981, p. 27)
…density	(Gunstone et al., 1981, p. 31)
Concept: 1. What is a force? 2. General idea of force as a push or pull. 3. Directionality is important. {*and further statements*}	(Entries in one column of the table in Appendix I of Gunstone, Gray, & Searle, 1992, p. 191)
…movement, force… energy	(Gutierrez & Ogborn, 1992, p. 209)
…a model of Linear Causal Reasoning… The concept of Linear Causal Reasoning	(Gutierrez & Ogborn, 1992, p. 216)
…gravity, balanced forces… projectile motion	(Guzzetti, Williams, Skeels, & Wu, 1997, p. 706)
…counterintuitive concepts about free fall and gravity	(Guzzetti et al., 1997, p. 707)
Although Wayne had defined inertia in a lecture, several students were also unfamiliar with the term and/or the concept.	(Guzzetti et al., 1997, p. 713)

In summary, an analysis of the way the word 'concept' is used in science education research literature accords with its usage in everyday contexts better than it accords with any one author's explicit descriptions of concepts. Furthermore, both the general usage, and the usage in science education research papers illustrate that the concept *concept* has a prototype structure with graded membership. In fact, the evidence gathered in this chapter suggests that the word 'concept' is used in science education research literature with much the same meaning as it is used in general contexts, irrespective of the theoretical accounts discussed in Chapter Two.

The usefulness of prototype theory has been established in this chapter in discussing the terms used by science education researchers. The next chapter looks at a concept which plays an important role in science education itself: *acceleration.*

An investigation of the structure of the concept of acceleration

Technical concepts

Some might agree that the use of the word 'concept' does in fact show that it has a prototype structure, but still disagree about the applicability of prototype structure to technical concepts like *acceleration*. DiSessa and Sherin, for example, are willing to concede that the prototype account of concepts works well for lifeworld categories like *bird* and *bachelor* but claim that it is a *prima facie* absurdity to apply this to technical concepts like *force* and *number* (di Sessa & Sherin, 1998, p. 1164).

Just as Rosch and others discovered in the 1970s that concepts like *bird*, or *furniture* or *animal* did not clearly delineate classes of objects in the real world, in so far as their informants would classify certain instances of these classes as better and others as worse examples, so later researchers discovered that the same conceptual structure of prototype and marginal cases applied to such apparently unlikely cases as *odd numbers*, where 3 and 7 are prototypical examples and numbers like 23, 57 and 447 are successively less good examples (Armstrong, Gleitman, & Gleitman, 1983). Given that an odd number is defined mathematically, Armstrong interpreted this result as showing that concepts do not in fact have a prototype structure: that the typicality ratings are simply a by-product of the routines our brain uses to identify instances of a concept. But this argument does not affect the importance of prototypes for science education: whether or not technical concepts are completely clear-cut (as the concept *odd-number* presumably is) the fact that we find a prototype structure and graded membership within them is irrefutable. Given firstly that some concepts are vague (like *bald*) and, secondly, that some are clear cut (like *odd-number*) and, thirdly, the evidence for the prototype structure of the concept *concept*, one can conclude that clear-cut categories are simply one sort of concept. Although there are different sorts of concept, the human mind nevertheless seems to deal with all of them in a similar way: initially building upon a foundation of a prototype and only later clarifying the boundaries which separate examples from non-examples.

This can be seen to be the case if one looks at another technical concept: *acceleration*, the rate of change of velocity. Acceleration is not, by and large, a word in everyday use. People would normally encounter it often only in restricted contexts.

Evidence for this can be seen in the results of two large-scale projects carried out to determine word frequency, one in the USA and another in Britain. Each of these took five hundred extracts, each of 2000 words, from texts selected from fifteen genres. The word 'acceleration' occurred in each of these million word samples of written text both infrequently and in restricted contexts. It was present only four times in British texts (Johansson & Hofland, 1989) and nineteen times in US texts (Francis & Kucera, 1982). It was present in only two out of fifteen genres in the British sample, and in four out of the fifteen genres in the US sample. In fact, the US figures overestimate the frequency with which the word 'acceleration' is used, as one of the articles selected for the corpus happened to be about accelerometers.

Further evidence can be seen by consulting linguistic corpora. These are very large computerized databases of text samples which can be searched in a variety of ways. The COBUILD corpus, for example had (at the time this chapter was written) 137 instances of the word 'acceleration' in it.

Further evidence can be seen by consulting corpora such as that of COBUILD (2000). This consists of a large computerized database of texts that can be searched conveniently. When using the corpus of texts one supplies the word for which one is searching and receives a list of extracts from the corpus, each centered around the word of interest. The COBUILD corpus contains:

- American books, ephemera and radio (nine million words)
- British transcribed speech (ten million words)
- British books, ephemera, radio, newspapers, magazines (twenty-six million words)

In this total of 45 million words of text, 'acceleration' occurs only 137 times. Roughly 40% of these uses are in the context of cars (e.g. "...gives a top speed of 130mph and acceleration to 62 mph in 8.5 seconds."), 20% are in the context of politics (e.g. "...freedom movement also produced an acceleration in militancy."), 10% in the context of sport ("... from short range, and excellent acceleration by Arnold produced a try.") with the remainder harder to classify (e.g. "...and tomorrow – thoughtless acceleration, unrestrained speed."). In each case, it should be noted, the sense is *go faster* where some of the usages (in politics, for example) are metaphoric (the "militancy" doesn't actually "go" anywhere).

In so far as 'acceleration' is an abstract, formal term and, as Wierzbicka (1985) notes, formal abstract terms are generally easier to define than concrete ones like 'cat', it seems to be an excellent test case for applying the Natural Semantic Metalanguage.

To get an idea of how the NSM deals with definitions, one can first attempt a definition of 'fast':

> Something is fast =
>
> Something can move from some place to a far place in a short time.

Each of the "words" in the above definition is taken to represent an innate concept. (The scare quotes around "word" are there because the innate concept of *a short time* is represented in English with a short phrase.)

One can then attempt a definition of 'acceleration' as it is used in the quotations represented above. (The line numbers do not form part of the definition, but are included for ease of reference in the following discussion.)

> Acceleration =
>
> Something that happens like this: 1
>
> Something cannot move from some place to a far place in a short time. 2
>
> After a short time it can move from some place to a far place in a short time. 3

This NSM definition describes the lifeworld usage of the word 'acceleration:' that is to say, the meaning of 'acceleration' which students will bring with them to the classroom. Loosely, what we call acceleration is the process (i.e. "something that happens", line 1) of something rapidly changing ("After a short time…", line 3) from going slowly (line 2) to going fast (line 3). This definition can, of course, be criticized. For example: it does not *explicitly* say that we can speak of acceleration even when something is moving fast already. However, in defence of this definition, the situation of accelerating, when already moving quickly, falls under the rubric of "like this" (line 1). Accelerating when already moving quickly is *like* accelerating from slow movement to fast movement. The key point to be made is that it is possible to define the everyday, lifeworld meaning of acceleration non-technically and without circularity.

Given that the idea of acceleration is infrequently referred to in everyday life, and yet is a key concept in Newtonian physics, it is clearly a technical term. A second investigation was therefore carried out. Its purpose was to determine if the technical concept 'acceleration' had graded membership, and whether it possessed a prototype.

It was hypothesized that situations falling under the technical definition of *acceleration* would be rated consistently as better or worse examples of *acceleration*. It was further hypothesized that there would be a prototype which all would accept as *acceleration*, while less good instances would be rated sometimes as examples and sometimes as non-examples of *acceleration*.

Method

Subjects

Sixty-four students in intact Physics classes participated. They were students at a senior secondary college for international students in Melbourne, Australia. Apart from English, almost all of these students spoke at least one other language, mainly Chinese (Mandarin, Hokkien or Cantonese), Indonesian or Malaysian. The students had all completed units on kinematics and Newtonian mechanics.

Pilot study

Eighteen of these students in one Physics class participated in an informal pilot study.

The pilot study was to determine whether there was any *prima facie* evidence to indicate that it might be worth investigating if the concept *acceleration* had a prototype structure. The students in one class were simply asked to put up their hands to vote for the best, the average, and the worst examples of acceleration out of 'getting faster', 'getting slower' and 'changing direction'. The unanimous response of this class was to classify these as best, average and worst respectively. Clearly it seemed possible that this concept had a prototype structure. However, there were some obvious problems with this as evidence: the students could see each others responses and might be influenced by peer pressure, the students might have been influenced by the order in which the options were presented, and the students might have answered what they felt the questioner wanted to hear.

Questionnaire study

The investigation proper involved three different classes of students one week later. They were asked to write out their answers to a questionnaire written on the blackboard regarding the degree to which they felt that different situations of acceleration were good or bad examples of acceleration. Forty-six students in three different intact Physics classes answered this questionnaire. Putting one's name on the questionnaire was optional, and twenty-six students did so.

The questionnaire

This questionnaire was in two parts: the first part attempted to concretize the situations by dealing with examples of the motion of a car. (The situations are given in table 2.) For each of these the students were asked to complete a scale from 0 to 10, with higher scores indicating better examples of acceleration. The second part of the questionnaire repeated the earlier pilot study, but, instead of giving three types of acceleration to select from students were told 'There are three types of acceleration. Please list them in order from best to worst.' The intention here was to avoid giving the three types of acceleration in any particular order.

Results

The ratings were analyzed to provide a summary of grouped frequency distributions, histograms, and were also compared using ANOVA. For purposes of analysis the 11 possible scores were collapsed into five categories: 0 = non-example; 1-3 = poor; 4-6 = average; 7-10 = good; 10 = perfect. The results are shown in Table 6 and Figure 8.

Table 6 Percentage of student classifications of different situations as exemplifying 'acceleration', together with mean and standard deviations of the ratings for each situation.

	A car getting faster	A car starting from rest when the lights go green.	A car getting slower.	A car falling after going over a cliff edge.	A car turning a corner	A parked car.
Perfect Example	77	51	26	28	26	6
Good Example	13	23	38	34	13	11
Average Example	4	15	28	21	32	11
Poor Example	4	9	2	13	21	9
Not an example	2	2	4	4	9	64
Total	100	100	100*	100*	100	100*
Mean rating	9.2	8.0	7.3	5.7	7.0	2.2
Standard Deviation	2.1	2.9	2.7	3.4	3.0	3.4

*Figures shown in the columns do not sum to 100 because of rounding.

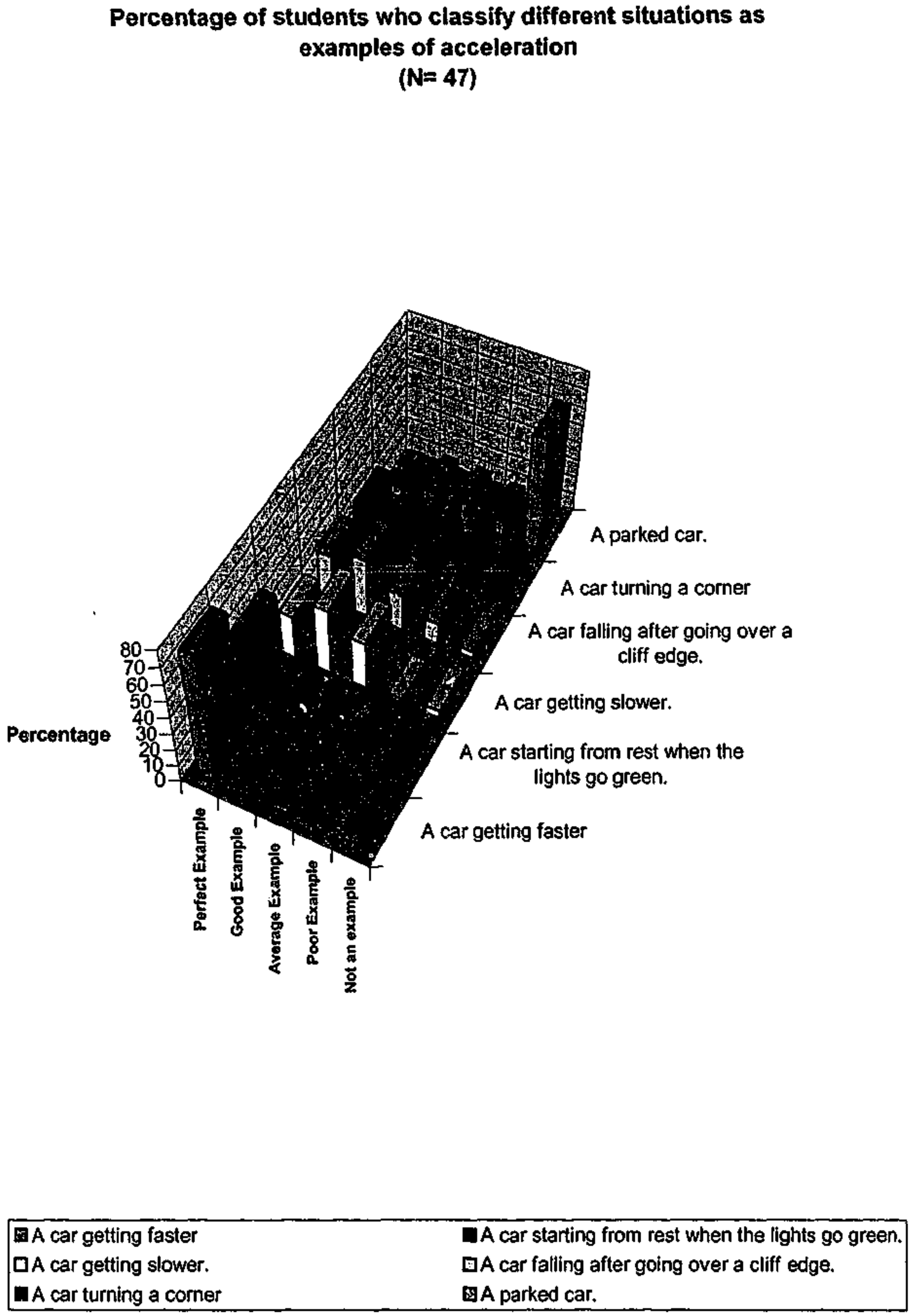


Figure 8: Ratings for situations involving acceleration

Although the second part of the questionnaire was intended to avoid presenting the options in a particular order, the instruction's wording ('There are three types of acceleration. Please list them in order from best to worst.') proved to be obscure to students, and many responses were left blank, were uninterpretable, or went off in a different direction (eg. one response was 'linear, gravity, centripetal'). As a result of this, only 27 useful responses were obtained from 47 students.

These results are shown in Table 7, and Figure 11.

Table 7: Percentage of student classifications of different situations as exemplifying 'acceleration'.

	Faster	Slower	Change Direction
Best	78	11	22
Average	19	63	15
Worst	4	22	63

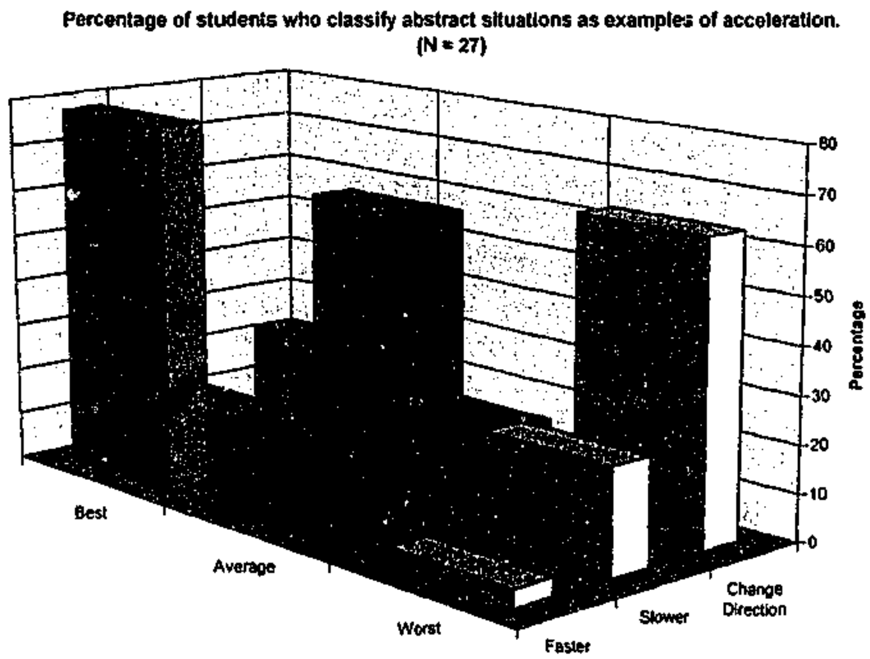


Figure 9: Comparison of ratings for three types of acceleration.

While it is clear that there is considerable variation from the unanimity of response obtained in the pilot investigation, it is also clear that the general picture remains the same: students take 'getting faster' to be the best example, or prototype, of acceleration. Discussion with students showed that deceleration is accepted as an example of acceleration on the basis that it is *negative* acceleration. Change of direction is clearly a marginal example of acceleration for these students.

Comparing these results with those on the previous page one sees some commonalties, and some surprises. As would be expected from the results in Table 2, the situation of 'a car getting faster' is clearly the best example for most students. 'A car starting from rest when the lights go green' is less central an example, perhaps because the idea of acceleration includes the component of increasing the quantity of speed, or perhaps because the idea of acceleration is not clearly distinguished from that of velocity so a car at rest is harder to consider as accelerating. After these two situations comes slowing, classified as average or better by most students. The car going over a cliff and falling and the car turning a corner both involve change of direction and students are less certain over their classification, and finally most students consider a parked car a non-example of acceleration. Those who do not consider it a non-example may think of it as having an acceleration of zero, rather than not fitting into the category of acceleration. Or, like those informants who classify bats as birds, they may have an idiosyncratic definition.

In any case, the ratings provided by students for each situation were significantly different from one situation to the other, but consistent across classes. These results are clearly statistically significant.

For both the studies whose results are reported in the tables above, comparing the ratings of the different situations using ANOVA gives $p < 0.0001$ (Table 6: $F_{crit} = 5.37$, $F = 27.86$. Table 7: $F_{crit} = 10.39$, $F = 21.75$). Hence one can be confident that it is very unlikely that the differences between the ratings of situations were due to chance. Furthermore, there were no statistically significant differences between the ratings provided by the three different classes surveyed. (ANOVA gives p values ranging between 0.204 and 0.963, with the exception of the ratings of the parked car situation where p is 0.079, approaching significance at the 0.05 level. Given the number of comparisons made, this is likely due to chance.)

There is clearly a prototype structure to the concept of acceleration, and this is in spite of the fact that this is a technical term that is encountered mainly in scientific contexts.

Historical analogues to conceptual development in learning science

It has been a standard approach in science education to compare students' conceptual development to the historical development of scientific theories during scientific revolutions (e.g. Hewson, 1981a, 1981b; Piaget, 1970; Posner et al., 1982). In some ways this has been a fruitful approach, suggesting useful ideas for research into aspects of science education. However, critics have argued that there are important differences between the conceptions that students exhibit and the conceptions put forward by historical figures (Viennot, 1985). Scientific theories, too, are much more complex entities than concepts.

An alternative model would be to compare student conceptual development with the historical development of the senses of words. Changes in the senses of words come from two main, often opposed, sources, "the effort of individual speakers to express and communicate their thoughts" which drives speakers toward extending the usage of the words and concepts they already possess, and the drive toward having one word, one meaning ("the isomorphy principle") (Geeraerts, 1997, p. 113). On this model, while the generalizing development in a concept such as acceleration (to cover all cases of change of velocity, not just increase of speed) is motivated by students' and teachers' efforts to communicate, the isomorphy principle explains some of the difficulties students have. When they enter the classroom, students already have the lifeworld concept of acceleration as increase of speed, but they also have concepts of slowing and changing direction and words to express them. So, for example, "slow down" and "decelerate", and "change direction", "change course" and "change tack" are all found in the COBUILD corpus with frequencies of the same order as "acceleration".

Students cannot be expected to simply adopt the physicist's notion of acceleration without a compelling reason. The reason for adopting the technical definition of acceleration is usually expressed in terms of all three situations (faster, slower, change direction) falling under the single mathematical definition of the vector acceleration as $a = \frac{\Delta v}{\Delta t}$. Other reasons for grouping these situations together can be found in terms of the relativity of frames of reference (Bowden et al., 1992).

However, when looked at abstractly, what is being asked of students is that they should group acceleration and deceleration under the single term acceleration: this is to put together what are two opposites in lifeworld terminology. It is of the same order of counter-intuitiveness as asking people to group black and white under the single term white. And when we proceed to the study of circular motion and ask students to add changing direction in, we are creating a problematic mix: acceleration is getting faster *or* getting slower *or* changing direction. It has been known since at least the fifties that disjunctive concepts (categories formed by having this property *or* that property) are amongst the hardest to learn (Bolton, 1977, pp. 102-105). Students are generally cooperative and will try to do their best, but integrating the mathematical definition of acceleration with its lifeworld meaning(s) is bound to be difficult. As the results above indicate, changing direction is for most students a very peripheral member of the category delineated by the concept acceleration. It is hardly surprising, then, that students have many difficulties in understanding circular motion (Gardner, 1984; Gunstone, 1984; Searle, 1985).

However, it would be wrong to give the impression that it is the prototype structure of concepts that is the source of problems: that students are somehow imprisoned behind prototypical barriers to learning. On the contrary, as Geeraerts (1997, p 113) has argued in his analyses of semantic changes of Dutch words as varied and as apparently totally unrelated to physics as the modern *legging* (English: legging), or the centuries old *vergrijpen* (English: to lay violent hands upon), the prototypical structure of concepts allows people to communicate by their shared possession of the prototypical meaning, and still to express themselves and to develop new senses by means of modifying the boundaries: generalizing, specializing, or hiving off peripheral instances to form the prototype for new concepts.

> … the cognitive system should combine structural stability with flexibility. On the one hand, it should be flexible enough to adapt itself to the everchanging circumstances of the outside world. On the other hand, the categorical system can only work efficiently if it does not change its overall structure every time it has to cope with new circumstances. … the development of peripheral nuances within given categories indicates their dynamic ability to deal with changing conditions and changing cognitive requirements. … the fact that marginally deviant concepts can be incorporated into existing categories indicates that the latter have a tendency to maintain themselves as particular entities, thus maintaining the overall structure of the system.

Prototypically defined concepts acts as the source upon which the new concepts, ones which we find to be necessary in our interactions with each other and with the world, can be developed. They are our opening through which we deal with the world: our concepts at any stage form the horizon of our outlook, and by developing them we expand our horizons. Once we have grasped the idea of acceleration by assimilating its prototype we can immediately begin to build up the Newtonian notion of net force as mass times acceleration without having to be fully clear about all the details. By clarifying and expanding the notion of acceleration we then enrich our understanding of force too. The problem, then, in learning concepts like acceleration is not the fact of the concept's prototype structure, but that the efficiency of the system can mask lack of complete understanding.

> ... the conceptual organization is not drastically altered any time a new concept crops up, but new facts are as much as possible integrated into the existing structure, which can thus remain largely unchanged.

In the above quotation Geeraerts (1997, p. 114) is discussing semantic changes in the meanings of words from the times of Middle Dutch to the present, but his words have obvious application to conceptual development in mechanics. Prototype theory has found widespread applicability in many, apparently unrelated, fields. It has been argued here that it is clearly applicable to science education research as well.

The development of the concept *acceleration* over time

Much research has been done which has looked at the way concepts like *force*, or *mass* (and many others) have fitted in with student beliefs: how they are utilized in talk about real situations; the extent to which they are used consistently; whether they form part of a framework or worldview; whether there are gender or cultural differences; whether there are common developmental sequences, and so on. Analogies with historically recorded theories have been common: people talk of students being misled by their *Aristotelian concepts of motion*, which is to say that students say things that can be interpreted as similar to things that Aristotle said. What all these various (and important) investigations have in common, however, is that they have looked at the concept from the outside: from how it is linked to other things. The assumption has been that the concepts of mechanics are classical and clear cut: researchers have themselves assumed, somewhat paradoxically, that the Newtonian *concepts of motion* are *Aristotelian concepts* needing no further investigation as they have no internal structure. White (1988a; 1992) is a partial exception to this generalisation, as he does not conceive of concepts as classical. However, he does not look at the internal, prototype, structure of concepts.

This chapter has demonstrated that such assumptions are false. The concept *acceleration* does have a structure, and this structure must develop over time.

An order of development of the concept of *acceleration* which is consistent with the evidence presented in this chapter is illustrated in Figure 10. Students start with the lifeworld concept that they have encountered in contexts outside the Physics classroom, and soon learn to include examples of slowing. Only a small number of students, however, initially extend their concept of *acceleration* to include cases of changing direction.

In so far as the study reported in this chapter was not a longitudinal investigation of the development of student conceptions, there is no direct evidence for the order in which the concept extensions occur. Nevertheless, it is typically found that less prototypical representatives of categories are the last to be added, so given the ratings shown in Table 7, the order indicated in Figure 10 is predicted.

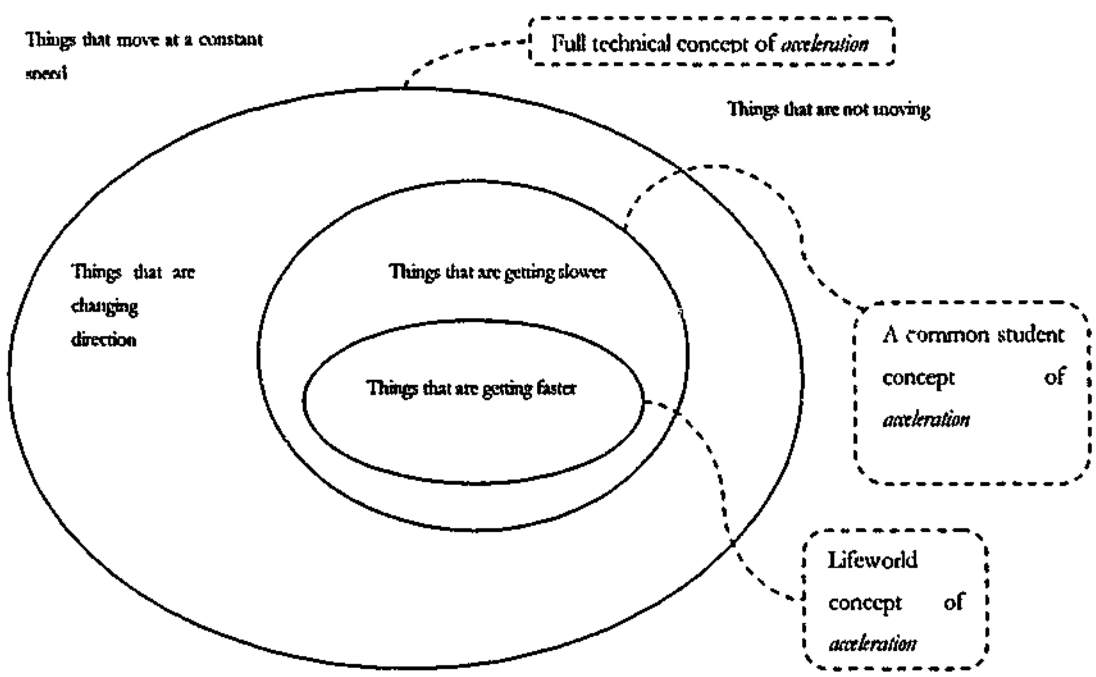


Figure 10: Development of the technical concept of *acceleration*.

Contrast with earlier investigations

This chapter has looked at the structure of the concept *acceleration*. In this, it is different in approach to the earlier investigations discussed in Chapter Two which had been largely restricted to looking at the use of this concept. Furthermore this investigation did not start with a fixed notion of what *acceleration* was.

The evidence shown in this chapter indicates that the meaning of 'acceleration' is not that of the ratio $\frac{\Delta v}{\Delta t}$ contrary to the assumptions of earlier investigators. "But how," one might ask, "can such a conclusion be reconciled with the definitions in a thousand Physics textbooks?" An answer to this question can be found in the ACT-R theory of cognition (Anderson, 1983; J. R. Anderson, 1996): calculating the ratio $\frac{\Delta v}{\Delta t}$ is one procedure by which one may calculate the measure of acceleration (others include solving kinematical equations such as $v^2 = u^2 + 2as$, or finding the slope of the tangent to a point on a *velocity-time* graph). Such procedures ('productions' in the nomenclature used by ACT-R theorists) are, or can be, activated when the concept of *acceleration* is in working memory, and one has the goal of determining its value. They do not represent the meaning of the concept, although they are closely linked to it.

In other words, acceleration is something that happens in the real world, and we can recognise it as such because we have a concept of *acceleration.* The formula for acceleration is a procedure we can subsequently use – if we know it, and if we wish to find the measure of the acceleration.

This is not to say, however, that such technical definitions do not play an important role in the development of the concept of *acceleration.* Clearly they do: using the fact that the formula $\frac{\Delta v}{\Delta t}$ applies equally well to all three situations is one method that a teacher can use to show students that the situations of getting slower and changing direction are like the situation of getting faster. It serves, in other words, to motivate the extension of the concept.

Having shown that the concept of *acceleration* has a prototype structure, the next chapter reports on an investigation of the related notion of *force.*

The lifeworld meaning of force

Introduction

As noted in Chapter Two, language has been cited as one possible source of misconceptions (e.g. Duit, 1990; Gunstone & Watts, 1985; Williams, 1999). Given that misconceptions abound in the area of Newtonian mechanics, and that it is one of the most thoroughly researched areas in physics education, this investigation of the role played by language in the understanding of physics concentrates on this area. Specifically, this chapter aims to characterize the concept represented by the word 'force', as it is used in everyday language to refer to physical force. It is argued that some of the difficulties that students experience in gaining a conceptual understanding of Newtonian physics are due to the need to integrate the Newtonian concept of *force* within the structure provided by this lifeworld concept. This chapter is focussed on the development of the meaning of the concept of physical *force* that students bring to school from their lifeworld experience. It details four studies.

Overview

The four studies employed different methods. The first study used the results of a questionnaire to elicit the prototype structure of the *force* concept in a group of senior physics students, average age 17 years. The second study investigated the usage, among children of ages from one-and-a-half to seven years, of words related to *force* (using the CHILDES corpora). The third study investigated the lifeworld concept of *force* by examining the definitions of force in dictionaries. The fourth study again investigated the usage of the word *force*, this time amongst the population in general (using the British National Corpus). These studies will now be described in more detail.

The first study investigated the structure of the concept *force* amongst senior secondary physics students by means of a questionnaire. Students were asked to rate various situations as good to poor or non-examples of *force*, and the responses were examined in the light of prototype theory discussed in Chapter Two (Rosch & Mervis, 1975).

The second study looked at children's usage of words related to force and motion. Prototypical examples are typically learned before other examples: looking at which words are used at different ages provides additional evidence for the prototype structure of the force concept. It is possible to examine word usage at different ages because developmental linguists have established common formats for transcription of children's discourse, and large numbers of transcripts have been made available by the original researchers for reanalysis, so providing rich source material for this study.

The third study examined definitions of 'force' provided by a number of dictionaries in the light of the results of the first study, with the aim of obtaining a clear characterization of the lifeworld senses in which we conceptualise *physical force*. The definitions from English dictionaries are also compared with those in several other languages to see which connotations are specific to a particular language and which are present cross-linguistically. However, while dictionaries are excellent sources for the explication of word meanings they are, in a sense, secondary sources: the usage of words by the general population are analysed and classified by lexicographers but the dictionaries contain only the final results of this work. It will become plain that dictionary meanings disagree with one another in different ways. It is also obvious that the analysis of this particular one word cannot have been a priority for lexicographers, who must deal with an entire language's vocabulary.

Thus the final study returns to the primary source material, using a random sample of 1000 usages of the word 'force' to examine them in the light of the earlier three studies, with the aim of clearly delineating this word's meaning. Naturally occurring uses of the word 'force' in adult discourse (both written and spoken) are examined in order to identify the patterns of usage of the word 'force' in the adult population. Once more, this investigation is made possible by the availability of corpora developed by linguists.

Prototypes

As discussed in the previous chapter, there is clear evidence for the relevance of prototype theory to the technical concept of *acceleration*. This chapter extends this work to the related concept of *force*: a concept that is shared by the lifeworld and technical vocabularies.

As discussed in Chapter Two, basic level, prototypical exemplars of a concept (e.g. *dog* as an example of *animal*, or *chair* as an example of *furniture*) are typically (see, e.g. Lakoff, 1987, pp. 41-46):

- produced first when examples are asked for,

- rated highly,
- easily visualised,
- amongst the earliest words to be learnt.

It will also be recalled from Chapter Two that where a concept is complex, it may have distinct multiple prototypes. An example discussed there was that of the Dani language of New Guinea, which uses just two terms to divide the whole of colour-space: 'mola' (referring to the warm colours) and 'mili' (the cool colours). In this language, some informants asked to identify the best examples of *mola* identify the prototypical *red* and others the prototypical *white* of English (see Kay & McDaniel, 1978, pp. 616-617).

This chapter presents evidence of a prototype structure for the concept of physical *force*. The above dot point criteria are used as a guide. The studies reported here provide evidence that the concept of physical *force* is complex and has multiple prototypes.

Study 1: The prototype structure of *force*

The subjects

Study 1 was carried out with thirty-five Physics students who were mainly from an Asian background but were fluent in English. They had been studying the final year of secondary physics (in English) at a senior secondary school in Melbourne, Australia.

They were asked to answer a questionnaire (Appendix F), which is described below. Filling in one's name on this questionnaire was optional, and twenty-one students did so. Of these, the second language of twelve was Chinese (four Mandarin, three Cantonese, and five unspecified), six students spoke Malay or Indonesian, and two spoke Vietnamese.

The questionnaire.

The students were asked to complete a questionnaire at the end of the course: this is reproduced in Appendix F. There were two main parts to the questionnaire, each subdivided into subsections. The first part was subdivided into four subsections. The first three of these subsections asked the students to rate each of ten words as to how good it was as an example of *force*, as used in physics. These first three subsections investigated ten everyday nouns, ten everyday verbs, and ten nouns taken from a technical context: see Table 6. The words were selected from entries in a thesaurus under 'force'. The fourth subsection of Part One asked the students to write a sentence describing a situation that was a good example of a force.

Part Two asked students to rate the forces present in various situations, which were described in one or two sentences. There were three subsections: the first subsection dealt with forces exerted **by objects** (specifically, by cars), the next subsection dealt with forces **by people**, and the third subsection dealt with forces **on objects**. The situations in all three subsections were chosen to vary along five dimensions labelled here as: transitivity, aspect, effectiveness, speed and acceleration. See pages 116 ff. for descriptions of the meanings of the five dimensions so labelled and Table 10 for the details of the ratings.

Results

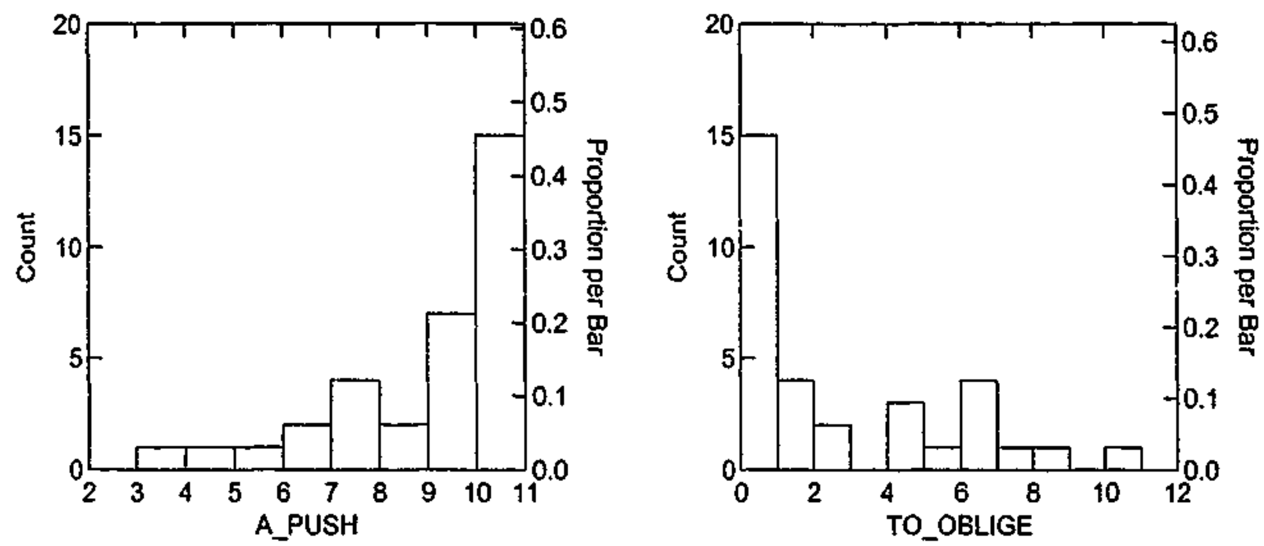
The following analyses pool all the data for all students irrespective of background, as it was found that there were no statistically significant differences between the responses of students from different language backgrounds. Neither were there statistically significant differences between responses from those who gave their name (hence allowing their background to be included in the analyses) and those who did not.

Study 1, Part 1: rating words

As expected, students had no difficulty with the request that they rate different words according to how good an example of force they were. In

Table 8 the words are listed in order of their ratings, from best to worst examples. The ratings were fairly consistent between students. Most items had a similar spread of ratings (the standard deviation is given in the table). The skewness in the distribution of ratings indicates that ratings were not randomly distributed about the mean: the higher rated items had a few low ratings but mainly bunched at the high end. Similarly the low items were skewed the other way. See Figure 11 for two examples.

Figure 11: Distribution of ratings for a good and a poor



example of physical force.

An analysis of the data from the first three sub-sections of Part One of the questionnaire indicates that the most prototypical exemplars of *force*, amongst those tested, are *pushing, pressing, punching,* and *strength.* Since there is no single factor common to all of these, it is argued that *force* is a complex concept, with multiple components to its meaning.

Table 8: Ratings of words as examples of physical *force*, ordered from best to worst.

Word	N	Mean	Std Dev	Skewness	Word	N	Mean	Std Dev	Skewness
a push	33	8.5	1.9	-1.4	to compel	30	5.2	3.1	-0.1
to press	33	8.5	1.5	-0.9	a propulsion	32	5.0	2.9	0.1
a punch	33	7.5	2.3	-1.3	to shove	31	5.0	3.2	-0.2
strength	32	7.1	2.9	-1.2	energy	33	4.9	3.8	-0.1
thrust	32	6.7	3.0	-0.7	to cram	29	4.9	3.1	0.1
to thrust	33	6.6	3.2	-0.8	drive	33	4.3	2.8	0.1
a strength	33	6.5	3.2	-0.6	to make	31	3.7	3.3	0.6
resistance	33	6.5	3.1	-0.9	impetus	29	3.6	2.4	-0.1
a pressure	33	6.1	2.6	-0.5	to pack	32	3.3	2.8	0.3
a squeeze	33	5.9	3.1	-0.7	to oblige	32	2.3	3.0	1.0
power	33	5.9	3.5	-0.6	to require	33	2.1	2.7	1.3
momentum	33	5.8	3.4	-0.7	an army	33	2.0	3.1	1.5
to propel	32	5.8	3.0	-0.4	a sinew	28	2.0	2.6	1.6
impulse	33	5.5	3.1	-0.4	a squad	33	1.8	2.6	1.7
inertia	32	5.3	3.4	-0.4	a shame	32	1.2	2.1	2.2

A glance at Table 8 shows that the idea of *a push* has the highest mean rating, narrowly followed by *to press*. These two words do contain a common component of meaning: exerting a force on an object in a direction away from oneself. Interestingly, while both *push* and *press* are, like *force* itself, quite general terms (pretty much anything can push or press) the next most highly rated example is the highly specific *punch*. A punch is a violent impact (i.e. as opposed to *tap*, a non-violent impact), which can only be exerted by a hand (i.e. as opposed to *kick*, by a foot), which is rolled into a fist (i.e. as opposed to *slap*, by an open hand). Both the specificity of the details that the word 'punch' calls to mind, and its nature as a unitary event occurring at a particular point of time make it more easily visualisable and so may contribute to its high prototypicality rating. Again, the force of a *punch* on an object is typically directed away from the 'puncher,' so it shares this component of the meaning with the first two items. However, as this sample of thirty words did not include any which represented an action directed toward the actor (such as 'pull', 'draw' or 'attract') the importance of this component of meaning cannot be determined.

The next four items consist of two pairs that were present in both the technical and everyday contexts.

Whether in the list of everyday verbs or of technical nouns, *thrust,* was rated very similarly. Also, at least in its everyday sense, it again shares the meaning component of a force directed away from the self.

The term *strength*, however, does not share this component of meaning. As well, there are significant differences between the ratings in the everyday and technical lists (p = 0.008 on a paired samples t test with df = 31). The key meaning component which *a strength* possesses is *ability*, while the key component of the adjective *strength* is the *degree* with to a property is present (Sinclair, 1987). Note also that both *strength* (as one speaks of 'the strength of' coffee, or acid, or the wind), and *a strength* (as in "each firm has its particular strength") are properties, rather than actions like the previous items. It has often been reported that students treat *force* as a property of a moving object and the presence of this meaning component may help to indicate why. The fact that both *strength* and *a strength* differ in important ways from other highly rated items, yet are also rated very highly indicates that the concept of physical *force* is complex. The concept of *force* has multiple components.

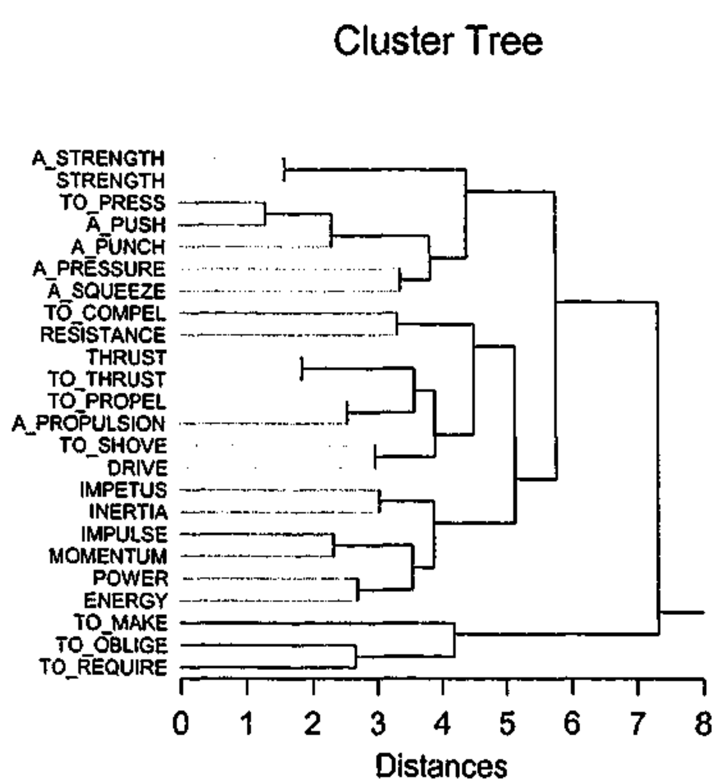
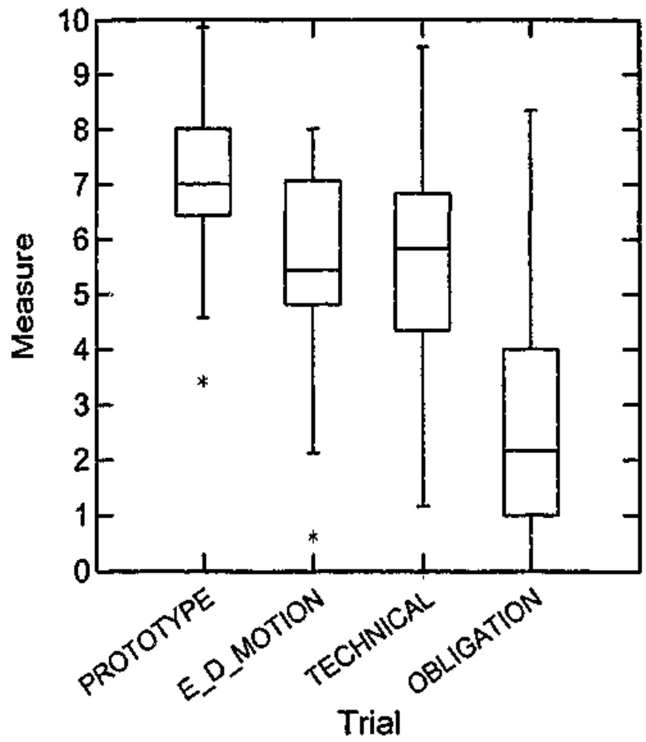


Figure 12: Cluster tree

This can also be seen from the results of hierarchical clustering applied to the matrix composed of the student responses: see Figure 12. The method used (Euclidean distance, furthest neighbours) performs well when the data forms natural clumps (StatSoft, 1999). Before analysis, a number of words (army, squad, pack, cram, sinew) were removed: they had received very low ratings because they were essentially unrelated to the concept of force in physics. (When included they mainly cluster together, but have little in common except their irrelevance.)

Starting at the top, there is the group containing the words identified above as the prototypes of physical force. Below this is a group of everyday words related to causes of and obstacles to motion: thrust, propel, propulsion, shove, drive, compel, resistance. Below this in turn is a group of six technical words also related in various ways to motion: impetus, inertia, impulse, momentum, energy and power. Finally, at the bottom of Figure 12, one sees a group of three terms related to obligation: to make, to oblige, to require. This is presumably because force can be used in expressions like "force him to do his homework".

A box-plot of the distribution of mean ratings for the terms in these four groups, Figure 13, shows that while the sense of obligation is clearly distinct from the other senses, that at least some students perceive terms from mechanics like momentum, power and energy as, at least similar to, if not actually representing, genuine Newtonian forces. Rating *power*, *momentum* or *energy* highly as an example of *force*, indicates a lack of conceptual clarity, which could impact upon the students' understanding of mechanics. There is some evidence that is consistent with such an expectation.



PROTOTYPE = best examples of forces;

E_D_MOTION = everyday terms related to causes of motion;

[illegible]

Figure 13: Box-plot showing distribution of mean ratings of main subgroups identified in Figure 12.

The scores for items on the *Force and Motion Conceptual Evaluation* instrument (Thornton & Sokoloff, 1998) were available for eighteen of the students. These students were divided into two groups: those whose mean rating for the six technical terms (impetus, inertia, impulse, momentum, energy and power) was above the group average (TERMINOL = High) and those below this value (TERMINOL = Low). Items on the *Force and Motion Conceptual Evaluation* were classified as relating to those involving constant velocity (NEWTON1ST), those involving non-zero acceleration (NEWTON2ND), and those involving Newton's third law (NEWTON3RD). In each case there was a lower score for those who had rated the technical terms like momentum more highly as examples of force: see Figure 14. However, the difference was only statistically significant at the 0.05 level for the Newton's first law examples (two sample t test, $df = 16$, $p = 0.037$), and this, while suggestive, obviously does not show causality.

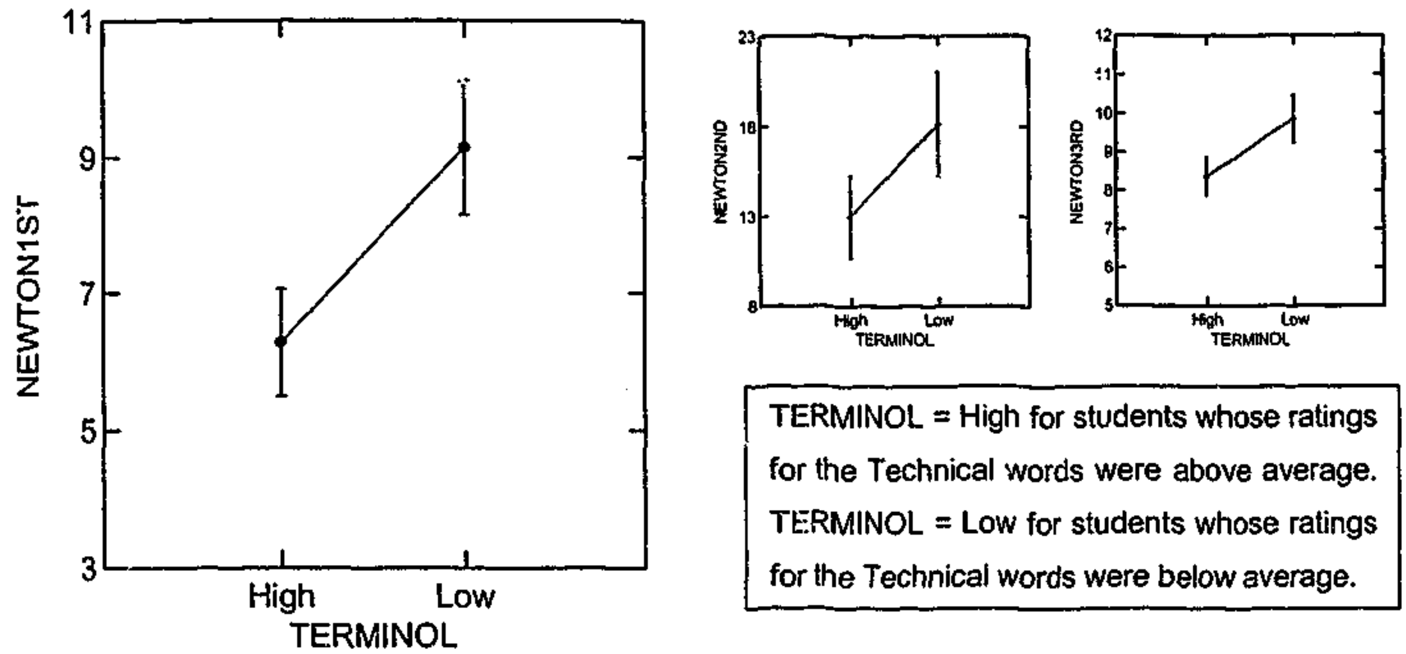


Figure 14: Comparisons between performance on test items related to Newton's laws.

Study 1, Part 2: Student produced-examples

The second part of this first study examined the results of the Questionnaire's Part One subsection four, where students were asked to write a sentence describing a situation that was a good example of a force. Thirty students did so: see Table 9. The sentences averaged nine or ten words in length and altogether 279 words (149 different word types) were used. The relative frequency with which various concepts were represented was examined.

The most commonly used word (apart from the articles) was 'push' (twelve times: four times as 'push', four times as 'pushing', and four times as 'pushes'). It was used more often, even, than 'force' itself (ten times).

The next most common idea (eight times) was of one object hitting another violently ('punch,' 'hit,' 'collide' were each used twice, 'banging' and 'bounce' once each). Possibly, given the context, 'fight' belongs to this category.

Motion was referred to six times ('move', 'moved', 'movement,' 'moving' twice, and 'travelling'). Friction, falling and strength (once as 'strong', once as 'strength') were each referred to twice.

These results are consistent with the first set: *pushing* and various forms of *hitting* are prototypical forces, but other components are present. One cannot therefore characterize the lifeworld conception of force as simply 'something like a push or a hit'. Amongst possibly other components of the meaning of physical force, there is some evidence for components related to *motion*, and *strength*.

Table 9: Responses to request for good examples of a force.

Example
When we lean on the wall, we exert force on the wall and the wall also exert the equal force on us.
A person pushes a box along the floor
A person pushes another person, so the former is exerting a force on the latter and vice versa.
A push on a thing.
Charles pushes against the wall.
Push a heavy stone.
Pushing a box on the floor.
Pushing by something.
When you push an object along a surface.
When you push somebody.
A moving car is pushing a standing car to move in front.
Any pushing force and frictional force and pulling force are good examples.
Something {pushes, pulls, collides} something.
A car banging into a tree.
A physics teacher was hit by a car traveling at 500 miles/hr.
When 2 cars come crashing together, they actually exert force on each other.
When two identical cars collide.
Two men fight with each other.
When we hit a ball with a bat the bat applies force on the ball.
Getting a punch on the nose is really hurt.
Punch Mr Taylor with our muscular and strong arms.
A ball free fall to the floor and bounce up again.
Attraction b/w a boy & a girl. 'Cause they've all got mass $F = GM_GM_E/R^2$
Some people fail the school result and jump from the 5/Fl to Ground.
Free fall Twisting something
In Physics force is mass × acceleration and the S.I. Unit is Newton (N). So a good example is the frictional force.
Reduce the speed of the car by braking.
Refusing to do anything
A force is when an applied strength will create movement.
Moving, from a stationary position. More accurately, being moved.

Study 1, Part 3: ratings of situations.

Similar results were obtained in Part Two of the questionnaire, in which the students were asked to rate fifty situations using the same scale with which they had rated the words.

Again, in this section of the questionnaire the prototypical examples of force are found to be *pushing*, and various forms of hitting (*punching*, *kicking*, and *collision*). The mean rating for each situation is shown in Table 10. Also shown are the classifications of these situations along six dimensions (the coding of each dimension is discussed below) which are labelled here as transitivity (see p. 119), aspect (see p. 120), effort and effectiveness (see p. 121), and speed and acceleration (see p. 117). The analysis of the data once more shows that a number of factors are at work when students rate these situations.

Table 10: Ratings of Situations, Ordered from Best Example to Worst, Together with Classifications as to Transitivity (p. 119), Aspect (p. 120), Effort and Effectiveness (p. 121), and Speed and Acceleration (p. 117).

Mean Rating	Example	Transitivity	Aspect	Effort	Effectiveness	Speed	Acceleration
8.7	A soccer player kicks a ball.	Active	Instantaneous	3	3	2	3
8.6	A boxer punches a boxing bag.	Active	Instantaneous	3	3	2	3
8.6	A door is pushed hard and it opens.	Passive	Beginning	3	2	1	3
8.3	A car collides with a truck at an intersection.	Intransitive	Instantaneous	.	4	2	2
8.1	A car is driven into the back of a parked truck. The car is wrecked. The truck is badly damaged.	Passive	Instantaneous	.	4	0	2
7.9	A weight lifter lifts 100 kg from the floor to shoulder height.	Active	Final	4	3	1	3
7.8	A door is pushed gently and it opens.	Passive	Beginning	2	2	1	3
7.8	A door is pushed hard but it doesn't open.	Passive	Beginning	3	0	0	0
7.4	A dentist pulls out a tooth.	Active	Final	3	3	1	3
7.4	A girl presses a button on a camera to take a picture.	Active	Instantaneous	1	2	1	0
7.4	An iron bar is compressed.	Passive	Final	4	3	0	0
7.4	An iron bar is twisted out of shape.	Passive	Final	4	3	1	0
7.3	A car gets faster.	Ergative	Final		2	2	3
7.3	A rubber band stretches.	Ergative	Stative	2	2	1	0
7.1	A car pulls a truck out of a mud hole where it was stuck.	Active	Final		3	1	3

Mean Rating	Example	Transitivity	Aspect	Effort	Effectiveness	Speed	Acceleration
7.0	A car is driven past a parked truck, hits the truck's taillight and breaks it. The car doesn't stop.	Active	Instantaneous		2	2	0
6.9	A rock falls 100 m.	Intransitive	Final		3	2	3
6.7	A criminal squeezes a gun's trigger.	Active	Final	1	4	1	0
6.5	A car is driven into the back of a parked truck. The car is wrecked. This truck is not damaged.	Passive	Instantaneous		0	1	2
6.4	A car gets slower.	Ergative	Final		2	1	2
6.4	A car goes around a curve.	Ergative	Ongoing		1	2	1
6.4	A door is pushed gently but it doesn't open.	Passive	Beginning	2	0	0	0
6.2	A bomb explodes.	Intransitive	Instantaneous		4	2	3
6.2	A cyclist rides around a circular cycle track.	Intransitive	Ongoing	2	1	2	1
6.2	The Moon orbits the Earth.	Ergative	Ongoing		0	1	1
6.1	A string vibrates.	Ergative	Ongoing		1	2	
6.0	A roller coaster goes over the top of a hill	Intransitive	Instantaneous		2	2	1
5.9	A cook chops up a cabbage.	Active	Repetitive	2	3	1	0
5.9	A pendulum bob reaches its highest point.	Ergative	Final		0	1	2
5.9	An avalanche of snow goes down a mountainside.	Ergative	Instantaneous		3	2	3
5.7	A car starts when the lights turn green.	Ergative	Beginning		2	1	3
5.7	A student spits out a watermelon seed.	Active	Instantaneous	1	3	2	3
5.6	A ball rolls down a ramp.	Intransitive	Ongoing		2	2	3
5.5	Water comes out from a hose nozzle.	Ergative	Ongoing		1	2	0
5.3	A car has a truck resting on its roof for an advertisement.	Ergative	Stative		0	0	0
5.3	A teenager chews gum.	Active	Repetitive	1	1	1	0
5.2	A ball rolls along a horizontal table.	Intransitive	Ongoing		2	1	0
5.2	A ball rolls up a ramp.	Intransitive	Ongoing		2	1	2
5.1	A rock drops 10 mm.	Ergative	Final		2	1	3
5.1	A tired shopper sits on a park bench.	Ergative	Stative	0	0	0	0
4.9	A boy sucks a milkshake.	Active	Ongoing	1	2	1	3
4.9	A car is on the back of a truck. The truck is getting faster.	Ergative	Stative		0	2	3
4.9	A metal bar expands as the temperature rises.	Ergative	Stative		2	1	0
4.8	A boy picks his nose.	Active	Ongoing	1	0	0	0
4.8	A car is on the back of a truck. The truck goes around a curve at a constant speed.	Ergative	Stative		0	2	1
4.5	A person walks down the street.	Intransitive	Ongoing	2	0	1	0
4.4	A cat jumps into the air.	Intransitive	Beginning	3	3	2	3
3.4	A car is on the back of a parked truck.	Ergative	Stative		0	0	0
3.4	An elephant sleeps.	Ergative	Ongoing	0	0	0	0
1.1	A car is driven past a parked truck. They do not touch.	Passive			0	2	0

Force and motion: classifying the situations by speed and acceleration

Since it has often been reported that students confuse force and motion (e.g.Gunstone & Watts, 1985) the sentences were classified according to:

- speed: none (coded 0, e.g. "sleeps"), slow (coded 1, e.g. "pushes"), fast (coded 2, e.g. "kicks")

- type of acceleration involved: get faster ("a car starts"), slower ("a ball rolls up a ramp"), direction-changing ("a cyclist rides around a circular cycle track"), no acceleration ("an elephant sleeps"). These were coded as 3, 2, 1, and 0 in order of their ratings as examples of acceleration (see previous chapter).

Perhaps surprisingly, particularly in the light of the evidence for a component of meaning related to motion in the first two analyses, but also because all the students had studied Newton's laws including $\Sigma F = ma$, there was no evidence to support a direct relationship between the concept of motion (whether as speed or acceleration) and the concept of force. Analysis of variance for the classification by speed ($df = 2, 47$, $F = 0.534$, $p = 0.590$) and for the classification by type of acceleration ($df = 4, 45$, $F = 1.469$, $p = 0.228$) showed no evidence for differences in ratings of situations as examples of forces where different types of motion occurred (see Figure 15). If there were confusion between the two concepts then differences in one should correlate with differences in the other. Admittedly, it is possible that such differences do occur, but that the size of the effect is small enough that the sample here was insufficient to allow them to be detected at statistically significant levels.

Although this finding (i.e. that the force-ratings of situations are unrelated to the type of acceleration) may be due to lack of statistical power, this seems unlikely. There are very clear and robust differences between the ratings of different types of acceleration as examples of *acceleration*: getting faster is clearly prototypical, getting slower has a much lower rating, and changing direction is at best rated as a marginal example. If the students' concept of physical force were closely linked with that of acceleration then one would have expected a similarly marked ordering in the ratings of situations as examples of forces. Although an upward trend is apparent in Figure 15, it is not significant statistically, and forces that cause slowing were rated marginally (non-significantly) higher than those involved in increasing speed.

Note that this lack of a relation between the concepts of *motion* and of *force* does not in itself conflict with the oft-reported association between force and motion, illustrated for example by students drawing force vectors in the direction of velocity. What it implies is that this association is not caused by conceptual confusion, but by other factors.

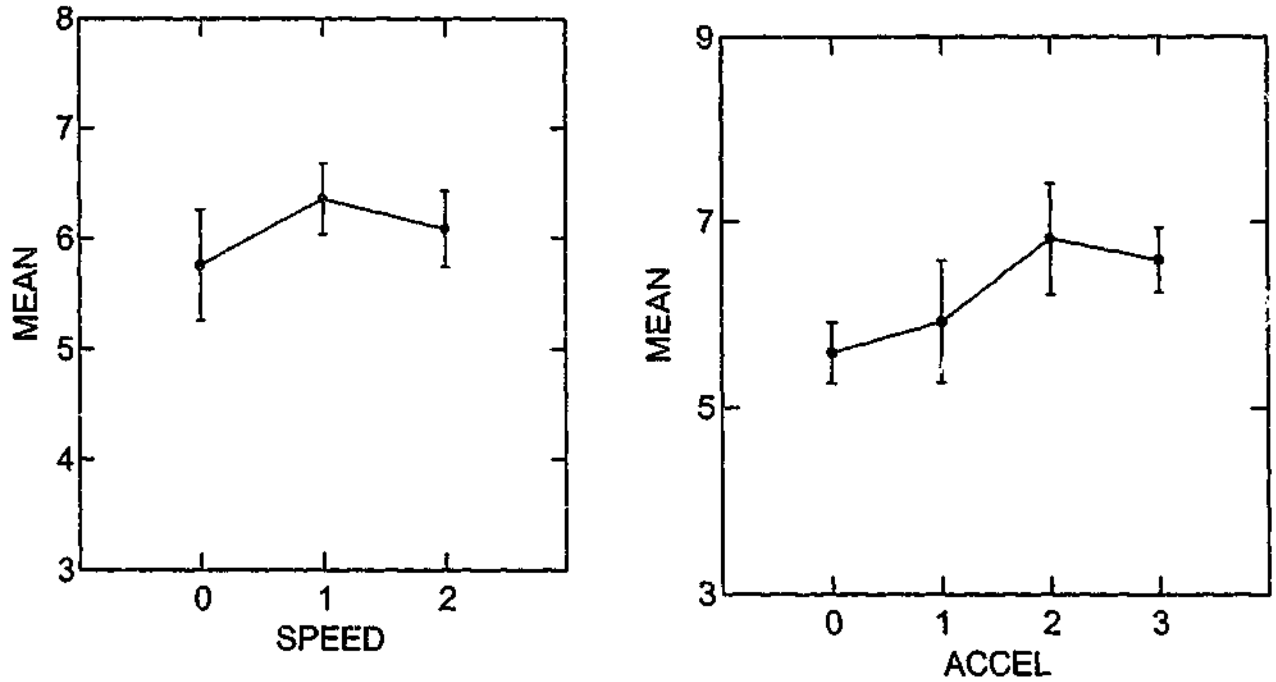


Figure 15: Mean ratings of situations classified by speed and by type of acceleration

By contrast with the above results, there was clear evidence for the relevance of the remaining variables.

Force as interaction: transitivity

Newtonian force is an interaction between two objects, not a property of a single object. To the extent that the concept of a physical force coincides with the Newtonian concept of force, situations involving two participants should be rated more highly than those involving one. Since the number of participants is coded grammatically as transitivity, the situations were classified by the transitivity of their description, as follows:

- active (where the subject of the sentence is the agent, and performs some action upon the object e.g. "player kicks ball"),
- passive (where the subject of the sentence has some action performed upon it by an agent, which need not be specified, e.g. "iron bar is twisted"),

- intransitive (the subject represents an agent that performs some action that does not require an object, e.g. "a cat jumps"),
- ergative (the subject is the undergoer not the agent, e.g. "a rubber band stretches")

A comparison of the mean ratings is shown in Figure 16. It is apparent that those situations with two participants are rated as better examples of force than those involving only one participant, irrespective of whether the agent is mentioned (active sentences) or not (passive sentences). Although the differences by grammatical form fall short of significance at the 0.05 level (ANOVA with $df = 3, 46$, $F = 2.687$, $p = 0.057$), the difference by number of participants is statistically significant (ANOVA with $df = 1, 48$, $F = 8.008$, $p = 0.007$).

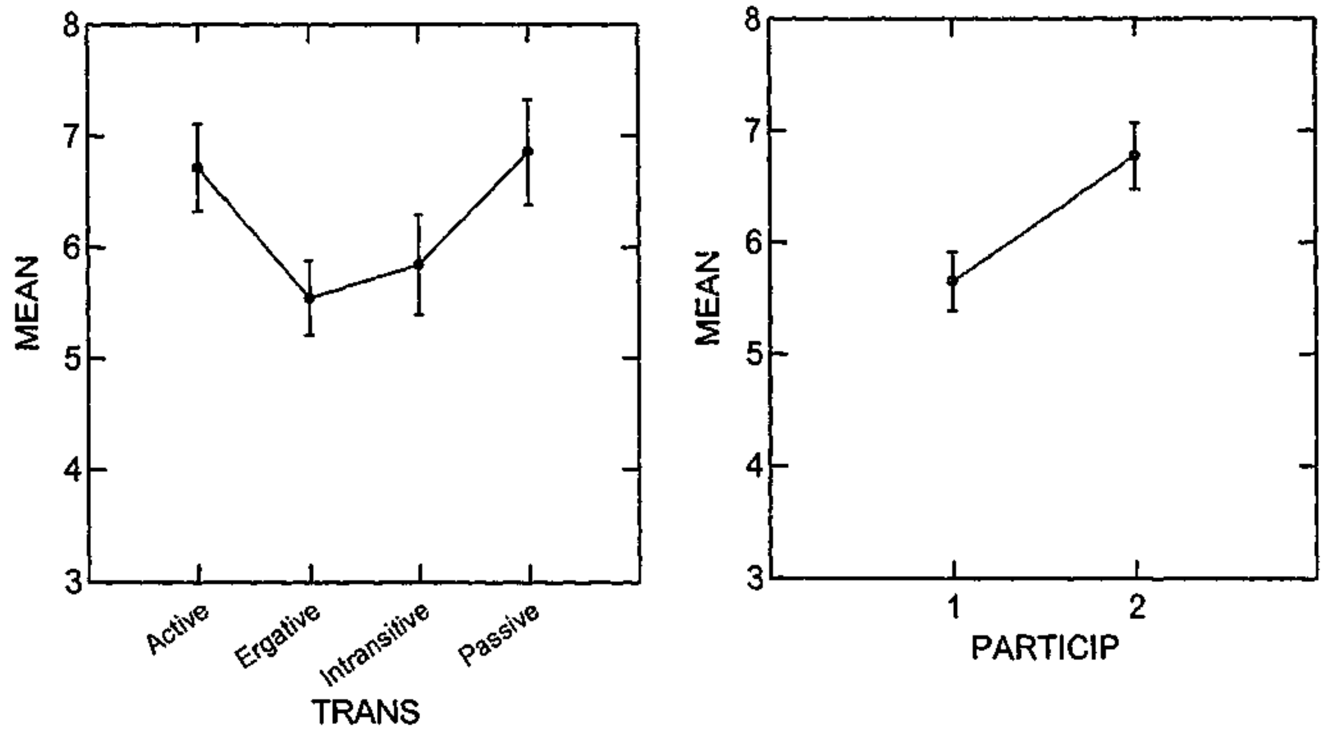


Figure 16: Mean ratings of situations classified by grammatical form of sentence (TRANS), and by number of participants (PARTICIP).

Force as an event: aspect

It was noted above that a punch was rated very highly as an example of a force, and it was speculated that this might be due to, amongst other factors, its occurrence at a particular point of time. In order to investigate this, the aspect (i.e. progression over time) of the descriptions of the situations was classified as follows:

- ongoing state (**stative**, e.g. 'sit'),
- ongoing action (**ongoing**, e.g. 'rolls'),

- repetitive action (**repetitive**,'chews'),
- action characterised by its initial stage (**beginning**, e.g. 'jumps'),
- action occurring at an instant (**instantaneous**, e.g. 'explodes'),
- action characterised by its final stage (**final**, 'reaches')

(One case was not classified because it represented a situation where nothing happened: the car passing a truck without touching it.)

The differences in mean ratings amongst these classifications are shown in Figure 17. They are statistically significant (ANOVA with $df = 5, 43$, $F = 6.218$, $p < 0.001$). There are two main groupings: those situations that occur over an extended period of time, and those that are characterised by an instant of time. The mean ratings obtained by grouping together all the instantaneous events (coded EVENT = 1) and all the things that happen over an extended duration (coded EVENT = 0) are also shown in Figure 17. The difference is clearly statistically significant (ANOVA with $df = 1, 48$, $F = 37.241$, $p < 0.001$).

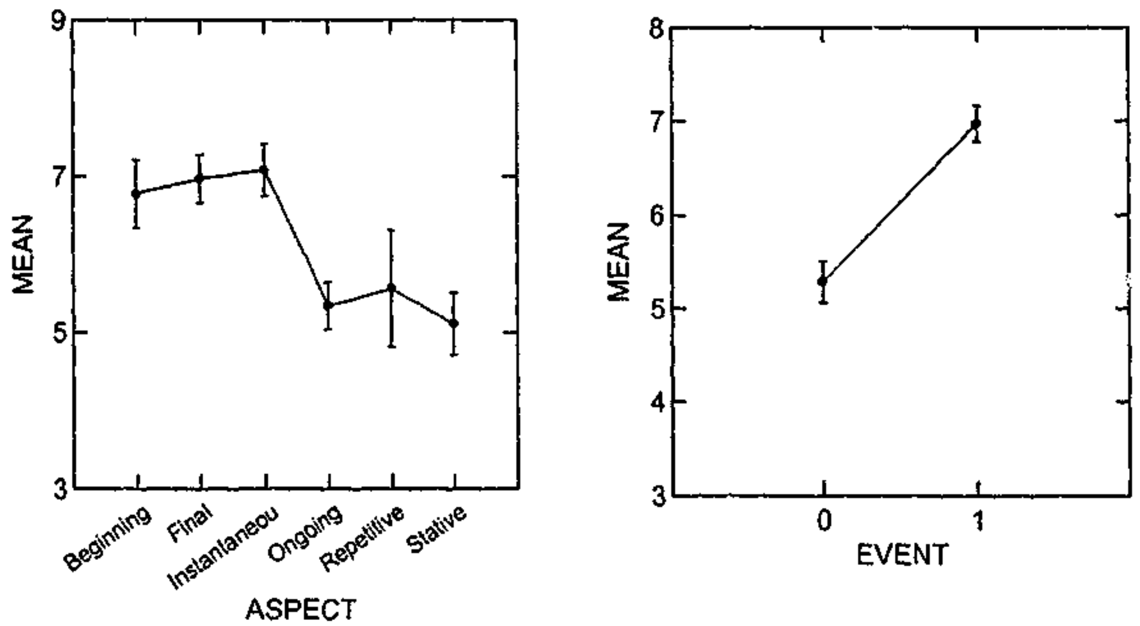


Figure 17: Mean ratings of situations classified by aspect, and when grouped as Events (Beginning, Final and Instantaneous) and non-Events (Ongoing, Repetitive, and Stative).

Force and intensity: effort and effectiveness

It was noted above that *strength* was highly rated as an example of *force*. Since strength is a property of a single participant, and is a state rather than an event this seems inconsistent with the findings reported above. It may be, however, that *strength* represents a different facet of the concept of *force*. As discussed above, strength is widely used to indicate the degree or intensity of entities: 'strong wind', 'strong acid', and 'strong argument', are just three examples from widely varying contexts. It may then be that strength relates to the magnitude of force, a facet of the meaning that is not obviously related to the number of participants or to the localisation of an event in time. To investigate the importance of the strength of the force on the ratings, the situations were classified both in terms of the degree of effort, and the degree of effect obtained.

The effectiveness, or extent of action, was classified as:

- violent (coded 4, e.g. 'explodes'),
- large (coded 3, e.g. 'falls 100 m'),
- medium (coded 2, e.g. a door is pushed and it opens),
- small (coded 1, e.g. 'suck a milkshake'),
- none (coded 0, e.g. a door is pushed but it does not open)

In so far as it is possible to exert a large force without obtaining any noticeable effect, when there is a great deal of resistance, for example, the degree of effort was also used to classify situations. Effort was classified as:

- high (coded 4, e.g. 'lift 100 kg to shoulder height'),
- medium high (coded 3, e.g. 'push a door hard'),
- medium low (coded 2, e.g. 'push a door gently'),
- low (coded 1, e.g. 'suck a milkshake'),
- none (coded 0, e.g. 'sleeps').

Where there was no agent (e.g. 'a car gets faster', 'a rock falls', 'metal expands') this classification was left uncoded. (Note that sentences where there is no agent differ from passive sentences where there is an agent, but it is omitted. Compare 'a car gets faster' with 'a car is broken.' The first of these sentences has no grammatical agent, while the second is a passive with the agent omitted. The difference is that while, logically, it is possible that a person may have been responsible for the occurrence of each of these situations, one cannot say 'a car gets faster by Fred' even though it is perfectly grammatical to say 'a car is broken by Fred.')

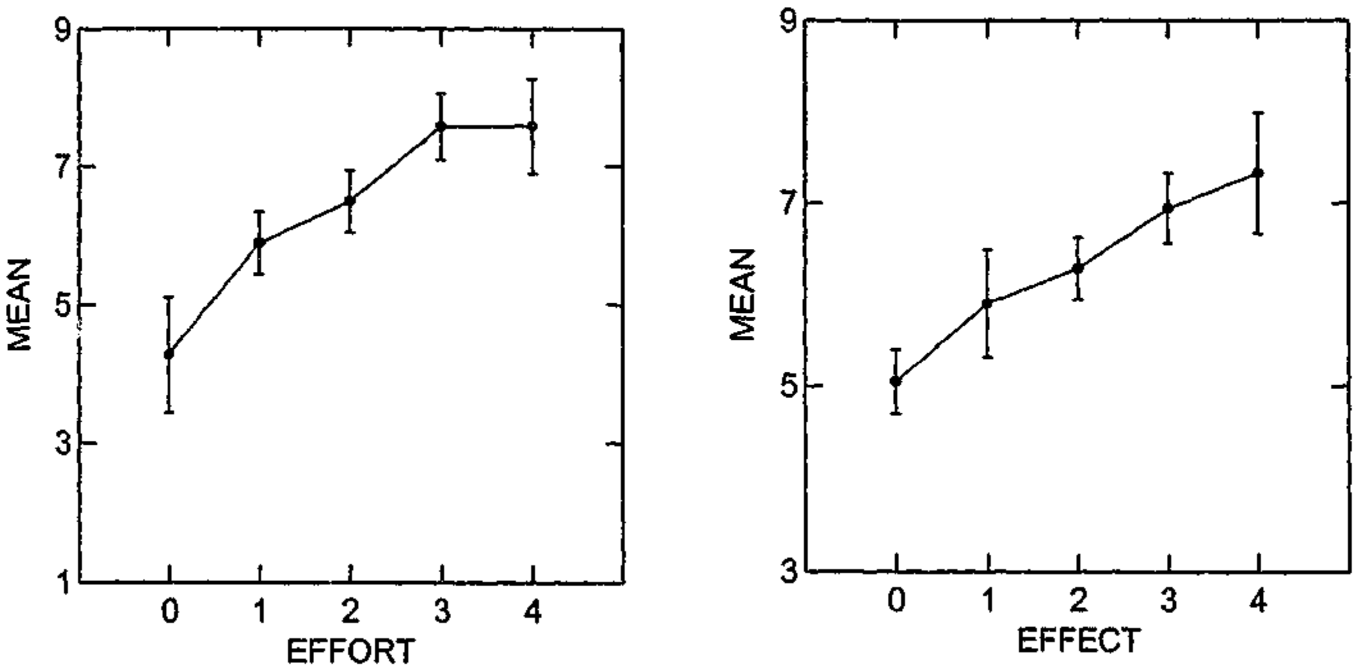


Figure 18: Mean ratings of situations classified by degree of Effort and Effect.

The results are illustrated in Figure 18 above.

There is a clea· trend for a situation to be rated as a better example of a force when either the effort or the effects were greater. The differences between groups were statistically significant at the 0.05 level. (Effort: ANOVA with $df = 4, 20$, $F = 4.080$, $p = 0.014$. Effect: ANOVA with $df = 4, 45$, $F = 4.394$, $p = 0.004$.)

An interpretation of this result, which is consistent with results that are to be reported in study four, is that *strength* is a component of the meaning of *force*. (In study four it is argued that forces are prototypically large – see p.154.)

Summary of Study 1, Part 3

Comparisons of student ratings of situations as better or worse examples of *force* indicate that a prototypical physical *force*

- involves two objects,
- occurs at an instant of time,
- requires a large amount of effort, and
- creates a large effect.

It is not possible with the data here to disentangle the relative contributions of these four components of the meaning as the components were quite highly correlated in the situations that were rated by the students. A factor analysis of the classifications of the situations presented for rating illustrates this: see Figure 19.

The situations presented to the students varied essentially along only two dimensions. Amongst the situations presented for rating, event-like situations and situations with two participants were also typically high effort situations: these all weighted on factor 1. The higher speed situations were typically higher acceleration: these both weighted on factor 2, which could be interpreted as the motion factor. And since motion is one type of result of force, the degree of effect had roughly equal weightings on both factors. Because of this one cannot disentangle their contributions to the meaning of force from this data set.

However, that the situations used essentially varied along only two dimensions, does not indicate that the classifications were just variants of one another. The classifications were not what Kline refers to as "bloated specifics" (1994, pp. 128-129): they are conceptually distinct. There is no conceptual necessity for situations with two participants to occur at an instant of time, or to involve large amounts of effort, for example. The fact that the various classifications happened to correlate is a property of this data set.

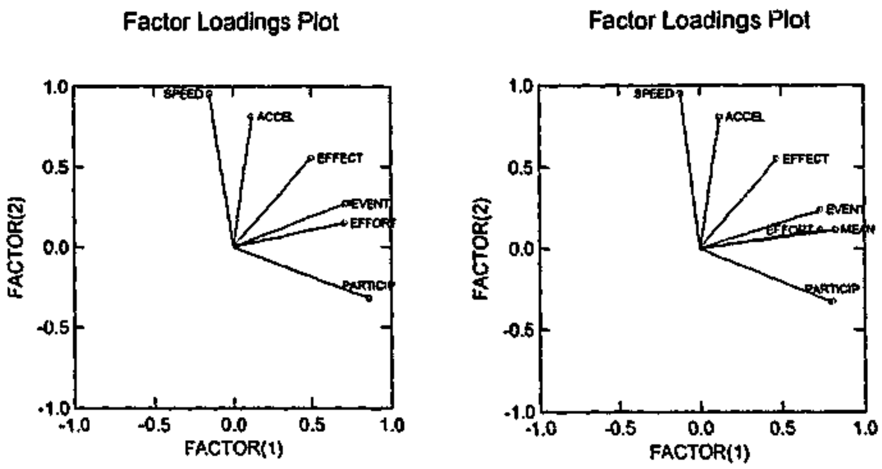


Figure 19: Factor analysis (Oblimin) of the classifications of the situations presented for rating, without mean rating (left) and with mean rating (right).

Adding the mean rating of the sentences to the list of variables to be factored helps to clarify this point, and to make clearer the meaning of factor 2: see Figure 19. Addition of the variable 'mean rating' leaves the factor structure almost unaffected. It can be seen that the mean ratings weight almost entirely on factor 2. Thus one could interpret factor 2 as the rating of *force*, which was investigated in this study.

That effort, effect, number of participants and event-status are weighted on this factor, while speed and acceleration are not, simply reiterates what has been argued above.

Study 2: The lifeworld use of the word 'force' and related terms by young children

The occurrences of words related to force were examined in a set of transcripts of discourse produced by young children. The data were obtained from transcripts on the CD ROM accompanying a guide to the CHILDES Corpus (MacWhinney, 2000a, 2000b). The original data used here were produced by many researchers over many years (Bellinger & Gleason, 1982; Bliss, 1988; Bloom, 1970; Bloom, Hood, & Lightbown, 1974; Bloom, Lightbown, & Hood, 1975; Bohannon & Marquis, 1977; Braine, 1976; Brown, 1973; Carterette & Jones, 1974; Cruttenden, 1978; Fletcher & Garman, 1988; Giles, Robinson, & Smith, 1980; Gleason, Perlmann, & Greif, 1984; Gleason, 1980; Gleason & Greif, 1983; Greif & Gleason, 1980; Haggerty, 1929; Johnson, 1986; Jones & Carterette, 1963; Kuczaj, 1976; Masur & Gleason, 1980; Menn & Gleason, ; Stine & Bohannon III, 1983).

The subjects.

The 153 children were aged from one year five months to seven years five months. The number of transcripts used for each age group is shown below in Figure 20. While the number of transcripts for six year olds (= 7) is far less than other age groups, the total word count for these transcripts was over 10000. Although transcripts from 153 children were analysed, a few children had been the subjects of detailed longitudinal study and some 64 % of the transcripts relate to just five children at different ages. At the other extreme, 21 % of the transcripts were due to single transcripts of some 136 children.

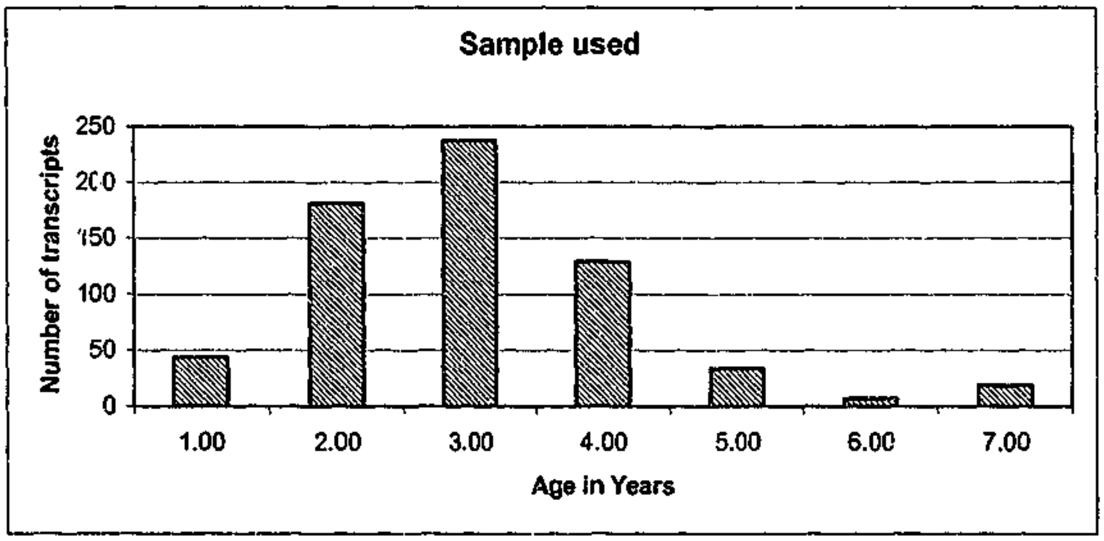


Figure 20: Number of transcripts used for each age in years

The amount of speech analysed was not evenly distributed among different age groups, but from the ages of 18 months to 64 months more than a thousand transcripts were available: see Figure 20.

The age at which those words discussed in study one are first recorded as being used is shown in Table 11.

Table 11: The ages at which force-related words are first used.

Age (months) when first used	Force-related words
18	sit, fall, pull, kick
19	hit
20	jump, run, roll, shake
21	push, chew
23	squeeze, spit
24	strong
25	squash
28	press
31	punch, chop, stretch
33	lift
38	twist
49	explode, strength

The total number of times these words were used is shown in Figure 21, and their rank order in terms of overall frequency is given in Table 12.

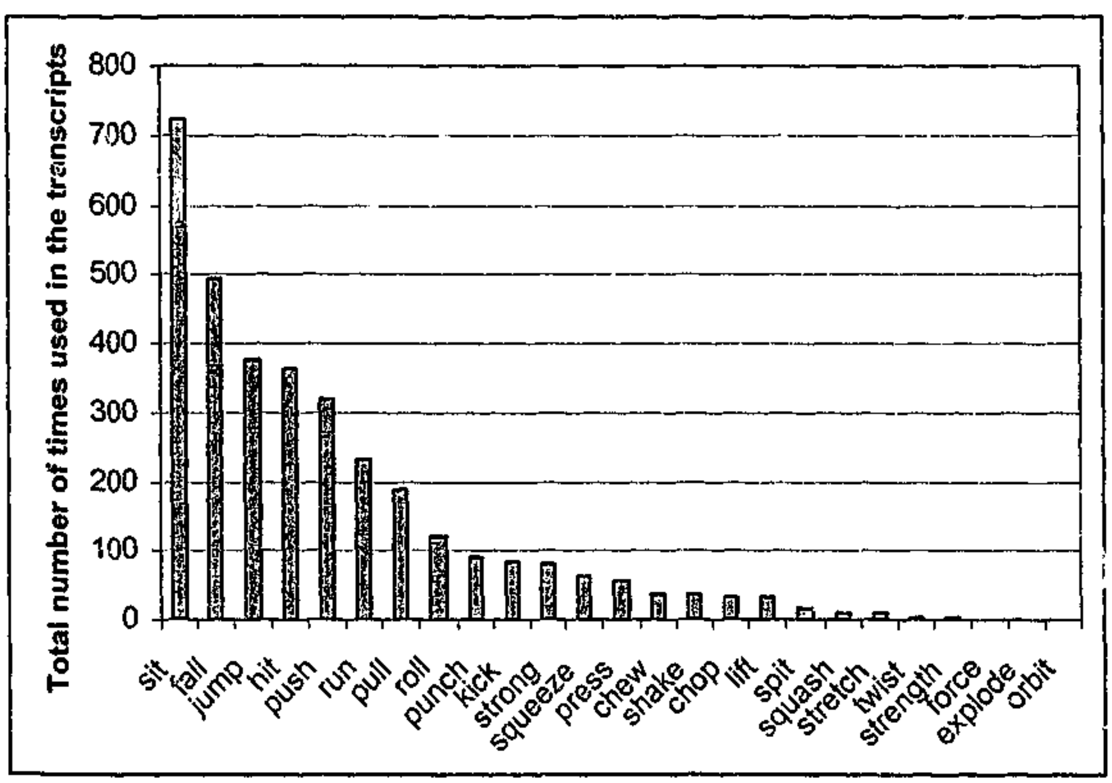


Figure 21: Total number of times words were used in the transcripts from age 17 months to 87 months.

It is apparent that the word 'force' has not entered the active vocabulary of these children before the age of seven.

It is also clear that three of the terms that have been found to be prototypical for force (push, punch, press) are amongst the earliest used and are used quite frequently (ranking between 331 and 1160 overall), while the fourth, 'strength,' is used far less often, and does not get its first use until comparatively late, although the related term 'strong' does appear early and is used frequently.

Note that it is not being argued that either an early appearance in the vocabulary or a high frequency of usage causes words to become prototypical. ('Sit' for example occurs both very early and very frequently, yet is rated very low as an example of force.) What is being argued here is that early high frequency terms are available to act as foundations, or prototypes, for later developing concepts. It may well be that terms need to be available for development of higher order concepts. For example, the late development of the nominalisation 'strength' from the adjective 'strong' may prevent young children from forming the higher-level concept *force*, although they are perfectly familiar with many examples of forces in their everyday experience.

Table 12: Rank of frequency of usage during first seven years.

Word	Rank	Word	Rank	Word	Rank
sit	161	punch	841	lift	1635
fall	203	kick	846	spit	2202
jump	262	strong	906	squash	3017
hit	275	squeeze	1077	stretch	3173
push	331	press	1160	twist	3376
run	396	chew	1517	strength	5724
pull	491	shake	1517		
roll	689	ch[illegible]	1635		

Summary of Study 2

The concept of physical force does not seem to have been attained by children under the age of seven years, although the prototypical actions which the concept of physical *force* will subsume are familiar to very young children It may be that the development of the concept of *strength* is needed before the full concept of physical *force* can develop.

On the other hand, some have argued that the concept of physical force never does develop naturally: that naturally occurring meanings are too vague to be labelled concepts.

Study 3: Concepts, proto-concepts and vague conceptions

While it is sometimes argued that the ideas referred to by everyday words like 'force' are too vague to deserve the title 'concept' (Trowbridge & McDermott, 1981), that they are merely 'proto-concepts' (or "vague and intuitive experiential notions" Deakin & Troup, 1981), there is a consensus amongst lexicographers that it is in fact possible to define both the lifeworld concept of physical force and, quite separately, the Newtonian one. Typically dictionaries list the lifeworld meanings first, followed somewhat later by the Newtonian definition of force in a separate, numbered, entry. (Similarly other meanings of 'force', such as its use together with a number as a measure of the speed of winds, or the rare, now obsolete, meaning *waterfall* also receive separate entries.)

Consider the English language dictionary-definitions of the noun 'force,' in the sense of physical force, in Table 13. It is plain, for instance, that there is a close link between the everyday concept of force and the everyday concepts of strength and power: all six definitions mention strength, and five of the six mention power. An attempt to arrive at a precise definition of the lifeworld concept of 'force' on the basis of linguistic evidence, such as dictionary definitions, is the basis of this study.

Although there are also ideas that are not common to all definitions – some like impact, or vigour, are mentioned in only one or two entries – this does not in itself prevent one from arriving at a precise definition. In fact, this phenomenon of having a prototype (a shared central core), and a more variable periphery has been argued to be a property of all meanings (see e.g. Lakoff, 1987; Rosch & Mervis, 1975).

Table 13: English-Language Dictionary-Definitions of the Noun 'Force' (Senses not Directly Related to Physical Force, Etymologies, Pronunciations, Examples of Usage and so forth are Omitted.)

A.	I. Strength, power 1. a. Physical strength, might or vigour, as an attribute of living beings (occas. of liquor). ... 2. As an attribute of physical action or movement: Strength, impetus, violence, or intensity of effect. ... 5. Physical strength or power exerted on an object; *esp.* the use of physical strength to constrain the action of persons; violence or physical coercion.	(Murray, Bradley, Craigie, & Onions, 1971)
B.	1. power; exerted strength or impetus; intense effort.	(COD, 1993)
C.	1. strength; impetus; intensity of effort	(Delbridge et al., 1997)
D.	1. strength; power; impact or impetus	(Robinson & Davidson, 1996)
E.	1. Strength, power	(Agnes, 1996)
F.	1. a. strength or energy esp. of an exceptional degree: active power: VIGOR b. physical strength or vigor of a living being... c. power to effect in physical relations or conditions.	(Gove, 1976)
G.	7. If you use force to do something or if it is done by force, strong and violent physical action is taken in order to achieve it. 8. The force with which someone hits or moves something is the amount of power that is used. 9. Someone or something which is referred to as a force has a considerable effect or influence on a situation or on people or things.	(Sinclair, 1987)

Of course, someone who doubts that everyday lifeworld word-meanings are sufficiently clear to deserve the title 'concept' might argue that writers of dictionaries have a greater command of language than the average person and that in everyday practice there is far more vagueness in usage than is evidenced in dictionary definitions. In terms of physics education, they might argue that there is evidence, for example, of confusion of force with motion (Gunstone & Watts, 1985), displacement with velocity (Trowbridge & McDermott, 1980), and velocity with acceleration (Trowbridge & McDermott, 1981). To this one could add the confusion of force with power, energy, momentum, and so on, evident in some students in Study 1, Part 1, above.

To answer such a critique requires, firstly, that one distinguish vagueness in reference from vagueness in meaning. To clarify this first point, consider a word unrelated to the issues under discussion: 'bald'. The word 'bald' could be defined quite precisely as 'being like a head that lacks hair'. Although this definition is in itself clear and precise, the word 'bald' can obviously be used not only to refer to the completely hairless person, but also to refer to a person with only one or two, or a few hundred hairs. In fact there is no exact number of hairs on one's head that qualifies one to be bald, and who (or what, eg. tyres) one calls bald will depend upon the context. The point is that a definition, to itself be precise and accurate, must correctly *specify* the indeterminacy in reference, and that if it does so successfully, then it can be used to *predict* what the word will be used to refer to in various contexts. (An analogy with statistics: the mean and standard deviation, analogous to meaning, are precisely defined and yet are very useful in characterizing a population of many varied measurements, analogous to usages.)

Hence, what one wants in a definition of the lifeworld concept of physical force, is a clear explication of the meaning that enables one to accurately predict the range of usage.

Linguistic evidence for lifeworld conceptions of force

Aristotle comes in for a fair share of criticism in science education research, and he has the dubious honour of having some common student errors tagged with his name. In his defence, however, it can be said that, whatever his errors in physics, he got at least one point correct in at least one field of study. In a translation of his *Topics* (1899), we find "It is the easiest of all things to demolish a definition, while to establish one is the hardest." The truth of this is illustrated here in an investigation of the definitions of force in a number of dictionaries. In the discussion that follows, these definitions are criticised. This should not be taken to reflect upon the abilities of the lexicographers involved, but rather upon the difficulty of their task, which, of course, goes well beyond the analysis and definition of the one particular word that is being examined here.

The key criticisms to be levelled here at dictionary definitions of 'force' are that they suffer from both obscurity and circularity – faults well characterized by Goddard (1998, pp. 26-35). The definitions in a number of dictionaries will now be examined to show that they exhibit these problems.

Obscurity

For a definition to be of use, it needs to use words that are simpler than those being defined. Not to do so is simply to delay the explanation of meaning, putting it off until the more complex words have in turn been explained.

So, for example, consider the term 'impetus' that is used in each of definitions A, B, C and D in Table 13. This is clearly a word less frequent than the word 'force'. It is hardly likely that people who did not know the meaning of force would have this term in their vocabulary.

Or consider entry G, also in Table 13, taken from a dictionary which uses a "carefully controlled defining vocabulary" (Sinclair & Coulthard, 1974, quoted from the rear dust jacket): even here use is made of the terms 'power', and 'strength', which are of comparable complexity to the word 'force.' (In fact both of these terms are present in all the dictionary definitions in Table 13, entry E consisting solely of them.) Since it is by looking up their definitions in turn that one finds that these words are of comparable complexity to 'force', this leads naturally on to the next problem with dictionary definitions: circularity.

Circularity

While it is possible to find examples where a term is defined in terms of itself (as for instance the Oxford English Dictionary does with *concept* defined in terms of *conception*), none of the dictionaries consulted did this with the word 'force'. However, all of them, without exception exhibited circular definitions (for example, where A is defined as B, B is defined as C, and C is defined as A). This sort of circularity was all pervasive, although camouflaged to some extent by the use of multiple explications within each definition. Tracking down the circularities requires some willingness to follow up the definitions of each content word in a given definition, and the patience to continue doing this for some time. Table 14 illustrates the extent of this problem using the Oxford Concise English Dictionary as an example. Consulting Table 14, one sees that force, for instance, is defined using

- *impetus*, in turn defined using *force*
- *power*, in turn defined using *force*
- *strength*, defined in terms of *strong*, in turn defined using *power*, and so back to *force*.
- *effort*, defined in terms of *exertion*, in turn defined using *exert*, which finally leads one back to *force*

Table 14: Edited definitions from the Concise Oxford Dictionary (COD, 1993). Only relevant senses are shown, and grammatical information, etymologies and samples of usage have been omitted.

Term	Definition
Force — n	1. power; exerted strength or impetus; intense effort.
Ability — n	1. capacity or power
Capacity — n	1. the power of containing, receiving, experiencing, or producing
Cause — n.	1. a that which produces an effect, or gives rise to an action, phenomenon, or condition. **b** a person or thing that occasions something.
Cause — v.tr	1. be the cause of, produce, make happen.
Change — n	1. **a** the act or an instance of making or becoming different. **b** an alteration or modification
Change — v	2. *tr. & intr.* undergo, show, or subject to change; make or become different
Consequence — n	1. the result or effect of an action or condition.
Different — adj	1. unlike, distinguishable in nature, form, or quality (from another 2. distinct, separate; not the same one (as another).
Drive — v	1. urge in some direction, esp. forcibly. 2. **a** compel or constrain forcibly **b** force into a specified state 5. **a** force (a stake, nail, etc.) into place by blows
Effect — n	1. the result or consequence of an action etc.
Effort — n	1.strenuous physical or mental exertion. 2. a vigorous or determined attempt. 3. *Mech.* a force exerted.
Energy — n	1. force, vigour; capacity for activity. 3. *Physics* the capacity of matter or radiation to do work. 4. the means of doing work by utilizing matter or radiation.
Exercise — n.	1. activity requiring physical effort, done esp. as training or to sustain or improve health.
Exercise — v	1. *tr.* use or apply (a faculty, right, influence, restraint, etc.). 2. *tr.* perform (a function). 4. *tr.* a tax the powers of.
Exert — v	1.exercise, bring to bear (a quality, force, influence, etc.). 2. *refl.* use one's efforts or endeavours; strive
Forcible — adj	done by or involving force; forceful.
Haste — n	1. urgency of movement or action. 2. excessive hurry
Hasten — v	1. *intr.* (often foll. by *to* + infin.) make haste; hurry. 2. *tr.* cause to occur or be ready or be done sooner
Hurry — v	1. move or act with great or undue haste. 2. *tr.* cause to move or proceed in this way. 3. *tr.* hasty; done rapidly owing to lack of time.
Hurry — n	1. a great haste.
Impede — v	retard by obstructing; hinder
Impel — v	1. drive, force, or urge into action. 2. drive forward; propel
Impetus — n	1. the force or energy with which a body moves. 2. a driving force or impulse.
Impulse — n	1. the act or an instance of impelling; a push.
Issue — n	4. a result; an outcome
Move — v	1. *intr. & tr.* change one's position or posture, or cause to do this. 2. *tr. & intr.* put or keep in motion;
Occasion — v	1. be the occasion or cause of; bring about esp. incidentally. 2. cause (a person or thing to do something).
Outcome — n	a result; a visible effect.
Power — n	1. the ability to do or act 11. the capacity for exerting mechanical force or doing work 14. a mechanical force applied e.g. by means of a lever. 15. *Physics* the rate of energy output.
Produce — v	4. bring into existence.

Term	Definition
	5. cause or bring about (a reaction, sensation, etc.).
Push — n	1.the act or an instance of pushing; a shove or thrust.
	2. the force exerted in this.
	3. a vigorous effort
Push — v	1. *tr.* exert a force on (a thing) to move it away from oneself or from the origin of the force.
	2. *tr.* cause to move in this direction.
	3. *intr.* exert such a force.
	4. *intr. & tr.* a thrust forward or upward.
	5. *intr.* move forward by force or persistence.
	8. *tr.* urge or impel
Resist — v	1. *tr.* withstand the action or effect of; repel.
	2. *tr.* stop the course or progress of; prevent from reaching, penetrating, etc.
Resistance — n	1. the act or an instance of resisting; refusal to comply.
	2. the power of resisting
	4. the impeding, slowing, or stopping effect exerted by one material thing on another.
Result — n	1. a consequence, issue, or outcome of something.
Retard — v	1. make slow or late.
	2. delay the progress, development, arrival, or accomplishment of.
Slow — adj	1. a taking a relatively long time to do a thing or cover a distance. **b** not quick; acting or moving or done without speed.
Strength — n	1. the state of being strong; the degree or respect in which a person or thing is strong.
Strong — adj	1. having the power of resistance; able to withstand great force or opposition; not easily damaged or overcome
	6. capable of exerting great force or of doing much; muscular, powerful.
	7. forceful or powerful in effect
Thrust — v	1. push with a sudden impulse or with force
Urge — v.tr.	1. drive forcibly; impel; hasten
Use — v	1. cause to act or serve for a purpose
Vigour — n (US vigor)	1. active physical strength or energy.
Work — n	1. the application of mental or physical effort to a purpose; the use of energy.
	9. *Physics* the exertion of force overcoming resistance or producing molecular change

Of course, many other terms are used in these definitions, and these too need to be followed up. Fre 22 provides a summary overview. Note that the fact that the diagram fits onto a single page should not mislead one into thinking all the cycles have been tracked down: the process was simply halted when space on the page ran out. Arrows point from words to be defined to words used in the definition. Dotted lines point toward new terms, while solid lines point back to terms that have already been used. Thus, for example, Force is defined (dotted line) in terms of power, which is in turn defined (solid line) in terms of force. (See Table 14 for details.)

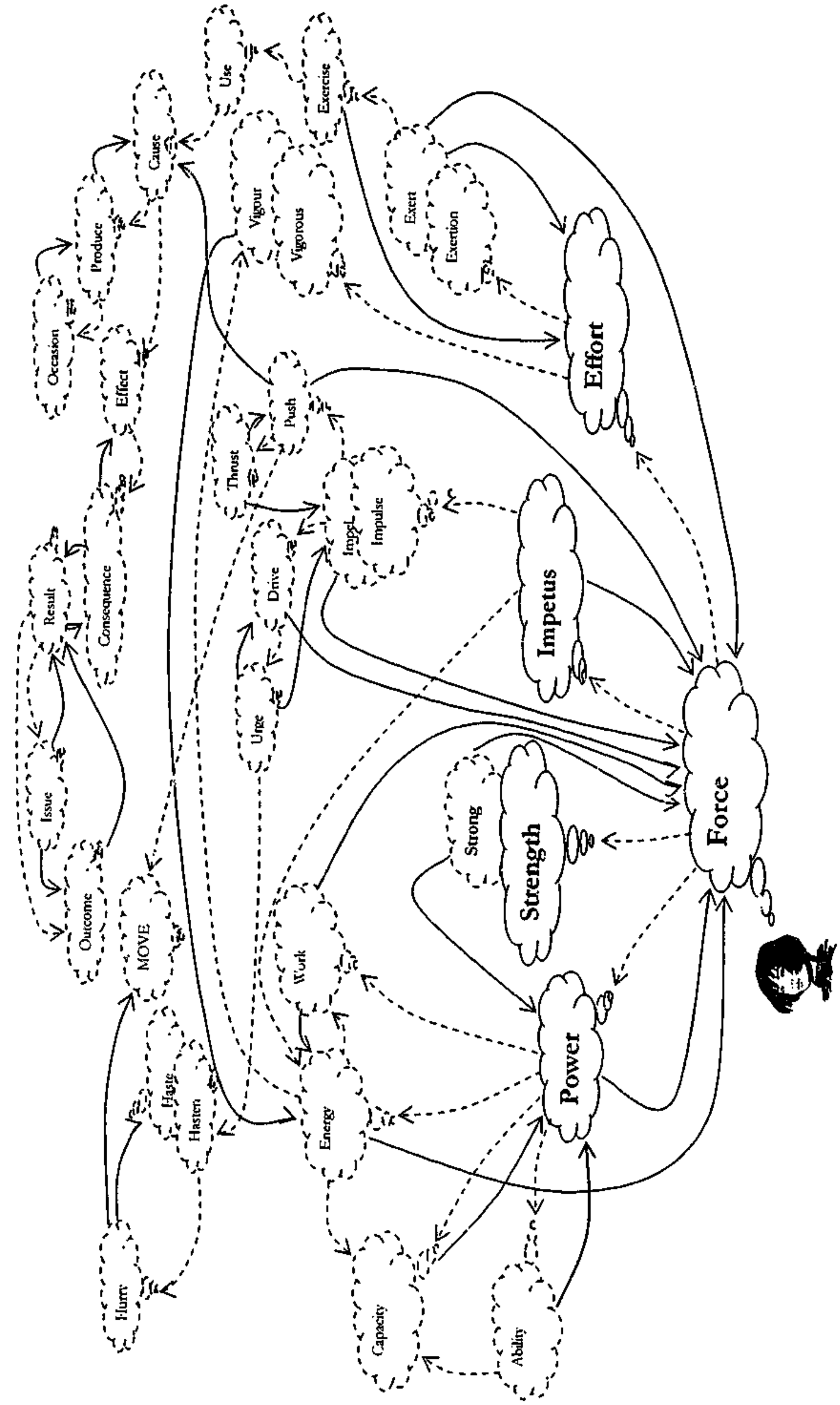


Figure 22: English Words used to define force

Very similar pictures are obtained by following up the definitions used in other English language dictionaries, although the resulting diagrams will not be reproduced here.

Summarizing: these dictionaries do not fully define the lifeworld sense of the English word 'force': the explications rely on our knowing other concepts that are of the same level of complexity. The most one could claim for these dictionary definitions is that they allow us to locate the place this concept, *force*, occupies amongst other concepts in our mental lexicon.

Languages other than English

A question that naturally arises from the above discussion is that of the extent to which this intertwining of concepts is a peculiarity of the English language, and the extent to which this intertwining is a property of the concepts themselves. That it is the concepts themselves that are inextricably intertwined is the position argued by the philosopher Jacques Derrida. He is well known for his argument that all words are defined in terms of their differences from other words, so that one can never arrive at the meaning of anything, meaning being endlessly deferred (Derrida, 1976). Whatever the truth of Derrida's position in general, the evidence obtained by investigating words for force in dictionaries, at least, seems consistent with his views.

Consider Derrida's own language, French. Table 3 consists of the definition of force – from the Larousse de Poche (1954) – and the definitions of all the words used in the definition of 'force.' The English language glosses come from Harraps Shorter French and English Dictionary (Mansion, 1977). A quick glance at the table is sufficient to establish that again there is extensive circularity within the definitions, and this is (partially) illustrated by Figure 23. It is interesting to note, also, that the French words involved in the French definitions, correspond to words that turned up in the English definitions.

Table 15: Definitions from *Larrouse de Poche*, (1954)

French Word	English gloss (Mansion, 1977)	Definition from Larrouse de Poche (1954)
Force	Strength, force, might, vigour	Vigueur physique, énergie vitale. Puissance capable de produire un effet... Puissance d'impulsion... Énergie, fermeté
Vigueur	Vigour, strength	Force physique. Energie du caractère. Puissance d'esprit. Autorité effective.
Physique	Physical	Matériel...Qui a rapport à la matière. ...Qui s'appuie sur une observation des sens.
Énergie	Energy. Force, vigour.	Puissance, force physique. Vertu, efficacité.
Puissance	Power	Autorité... Pouvoir... Domination... Force, influence
Capable	Capable	Qui peut faire une chose
Produire	To produce, bring forward, adduce (evidence etc.)	Engendrer, porter...Rapporter. Occasionner. Présenter.
Effet	Effect, result	Résultat d'une cause. Acte d'une agent. Réalisation, exécution. Impression...
Impulsion	Impulse	Mouvement communiqué

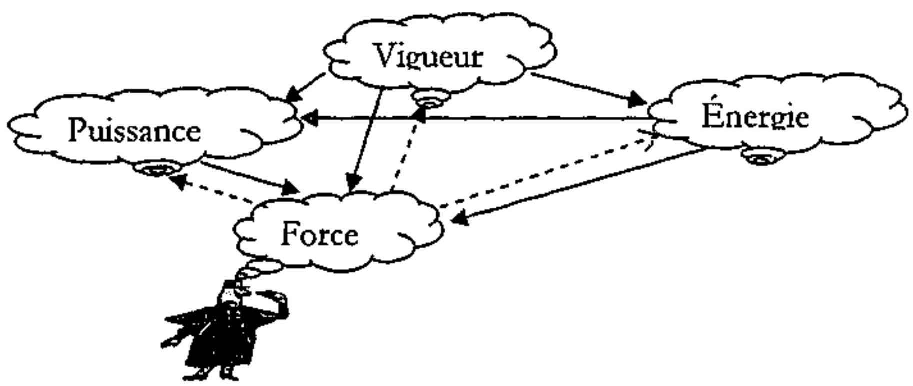


Figure 23: French words used to define 'force'.

The French language has shared a lengthy history of contact with English. It might be that the similarities are due to this.

A more distant language is Bengali. Yet one also finds the same circularity in definitions in *Chalantika Modern Bengali Dictionary* (Basu, 1942), illustrated in Figure 24.

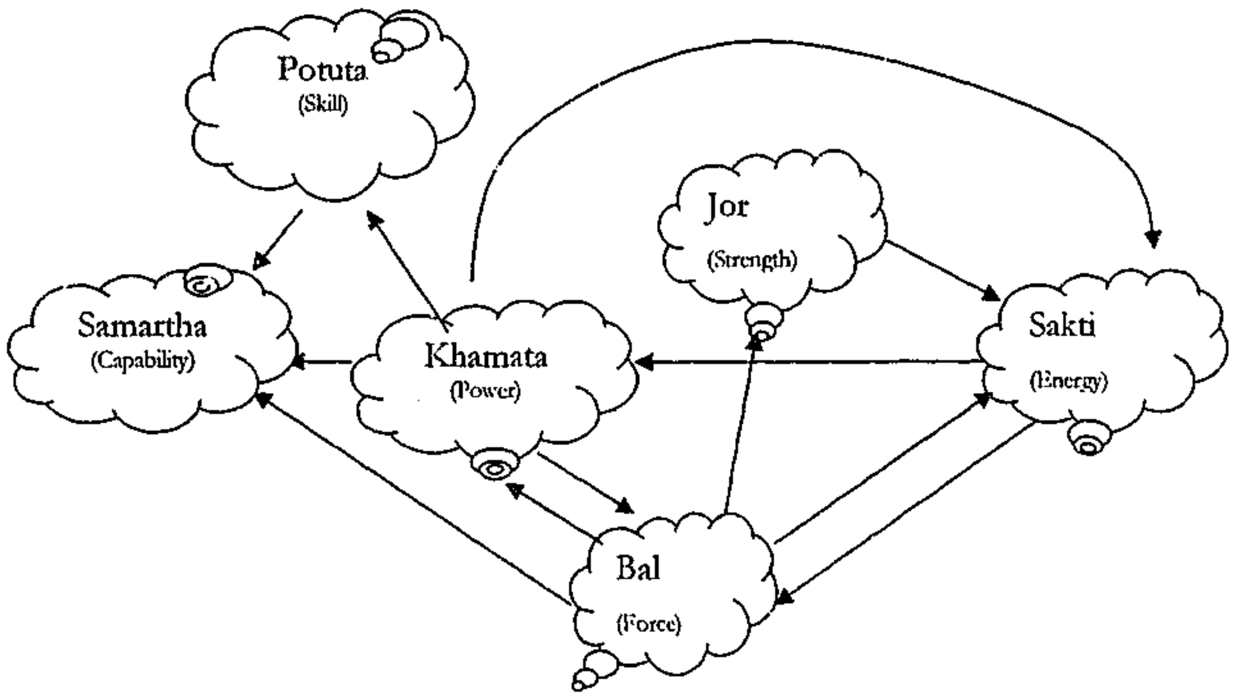


Figure 24: Bengali words used to define 'force'

Approximate English glosses are supplied in brackets.

These Bengali definitions of the word 'Bal' (meaning *force*) refer to certain concepts (*power*, *energy* and *strength*) which were also present in the English and French dictionaries. Hence it is clear that, cross culturally (at least within some groups speaking Indo-European languages), people have a qualitative understanding of the concept of physical force, recognisable in the definitions quoted from various dictionaries above, which is conceptually prior to any technical definition, such as that of Newtonian force.

Study 4: defining the lifeworld sense of physical force using the Natural Semantic Metalanguage

The mathematical definition of Newtonian force is quite straightforward in the case where there is only a single force acting: it is expressed Newton's second law, $F = ma$, here F is the force which is acting on an object of mass m, which has a as its acceleration. Although this equation provides a mathematically precise way of measuring force, the concept of that which is being measured is one that we identify by our prior qualitative understanding. As shown in the investigation into their use of words, children are well aware of prototypical forces a long time before they use mathematics.

In addition, the (very limited) examination of cross-linguistic data attempted above has shown that physical force is conceived in similar ways by (at least three) different cultures, and other work (Talmy, 1985), discussed toward the end of Chapter Two, also points toward the same conclusion: the idea of physical force is cross-culturally the same.

Admittedly, the definitions given by dictionaries all clearly suffer from circularity, but if one is not to abandon scientific understanding to some sort of Derridean play of signifiers, it should be possible to define the lifeworld idea of physical force, in such a way that it will make sense cross-culturally, using some solid foundation. As an empirical exercise, one can attempt to do this using the Natural Semantic Metalanguage (Wierzbicka, 1996). If it proves possible this has two consequences. Firstly, it will indirectly provide a small amount of additional backing for the theory of the NSM, in so far as it will have proved its usefulness in a field well outside its original context. Secondly, it will provide a clear indication of the meaning of force which students use as a foundation for their studies in mechanics.

It will be recalled that Wierzbicka (1998) defined the verb 'to force' using the NSM. Her definition is reproduced below for convenience. This definition is related to the interpersonal meaning of force (as in "force a student to do homework") rather than to physical force.

Person X forced person Y to do Z (e.g. to apologise)

X wanted Y to do Z	(a)
X knew that Y didn't want to do this	(b)
X thought that if X did something to Y then Y would have to do Z	(c)
because of this X did something to Y	(d)
because of this Y had to do Z	(e)
Y wouldn't have done Z if X had not done this to Y	(f)
when Y was doing Z, Y thought I don't want to do this.	(g)

This definition can be seen to be appropriate for many occurrences of the word 'force' in common usage. The following three illustrative examples, which are taken from the British National Corpus, are typical:

1. "This is not simply to force people into speaking blank verse, but to see how a person responds to the essential humanity of a character — for Shakespeare, of all the classical writers, is probably the most human, whose work is blessed with both grandeur and the common touch."
2. "I grow bored with the sheer size of the glass and have to force myself to continue, he wrote."
3. "You can't force her."

In each of these there is something (*speaking blank verse* in example 1, *continuing* in example 2, but unspecified in example 3) which someone wants someone else to do (although only metaphorically someone else in example 2), and something has to be done (or, as in example 3, cannot be done) to achieve this aim.

As a first attempt, one can modify this definition to define force, in the sense of physical force, by removing any references to the humanity of the thing affected by the force in the body of the definition. This involves removing any references to Y doing things, Y thinking things, or Y wanting things. The result reads as follows:

X forced something (Z) to happen to Y (e.g. to break, or to accelerate)	
X wanted Z to happen to Y	1
X knew that another thing W would happen to Y	2
X thought that if X did something to Y then Z would happen to Y	3
because of this X did something to Y	4
because of this Y did Z	5
Z wouldn't have happened to Y if X had not done this to Y	6

This definition conceptualises a force exerted by a person as being a goal directed (line 1), deliberate (line 3), and effective (lines 5 & 6) action (line 4), causing a change to what would otherwise have happened (line 2). Clearly these features are present in some usages, such as these two, again from the BNC:

1. "I found I could **force** the steering wheel a quarter of a turn and flex all sorts of things, like the casting for the gearbox mounting."
2. "He puffed softly, then strongly, to **force** the smoke from the smouldering fungus into the skep."

While these two examples are clearly related to the notion of physical force, as opposed to social force, they are nevertheless in the form of the verb 'to force' rather than in the form of a noun.

One way of rewriting the definition to make force a noun gives:

X is a force

X wants something (Z) to happen to Y	1
X can do something to Y	2
because of this, something (Z) can happen to Y	3
Z does not happen to Y if X does not do this to Y	4

In this case, however, force is characterised as an agent (X wants..., X does...). As such, it does not describe a Newtonian force, although this definition might well be appropriate for usages such as these (again from the BNC), where a force is identified with an entity – in these cases, respectively, "The Galleries", "She", or "Helena":

1. "I see The Galleries as a focal point, a powerful **force** aiding the city, a proper part of the centre designed to enhance its life and character."
2. "She is becoming Burma's most famous non-person, much to the anger of the military, who were hoping she would fade away as a political **force**."
3. "It is in the second half, when Helena becomes a quasi-symbolic **force** for life and renewal that Patricia Kerrigan's performance comes into its own."

The above definitions of 'force' could be characterised as theoretical, as they do not proceed from an encounter with evidence in the form of naturally occurring utterances, but rather from modifying the definition of the same word used in a different way. In order to make further progress in this attempt to characterise the life-world sense of 'force' in the sense of physical force it is important to examine its usage: this can be done by consulting linguistic corpora.

A large sample of usages of 'force'

The British National Corpus (1994) contains approximately 100 million words and can be queried to provide examples of usage for any word. Querying it for the word 'force' one finds that it contains 15845 instances in 2460 texts. Of these, 'force' is used as a singular noun 9838 times in 2031 texts. To reduce this to a manageable level, the BNC was queried to provide a random sample of 1000 usages of the word 'force' as a singular noun.

When these were examined they could be sorted into nine main categories

Groups of people can be called a force. For example, military or quasi-military groups such as the armed forces, naval forces, the police force, peace-keeping forces, combat forces and so on. Also, other groups with a common aim, such as a task force, sales force or teaching force; and workers, such as labour force, work force, or force of production.

Legal force: laws can be in force, or come into force, and they possess the force of law, or statutory or regulatory force.

Communicative force: in this group is found the force of an argument, logical force, the emotional force of writing, the force of music, or visual force of a painting, and so on.

Forces of nature: including, as well as the phrases 'force of nature', 'natural force' and 'physical force', examples like the wind or gale, tides and waves; as well as pushes, pulls, squeezes, and pressures. As this is the category in which we are most interested, further description will be provided below. It should be noted that although this category includes within it examples which fit with the physicist's notion of *physical force*, it also includes many metaphorical and other usages which do not fit this notion.

Influences: including the force of tradition, or of will power; the force of destiny, blind force, inner forces, moral force, and religious force. As well, there are forces to be reckoned with, dominant forces and driving and motive forces.

Violence: here one finds the use of force, force of arms, excessive force or reasonable force, and brute force, shows of force and the threat of force.

Effects: in this category one finds forces described by what they do. These include forces of abrasion, a shattering force and a force for rehabilitation. This category is sometimes hard to distinguish from the next category. The difference is between whether the effect is something which just happens (effect), or whether it is in some sense aimed for (goal).

Goals: forces described by their aim. These include various "forces for...," such as forces for change, forces for emancipation, and forces for stability. There are also forces of unification, guiding forces and leading forces.

Personality: an individual can be called a force. One can be a dynamic force, a "one woman shock-force," an active force, or on the other hand one may be not the force one was. A person also can possess force of character, force of personality or force of will power, and be subject to a force of repression or the force of habit.

Some passages could not be classified because they were too short: for example, "the existing force that was there." Others were insufficiently clear, for example:

> If an honest attempt is made in one's own little world to make this strive *(sic)* a living force without being crushed under foot by one's contemporaries, then one may consider oneself and the community to which one belongs lucky.

The nine categories are not represented in the BNC sample with equally frequency. Of the 978 instances which were classifiable, by far the most common usages were related to military or quasi-military groups. The distribution of these categories is illustrated in Figure 11.

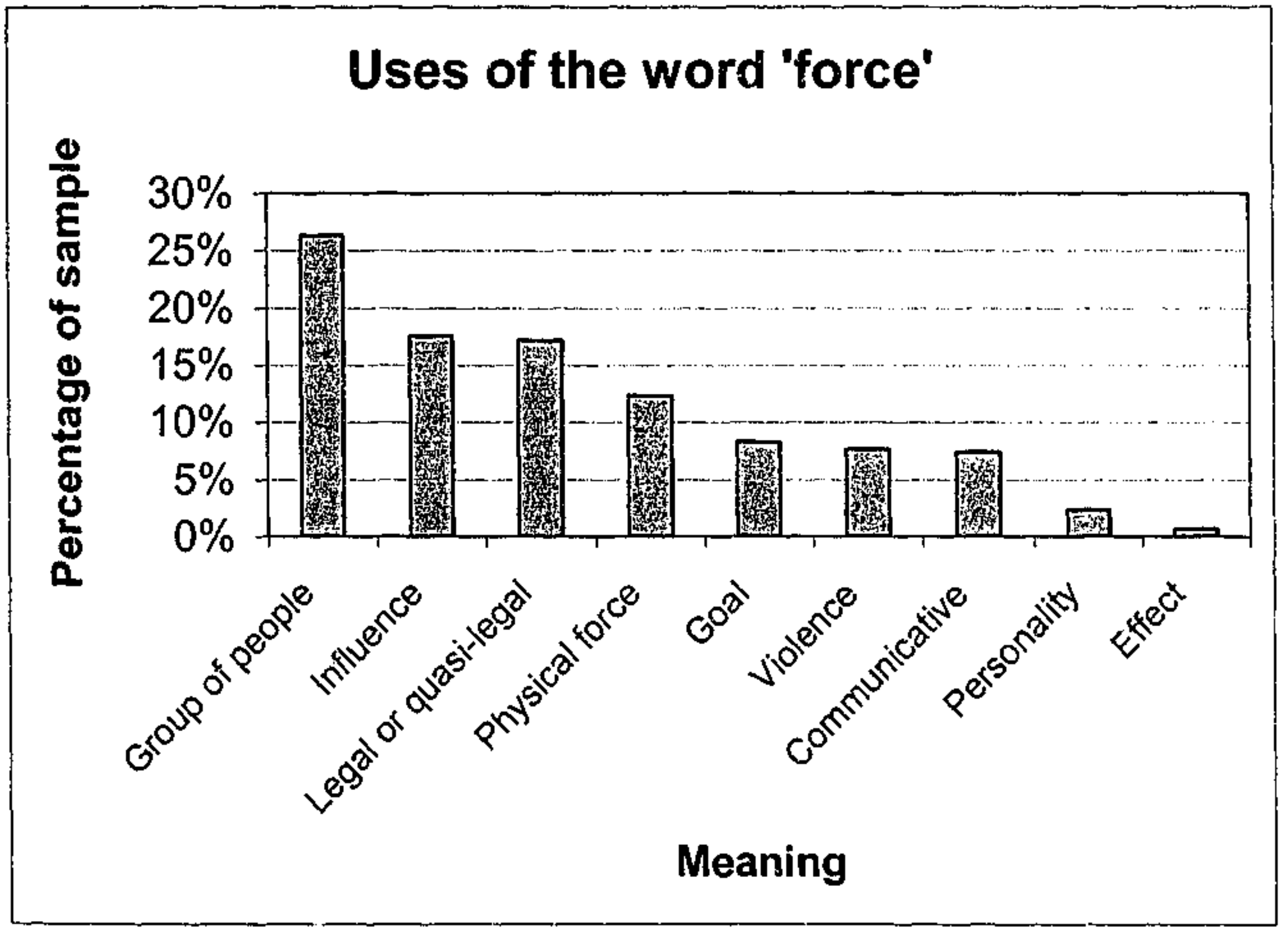


Figure 25: Percentages of 978 utterances using 'force' in different categories.

Usages related to the physicist's notion of physical force

It has already been shown, in Figure 25, that utterances which fall under the category of *physical force* represent only a small proportion of the total. However, the initial classification reported above was based on the form of the words: examining the sample more closely in terms of the meaning of the words shows that the initial classification is overly inclusive. Certainly the Figure 25 does not to underestimate the frequency with which the concept of physical force is used within lifeworld contexts. However, when one examines the 111 instances classified under this heading more closely, one finds that only a proportion of these in turn are related to *physical force* in the physicist's sense (apart from those taken from technical contexts).

Technical usages – 35 instances

Technical usages were identified on the basis that they came from within a particular specialised context: there were 35 instances of these which fell under the rubric of physical force. Some examples:

- If the rotor is slightly displaced from the step position a force is developed between the stator and rotor teeth (Harris et al., 1977) giving a torque which tends to return the rotor to the step position; a rotor displacement in the negative direction produces a positive torque and a positive displacement results in a negative torque.
- Cohesion, the force holding the water molecules together, causes more water to move into the leaf cells from the xylem.
- The plastic plate was suspended 3–5 mm above the abdominal wall from a force displacement transducer (Grass FT3C).
- Simply pouring the beer out from a can did not generate the force needed to initiate the gas break-out that creates the head.
- Energy is officially defined as 'capacity for doing work' and it has the dimensions of force multiplied by distance.

Clearly none of these technical usages will form part of the lifeworld meaning of *force* which is being sought here. They occur within particular contexts which would only be encountered after the lifeworld concept had developed. There remain only 76 instances (111 less 35) which might represent the lifeworld concept of *physical force*.

Natural forces – 36 instances

Of these, one might expect that the phrase 'physical force' would approximate the physicist's notion *physical force*, but this is in fact clearly not the case. Consider the four instances where this phrase was used.

- Maybe hatred could grow so strong that it became a force of its own, he thought — a real physical force.
- Violence 'involves the use of great physical force or intensity and, while it is often impelled by aggressive motivation, may occasionally be used by individuals in a mutual violent interaction which is regarded by both parties as intrinsically rewarding' (Siann 1985: 12). (*sic* – the reference forms part of the quotation)
- If we were asked to offer a definition, we would probably suggest something along the lines of 'the use of physical force against an unwilling party'.
- The constitution helps to shape politics and also regulates public access to, and the behaviour of, the various institutions of the state (such as the Cabinet, the House of Commons, the civil service, the military, the judiciary, and the police) that in their different ways have the right to exercise public power, if necessary through the use of physical force, within the United Kingdom.

The first is plainly metaphorical. The remaining three instances are best classified under the heading 'violence.' Even though they explicitly use the phrase 'physical force,' they do not approximate the idea of *force* as used in physics. In fact, the last three are really examples of the expression 'use of physical force,' which is a minor variation of 'use of force' – classified under 'violence' in the section above. The use of the word 'physical' is of little import. There remain 72 instances to be examined.

If the phrase 'physical force' is not used in the sense being looked for, then perhaps the phrase 'force of nature' or one of its variants might be (there are four such instances in the sample studied here). After all the word 'physics' derives from the Greek word for 'nature' (Lewis, 1961):

- People love to be awed when they enter a pub by a superior natural force – a strange sort of higher masochism.
- Christabel makes it into a force of nature.
- Silhouetted against the lightening sky, it seemed more majestic than ever; more a part of nature's raw force than the work of man.
- Public attention was focused on the force of nature as never before; successful campaigns were launched to foot the vast bill for the clear-up operation.

The first two of these are metaphorical usages. The third is also metaphorical, although one needs to read the passage in context to be certain of this: the "it" which is said to be "part of nature's raw force" is in this case a burial barrow.

Only in the last of these is there some connection with *physical force* in the physicist's sense: in this last quotation the topic of the discussion is a severe storm, which is something capable of exerting force. (In an aside, although these four quotations are not all linked to the physicist's notion of physical force, they are nevertheless linked together by the writer's attitude – one of "awe", toward the "majestic", or "vast" – towards the 'force of nature.' This is consistent with the hypothesis that it is part of the notion of a 'force' that it is, in some sense, big. This point will be returned to in the section beginning on p. 155.)

However, once the instances examined so far have been eliminated, the remainder contain a higher proportion of relevant usages. For example various natural phenomena – specifically light, gravity, and water or air in motion – are described in terms of force. There are 28 such instances in the sample and they will be discussed in turn, in ascending order of frequency.

Light – 1 instance

- Outside, the sun still shone, though with the approach of evening the light had lost some of its force and the room grew dimmer.

Although light is described as having *force*, this is as a way of describing its intensity, not the *physical force* which it can exert.

Gravity – 2 instances

- I think my race across the car park and the force of gravity just made the baby slip out!
- The time you eight, nine, ten, eleven, twelve hours later when you go to bed because the force of gravity is hasn't it to. *(sic)*

Although the second of these is taken from speech, and is somewhat hard to interpret, it becomes easier when one sees the context, which is that the speaker is arguing you are taller after a night's sleep. In both these examples, then, the force of gravity is a clear example of a lifeworld sense of *force* which is related to the physicist's sense of *force*. Note, however, that the phrase used is the 'force <u>of</u> gravity'. In these examples, gravity is not necessarily being identified as a force. Force could well be something gravity <u>has</u>. These usages can be contrasted with the same phrase amongst the technical usages mentioned above: there were seven of these, like this one:

- They are called the strong force, the weak force, the electromagnetic force, and the force of gravity; and the greatest prize in theoretical physics today is to uncover the underlying symmetry which has to exist between them.

In the technical usages it is generally plain that the 'force of gravity' is the name for a particular sort of force, while in the lifeworld usages this is not necessarily the case. This point can be made clearer by looking at parallel phrases where the expression 'force of X' does not necessarily indicate the name for a type of force.

Force of moving water – 4 instances

The expressions 'force of the water,' 'force of the downpour,' 'force of the waves' and 'force of the bore' are all used in the examples below, and all clearly have some relation to the sense of *physical force.* Compare these usages to 'force of gravity.' Gravity is not a tangible thing in the world, but water, and so on, are. If one reads the above expressions as naming forces, in the same way that 'force of gravity' is the name of a force, then one is committed to the position that lifeworld forces can be tangible things.

- They tried pushing it back into the hole but the force of the water was too great.
- When Kamacrli came back with the cows only minutes later, she was caught in the full force of the downpour.
- You are floating along a quiet river now, you don't see the water boiling at the foot of the great rocks, but one day you'll come to a point in life's stream where the wild force of the waves may destroy you, where the noisy rushing water may drown you!
- Rivermen were surprised at the force of the bore on a 9.3 metre tide, nearly a metre below maximum at Sharpness.

Force of moving air – 21 instances

Again, here, one finds that many of these instances are not related to the sense of *physical force* which is relevant. Of these 21 instances, 11 used force as a description of the speed of the wind. Although *speed* of the wind is indirectly related to *physical force,* since faster winds can exert greater forces, this type of usage is sufficiently formalised that this connection is irrelevant: a 'force eight gale' is simply a body of air travelling within a certain range of speeds irrespective of whether it exerts forces on things in its way or proceeds unimpeded over an empty seascape. Even though this is a specialised sense of the word, it was not classified into the Technical subcategory since it is likely to be encountered by anyone who listens to a weather report. These are two typical examples:

- Gales of force eight gusting to storm force ten can be expected in the area well after dawn.
- COASTGUARDS slammed organisers of a race in Swansea Bay after 65 teenagers were rescued when gale force gusts capsized their dinghies.

Apart from this specialised usage there were another ten instances where there were references to the force of moving air. Below are listed 5 instances which used the phrase 'the force of the wind.' Of these the last is metaphorical, and the first four are not able to solve the problem of whether 'force of the wind' is the name of a special type of force, or whether it indicates a property of the wind. The first two quotations tend to support the first hypothesis, while the remaining two tend to support the second hypothesis.

- The force of the wind shook the alarm bell.
- It does not seem that we necessarily are always simply taken, let alone overtaken, by our desires, so that they act themselves out in us as the force of the wind acts on a leaf.
- Miss Danziger put more pressure on the pedals; she knew that if she did not reach Sandweg before the force of the wind had accumulated its full strength, she would have very hard work before her.
- The islands present a bare and inhospitable face to the mainland, as their landward slopes receive the full force of the wind.
- This place must have stood here like this for all those years, and I imagined it retaining always this season, a pocket of perpetual spring — almost a source of spring from which the frozen bare earth in other places could be revivified, as those old maps depict in each corner a Wind holding in his bursting cheeks the force of the wind everywhere.

Two further quotations show that force can be considered a property of the wind by the way they refer to a 'gale' and '<u>its</u> force,' and to 'the wind driving hail' '<u>with</u> a force, respectively:

- ...like an Aldeburgh gale, changing the coastline, breaking down bridges, and generally modifying the landscape by its sheer force and energy.
- After what seemed like only a doze she woke to daylight, the wind driving hail against the window with a force which threatened to break it.

There were two more which talked of the 'force of air currents' and the 'force of an air-conditioning outlet' in ways that indicate that the 'force' is considered a property of the currents and the outlet.

- Migrating birds displ:y great sensitivity to the direction and force of air currents, changing altitude frequently to find the best conditions, and yet for many years their sensors have eluded detection.
- It is important to check that no seat is subjected to the full force of an air-conditioning outlet, and to ensure that no guests are distracted by waiter activity at an adjacent service station or kitchen door.

Finally, there is one example where the wind is clearly thought of as exerting a force:

- When we're sailing downwind, the wind is simply pushing the boat forward and there's no sideways force, so we don't need the daggerboard.

Only the very last of these is clearly consonant with the idea of Newtonian force. However, as its source is listed as *Royal Yachting Association instruction video: sailing (Business)* it seems likely that this in fact an example of what I have classified as a Technical usage, rather than a lifeworld usage.

Out of this section consisting of 28 instances of natural forces, then, there are 14 related to the relevant sense of *physical force*. These indicate that the lifeworld sense of *physical force* refers to either things, or to properties of these things, or possibly both.

Exerted forces – 17 instances

A better set of examples might be those to do with forces which we can exert ourselves – our experience of these is available throughout our entire life. And, of course, the same forces can be exerted by other things than people so this forms a natural basis for extending the idea of *force* to include inanimate agents.

For example, consider the most common of these: the *push*. In the first three of these examples a person is doing the pushing (in the third case, metaphorically so), while in the next two a psychological instinct is (metaphorically) pushing, in the second to last an abstraction is (metaphorically) pushing, while in the last inanimate matter is doing the pushing:

- At first it was a low moaning noise, and his mother tried to comfort him, but he pushed her away with a force that almost knocked her down.

- He saw a figure moving towards him, and felt himself pushed, with great force, into a wall.
- Successful schemes always had someone who could be identified as the driving force, someone who pushed and who steered the project through to a satisfactory launch.
- Again her driving force was animating him, pushing back an insidious sensation of inertia.
- At around two years, there is another driving force pushing her towards independence: the desire to be like everyone else.
- This is the economic force which pushes up national income.
- Beneath this the layer of ice was almost perfectly smooth, though occasionally an unseen force had pushed its plates together and upwards forming huge mountains amongst the hills and valleys.

In spite of all the differences between these sentences, they can be seen to fall into just two categories. In the first two sentences, with human agents, the word 'push' is supplemented by a prepositional phrase 'with ...force'. In the remainder, it is a force that is said to be doing the pushing. That is to say, pushing is something which is done by an agent, hence *force* is being conceptualised as an agent which does the pushing.

On the other hand, where an agent is otherwise specified, as in the first two sentences, there are two alternative interpretations of the meaning to be attached to the prepositional phrase containing 'force'.

The first of these is that *force* is being conceptualised as an instrument of some agent: in other words, just as one can push something 'with a stick,' so one can push something 'with a force.' On this interpretation, then *forces* are seen as instruments, by the use of which one can achieve one's aims. This seems unconvincing.

The second alternative is that a prepositional phrase like 'with...force' indicates the degree, strength or intensity of the push. This would also be consistent with usages like these, not related to physical force:

- I have never experienced the spring like this before, with such force, such awareness, such joy.
- The girls both stare at him, obviously impressed by the scale and force of his concern.

In these examples, like those where it is used to indicate wind speeds, *force* is, loosely speaking, a measure of how big something is.

Given these alternative analyses for pushing, one might like to look at the usage of the idea of pulling to compare the way it is conceptualised. Unfortunately, there is only one instance in this sample where the concept *pull* is linked with the concept *force*:

- If we all pull together, then we can be a successful force in world rugby again.

While this is not inconsistent with analysing force as an agent (since it can be interpreted as saying that "a successful force…" is one that pulls together), it is hardly convincing. To "pull together" is a dead metaphor meaning to cooperate, and the meaning of 'force' here probably belongs better in the Influence category noted above.

While pushing and pulling can be done by anything, animate or inanimate, the next two *exerted forces* can only be performed by people (or at any rate creatures similar to people – monkeys, say): *gripping* or *squeezing*, and *embracing*. In these examples, we approach the sorts of situations which would be classified as forces in physics, but it is noticeable that in each case the action is not described as a force. The force in these examples is not the action, but an attribute of the action:

- Nicholson gripped Merrick's hand hard and squeezed with unnecessary force, watching the flicker of pain cross the man's face.
- Even her body protector wasn't enough to stop her wincing at the force of his embrace.

With these, again, it seems that the force is seen as a measure of the intensity of the action rather than as an entity in its own right.

Sudden movement – 22 instances

The last set of instances has been grouped together because they all relate to sudden, short term change in motion. In so far as Newtonian force is related to acceleration, the rate of change of motion, these then are good candidates for lifeworld usages which fit with the physicist's usage.

Explosions seem excellent examples of sudden short term changes, and there are five instances in the BNC sample referring to the force of an explosion. However, a question similar to one discussed above arises: is the word 'force' here identifying a type of force (as it does in 'force of gravity') or an attribute of an action (as it does in 'force of his embrace')?

- Its basis is the naturally occurring substance uranium, an element which under certain circumstances can be made to become so unstable as to produce an explosive force.
- The question is the charge of gas, exploded in the cylinder head, which is the motive force of every piston-stroke.
- They'd seemed robbed of speech by the force of the blast.
- But in that very same instant, momentarily demolishing her composure, the door burst open with the force of an explosion and a tall dark-haired figure in a charcoal-grey suit came striding purposefully into the hall.
- About forty-five minutes later three huge explosions went off in the Wilkerson house, the force of which blew Dustin's $700 desk through a hole in the wall.

In the first two sample sentences above the word 'force' seems to be identifying a type of force: in the first it identifies an 'explosive force' and in the second it identifies the force with the explosion of the gas – '…gas, exploded… which is the motive force…' In the last three of these examples, by contrast, it seems to be identifying an attribute of the action. In each of these three cases the explosion is identified as a particular event. This can be seen from the underlined words: 'force of the blast,' 'force of an explosion' and 'three huge explosions…the force of which...' The 'force' therefore must be a property or attribute of this particular event: if it one wants to name a force it would have to be called the 'force of explosion' – the same way we speak of the 'force of gravity' not the 'force of a/the gravity.' Nevertheless, the attribute which 'force' is being used to express here is not just the degree or intensity of the explosion (cf. the discussion above of the force of a push). Rather, just as we can talk of the 'noise of the explosion' to indicate the sound created by the explosion, so here 'the force of the blast' refers to the force exerted by the blast: a force which has effects such as robbing people of speech, or blowing a desk through a wall. Thus all five of these sample sentences are examples of the lifeworld concept of *physical force* that is related to the physicist's.

The remaining 16 instances can be divided into five groups. The first three groups are directly relevant to the lifeworld concept of *physical force*: these are forces that are identified with impacts, force considered as the property of things, and forces identified with effects of impacts. The remaining two groups are less directly related: metaphorical uses where 'the force' of some abstraction 'hits' one, and examples where the word 'force' is used to indicate size, degree or intensity. These will be discussed in turn.

There are two instances where forces identified with their causes: in the first, 'force...is the... impact,' and the second speaks of 'the force of flail threshing,' a phrase similar in form to 'the force of gravity.'

- The only force generally believed to be sufficiently powerful is the high-energy impact of a large asteroidal fragment on the Moon.
- It would also catch grain bounding off the floor with the force of flail threshing.

There is also one instance, a rather gruesome description of an early version of the guillotine, which uses 'force' in such a way as to show that it is conceptualised as an attribute of a thing:

- As the block thundered down, its force propelled the severed head into the basket she carried in front of her on the saddle.

As well there are four instances where 'force' is evidently conceptualised as being caused by an impact: the expressions 'force of the blow,' 'force of it' and 'force of the impact' indicate this interpretation, for the same reasons mentioned above when discussing the force of explosions.

- As she stopped a few feet away from them she saw that his cap had been knocked some distance from where he lay and the force of the blow, which had thrown him through the air, had dislodged his fountain pen from his pocket.
- He must have taken the force of it on his shoulder, it doesn't swing very smoothly.
- The force of the impact knocked the breath out of her.
- A rake left lying on the ground is dangerous, not only because of its teeth, but because an unwary step on the head of the rake could suddenly swing the handle upright with shattering force.

While metaphorical uses of 'hitting' with 'force' are not directly related to the idea of physical force, they presumably base their meaning upon the physical sense of 'hit'. There were four instances of this sort of usage in the thousand downloaded from the BNC:

- The full force of being public property and pop royalty had not hit him — whereas Beatty, who was being chased by the news hounds from New York to Florida to the South of France, with Natalie in tow, had suddenly become the talk of the town.

- Neither of these events, nor other similar ones, have hit the imagination of the media or the public with the same force as the similar serious scandals involving children.
- Then, suddenly, just as Joanna was nodding a dubious assent, the full force of what she had been told hit her like a thunderbolt.
- He said then that the housing market recession had not hit the North East and North Yorkshire with anything like the force with which it had struck the rest of the country.

It is surprising that only metaphorical usages of 'hitting' with 'force' were found in this sample: if the metaphor is based upon the physical sense, then one would expect to find some instances of the latter. While none were present in this sample, they do exist, as for instance, the following, which was found in another search of the BNC:

- Although Jack was taller, Ho hit him with such force that soon Jack fell away from him and he sprang out.

The last grouping to be discussed here is where 'force' is used to indicate the degree or size of effect of some impact. The impact can be of many different types: some described with verbs such as 'smiting', 'slapping,' 'thumping,' 'banging', or 'slamming;' some described with a noun such as a 'blow;' some described with a phrase like 'bringing his fist down.' In each of these cases, however, the 'force' describes how much effect (usually a lot) the action had. This group includes the following:

- An open palm, as big and fattily solid as a Bradenham ham, smote the side of my head with horrific force.
- The delicate porcelain features split two inches above the right eye as a fountain of blood arced out, a trajectory of sheer surprise at the force of the blow.
- The force with which he brought his clenched fist down on the parapet showed just what he had held in check.
- The sound of the slap surprises her, she drops her hand, tingling with the force of it.
- Then whoever it was began thumping on the knocker with renewed force, and Jessamy forgot all about being cautious.
- She banged the receiver down with a force that made Claudia wince.
- I grabbed the door handle and pulled the door wide open and with all my force slammed the door into him before he could do anything about it.

Overview of the thousand sample sentences

After removing the technical usages, and lifeworld usages restricted to the specialised context of wind-speed and other usages where 'force' is used to indicate the degree, intensity or strength of some action, and, finally, discounting metaphorical usages there remained only 32 instances which showed 'force' being used with a sense related to the physicist's notion of *physical force*, or slightly more than 3% of the total.

Of these, a number of examples conceptualised force as an entity in the world: that *water*, *a storm*, *light*, and *wind* are thought of as forces is arguable from the evidence of the sample sentences discussed above in the section headed Natural Forces. This should not be surprising: many other usages of 'force' – *force* as influence, for example – conceptualise a *force* as an agent.

It was found that lifeworld examples of physical force were predominantly related to pushing, gravity, explosions, and impacts of various sorts. This is consistent with Study Two, where these were found to be highly rated as examples of force.

Quantified forces

It has been mentioned above (see not only the preceding section, but also p. 124 and pp. 144–146) that that there is evidence consistent with prototypical forces being large. Further evidence for this can be seen if one examines all the examples, out of the thousand in this sample, where the force has been described in quantitative terms, irrespective of whether in the sense of physical force, or force in one of its other senses.

When this is done it is found that out of 86 such instances the force is described as in some way large in 80 instances. Most commonly used (19 instances) was the expression 'full force,' closely followed by 'powerful force' (18 instances). Other common adjectives were 'great' (8 instances), 'sheer' (6), 'strong' (4), and 'huge,' 'irresistible,' and 'formidable' (each occurring 3 times). Forces could also be 'enormous,' 'potent,' 'too much,' 'exceptional,' 'tremendous,' 'invincible,' 'terrible,' 'horrific,' 'unstoppable,' and 'excessive'.

By contrast there were only five instances when the force could be considered of a medium size: the phrase 'reasonable force' occurred four times and 'sufficient force' occurred only once. Even here, the sense is surely *large enough* rather than *medium*.

There was only one instance where the force could be possibly counted as small: this occurred in a sentence quoted earlier:

- Outside, the sun still shone, though with the approach of evening the light had lost some of its force and the room grew dimmer.

An NSM description of physical force.

As discussed above, the idea of a physical force that is being sought here is not identified with actual entities. However, the evidence shows that forces are often conceptualised as being agent-like, as being like pushing or impacts, and as being large.

An NSM definition that covers these aspects is as follows:

F is a force

F is something that happens	1
Sometimes someone does something (X) to something (W) near and because of this something (W) moves far from this person.	2
A force is something like this (X).	3
Sometimes someone (S) wants something (X) to happen to something (Y)	4
S can do something (Z) to Y	5
because of this, something (X) can happen to Y	6
X does not happen to Y if S does not do this to Y	7
A force F is something like this (Z).	8
When someone thinks of a force (X) they think X is big	9

Line 1 states that a force is not an entity in the world, but rather an event.

Lines 2 and 3 describe pushing and say that a force is like this. That is to say 'pushing' can serve as a prototype for force.

Lines 4 to 8 indicate that forces are similar to the actions of agents – things that cause other things to happen.

Line 9 summarises the default assumption that forces are large.

Rephrasing the NSM definition in a more compact form, it says that forces are events like strong pushes, which make things do what we want them to. (While this is more compact, however, it has the fault of circularity – to "make something do what we want" is to "use force;" "to push" is "to exert a force away from oneself"; and someone who is "strong" is someone who can "exert great force.")

Conclusion

The four studies reported in this chapter converge in a definition of *Newtonian force* which shows that it is based on the lifeworld concept of *force.*

Study 1 focused on the denotation of the technical concept of *force*: that is to say, the situations which would be classified as showing forces. This study showed that for students the highest rating examples of *Newtonian force* were *pushing* and *hitting*, and that these prototypical forces were thought of as having large effects, involving two entities, and requiring considerable effort.

Study 2 showed that 'pushing' and 'hitting' were amongst the earliest words which children learnt. As prototypical examples are amongst the earliest to be learnt this was consistent with the claim that these were prototypical forces.

Study 3 showed that the sense of *force* was entangled with those of *power*, *strength* and *effort*, amongst other concepts.

Study 4 showed that the technical sense of *force* could be derived by modifying the lifeworld sense of "imposing one's will on another." A survey of a very large sample of usages of the word 'force' showed that situations including Newtonian forces formed only a small subset of things that people count as *forces.* This was reflected, too, in the fact that the length of the NSM definition of Newtonian force was longer than that of the NSM definition of the verb 'to force:' additional clauses in the definition were required to whittle down the number of situations which the definition covered.

Unlike the case of *concept*, which was found to be used in a way that was equivalent to its lifeworld usage, and unlike the case of *acceleration*, which was found to be an extension of the corresponding lifeworld concept, the concept *force* has been found to be a restriction of the corresponding lifeworld concept.

Students learning to understand the Newtonian concept of *force* need to learn not only what are the prototypical situations, but also what other situations are included (eg. those involving passive reaction forces (Minstrell, 1982); frictional forces (Stead &

Osborne, 1981)). As well they need to learn the classes of situations to exclude (eg. those involving momentum, energy, or powerful agents or influences).

Because students have to learn which cases of lifeworld *force* to exclude as well as which ones to include in the category of *Newtonian force*, a more complex sequence of conceptual development can be expected with this concept than is found with *acceleration*. Students need to learn to include in the category of *forces* situations in which the object applying the force is not salient: frictional forces, the normal reaction force of a supporting surface, and the weight force for example. They also need to learn to exclude situations where the object applying a force is salient, and the effect of the force is obvious, but where the force has ceased acting: for example, they must learn that the 'hit' on a golf-ball does not act throughout the flight. As well they need to exclude influences on motion like momentum, energy and power. Because each of these adjustments is independent of the others there are a wide number of possible sequences in the development of the Newtonian concept of *force*.

Even though they may not have achieved a final understanding of the technical concept of *force*, or of *acceleration*, the fact that they have grasped the prototypes enables students to have a rough idea of what is meant by their teacher, their textbook and their fellow-students (who will also, by and large, have a similar rough level of understanding). When asked to apply their knowledge to prototypical physics problems (eg. calculations using formulae applied to situations involving constant forces on frictionless surfaces) this level of understanding will be enough to get by. When occasionally they encounter a less prototypical situation they may learn to modify their concepts (for example, learn to recognise passive reaction forces), and if these situations are encountered frequently these modifications may become a permanent part of their understanding.

The next three chapters will show how dealing with non-prototypical problems influences students' concepts of *force* and *acceleration*. It will also illustrate the influence of the different concepts held by students on the way that these problems are dealt with.

The problems discussed by students in the next three chapters come from the *FCI* and were developed for the purpose of uncovering common misconceptions in the understanding of Newtonian mechanics. The next chapter, dealing with Newton's third law, will include discussions of the way students dealt with the relevant problems on the FCI. Chapters Seven and Eight will then discuss the way students dealt with problems involving Newton's first law and second law respectively.

Newton's third law

So far, this thesis has been focussed upon the structure and development of particular concepts: *acceleration* and *force*. The next stage is to look at how larger scale conceptions, ones that use these individual concepts, are structured and develop. An important set of conceptions in mechanics is Newton's laws. These refer to the concepts already discussed, and develop partly out of everyday experience and thought, as well as explicit instruction. Of course, everyday experience can often appear to be incongruent with Newton's laws and this is one of the main sources of difficulty that students face in aligning their study of Physics with their overall world-view.

In the previous two chapters it was argued that students hold prototypes for *acceleration* and for *force*. The importance of prototypes is again emphasised in this chapter which extends the idea of prototypes from simple concepts to situations. In particular, it is argued in this chapter that students hold prototypes for

- situations involving Newton's third law, and
- action/reaction pairs.

It is further argued that

- where these prototypes are incorrect, that predictable errors will occur, and
- the level of errors will be related to the 'distance' from the prototype.

This is argued by three studies: the first re-examines the data provided by earlier researchers to show that it can best be explained in terms of prototypes. The second study analyses transcripts of student discussions of the solutions to problems in the Force Concept Inventory (Hestenes et al., 1992) to show that these too are best understood in the framework provided by prototype theory. The third study sought to find whether teaching which explicitly took account of prototype theory could be effective.

Study One: reanalysis of earlier work

As noted in Chapter Two, questions intended to elucidate difficulties with Newton's third law were included in a survey by Watts and Zylbersztajn (1981), which was intended to find both what was the distribution of children's ideas about forces, and also the extent to which their teachers were aware of these ideas (children: N = 125; teachers: N = 5).

A diagram showing stick figures in a tug of war (Figure 26) was presented to the subjects, and they were asked to compare the sizes of the forces exerted on the rope by the left and right contestants when one or the other was winning. They found that most students (82%) claimed that the winner was exerting the greater force: 'If one person is winning a tug-of-war, then it is clearly very difficult for children to imagine that the forces on the rope joining the two people are equal' (p. 363). While this was taken as evidence for widespread difficulty with Newton's third law, there is a problem with the question, which in one way raises a doubt about their conclusion, and in another serves to confirm it. What raises doubt about their conclusion with respect to the answers to this question is that, in so far as the contestants were each pulling on the rope and not on each other directly, the forces they exerted were not an action/reaction pair. That is to say, both forces were acting on the rope, while an action-reaction pair, in Newton's sense of these terms, acts on two different bodies. While the authors may have made the tacit assumption that the rope was an idealised light inextensible string serving simply to transmit the force, and that therefore Newton's third law was applicable, they may have simply not understood the third law themselves.

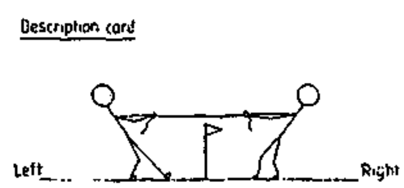


Figure 26: Diagram used by Watts and Zylbersztajn (1981)

That this ambiguity (forces on rope vs forces on each other) is present in the question can be seen by a brief comment made by one of the teachers, who clearly saw the question as being about Newton's second law: '...I mean its not clear when you say he is winning... do you mean he is accelerating that way...' Watts and Zylbersztajn quoted this teacher to suggest that confusion about Newton's third law was not confined to students, but was present in teachers. However, the confusion was implicit in the question itself and, quite possibly, in the researchers.

On the other hand, as suggested above, this in a sense confirms their conclusion: if the researchers themselves are not clear about Newton's third law then confusion is indeed widespread.

This particular problem of interpretation does not occur in the research presented by Maloney (1984). Here students were presented with diagrams showing two blocks touching (see Figure 27 for some examples) so that the forces of each block on the other were unambiguously an action/reaction pair.

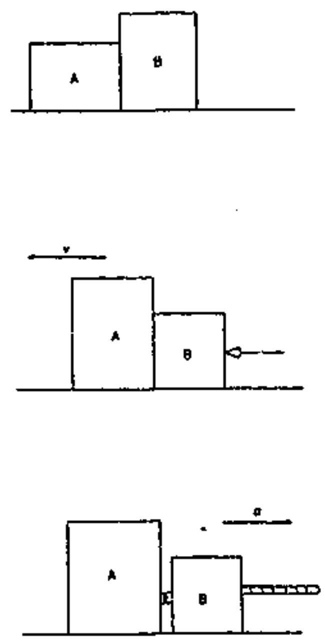


Figure 27: Diagrams from Maloney (1984)

The situations were varied along four dimensions: mass of blocks (equal or not), motion (rest, constant velocity, constant acceleration), direction of force (push or pull), and agency (which block was the 'cause'). Of these dimensions, the direction of force was not found to be of significance. Overall, then, eight situations were of interest. These are shown in Figure 28 . For each of these situations, the students were asked to compare the size of the forces of block A on block B, and of block B on block A. Only 9 of 112 students consistently answered according to Newton's third law. The remaining students all answered differently according to the situation. The main point Maloney was concerned to make was that, for most students (63%), the students' pattern of answers corresponded with the answers that would be expected if the students were using one of five simple rules (Table 16). This result provides evidence for the active involvement of the students in the conceptualisation of physics (none of these rules, obviously, had been taught), a point which will be returned to below.

Table 16: Common Rules Maloney (1984, P. 40)

Rule designation	Description of rule	Percentage of subjects overall using rule
1a	Mass is the only determiner for all states of motion. Greater mass exerts greater force.	11
2a	At rest the forces are equal, but for moving systems greater mass exerts greater force.	16
2b	At rest the forces are equal, but for moving systems the 'cause' exerts greater force.	19
3a	For rest and constant velocity the forces are equal, but for accelerating systems greater mass exerts greater force.	8
3b	For rest and constant velocity the forces are equal, but for accelerating systems 'cause' exerts greater force.	9
Total		63%

However, for the purposes of the present argument it is worthwhile looking not at the details of the hypothetical rules that the students were using but, instead, at which of the eight situations the students claimed had forces of equal magnitude. Doing this reveals the typical pattern for prototypes: all students agreed about the central member of the category – objects of equal mass exert forces of equal magnitude on each other – but different decisions were made about the boundaries (see Figure 28). It will be recalled from Chapter Two that when people were asked to pick out from an array of colour chips examples of a particular colour, for example, red, everybody picked out central instances but people made different decisions about whether or not to include colours near the boundaries where, say, red left off, and orange took over. Similarly, in

Labov's (1973) investigations of the meaning of 'cup', when subjects were presented with arrays of drawings where the height, width and other features of a container were systematically varied people agreed on the prototypical cup, but disagreed on the instances near the boundaries with bowl or mug and so on.

The same sort of pattern is clearly present in the results presented by Maloney (1984). Figure 28 shows the eight categories of situation used by Maloney, with boundaries drawn around them in order to indicate which situations were included in the category 'equal forces on each body' for each of the rules listed by Maloney. Note that the percentages do not total 100% as not all students' answers were captured by the listed rules: although '41 distinct rules were found' (p. 39), details were provided only for those listed here. Nevertheless, it is quite clear from the reported results that the prototypical exemplar of a Newton's third law situation for these students is one where the masses of the two bodies are equal: all students agreed that in this case the forces were equal.

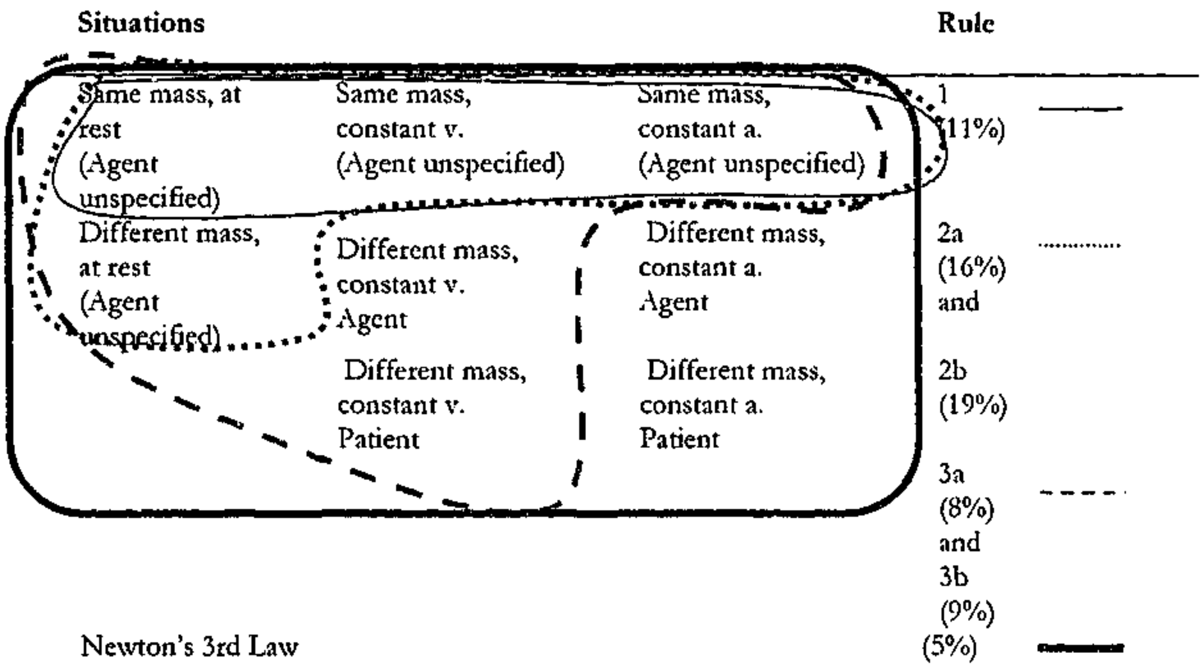


Figure 28: Situations for which students stated that the forces of the blocks on each other were equal (based on data from Maloney, 1984).

Furthermore, there is evidence consistent with the hypothesis that the extension of the category 'equal forces' is a result of teaching. Rules 1 and 2, which specify the most limited range of cases where the forces are equal are those favoured by 55% of those who have not studied high school physics ('novices'), but are used by only 16% of students of those who have studied high school physics ('experienced'). Just 7% of novices, but 50% of experienced students favour the more inclusive Rule 3. And Newton's third law, the most inclusive category, was used by no novices, but by 9% of experienced students (data from Table 3 in Maloney, 1984, p. 41).

Finally, a consequence of Maloney's argument is that students do not just add situations at random to the category 'equal forces'. Rather, they followed rules that they themselves must have generated (they were certainly not taught), to guide their extension of the category. This, in my view, is related to Lakoff's (1987) discussion of 'radial categories' where the extension of a category is motivated by likeness in some key feature. In this case, the grouping together of the state of rest, and the state of constant velocity by the 'experienced' students, plausibly provides the motivation for their move from Rule 2 to Rule 3.

Overall, then, Maloney's paper provides evidence of three characteristic prototype phenomena:

- Particular exemplars are agreed to by all.
- These are amongst the first to be learnt.
- Extensions to the category are motivated.

Further evidence of prototype structures can be seen in the research reported by Terry and Jones (1986). They used a set of seven questions to examine the understanding of Newton's third law by 16-year-old students (N = 39). From the first two questions they found that most students (95%) could not identify the reaction force to the action force of gravity acting on a person standing still. Furthermore, most students (90%) could not identify the reaction to the force of gravity acting on a falling stone. This was in spite of the heavy cueing provided by the form of the latter question where the weight force acting on the ball was referred to as 'the force *of* the *earth on* the *stone*' (p. 295, emphasis in original). The authors comment that the 'emphasis was clearly insufficient to prompt the pupils into giving the corresponding force *of* the *stone on* the *earth*'.

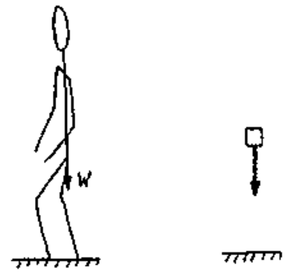


Figure 29: Diagrams for the first two questions used by Terry and Jones (1986).

Five more questions were asked. Two of these questions asked students to predict the effect of forces on the motion of a boat, and the remaining three asked them to compare the relative sizes of pairs of forces between people, and to compare the relative sizes of the action/reaction pair between a car windscreen and an insect. Only in one circumstance (two skaters of equal mass pulling on each other via a rope) did most of the students (90%) state that the forces were equal in magnitude. The authors suggest that this 'can probably be attributed to the use of this example to introduce Newton's third law' (p. 295).

These results can be interpreted simply in terms of prototype theory: for these students the prototypical third law situation was the one that was used to introduce it (see Table 17). As such, it was easily recognised, and Newton's third law, as they understood it, was applied. For situations that were not prototypical, the usual uncertainty that is likely at the boundaries of concepts occurred, and some students did, and others did not recognise them as third law situations.

Action/reaction pairs

It is suggested that, for these students, a prototypical Newton's third law situation is one where

- the focus of attention is on a pair of forces
- exerted by two people
- of equal mass
- on skates,

- both pulling on a rope (Figure 30).

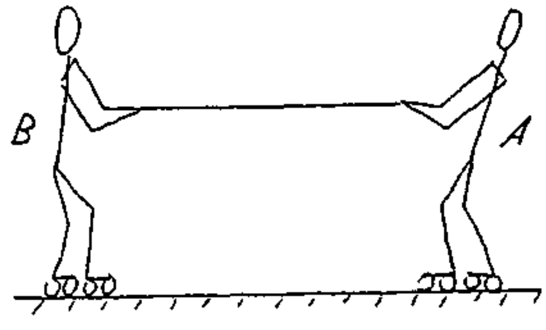


Figure 30: Diagram for question 5 in Terry and Jones (1986).

What were the prototypical features of action/reaction pairs in the minds of the students? From the evidence one can hypothesise that these are that the forces

- occur on the same object,
- have the same magnitude,
- have opposite directions, and
- fit the formula 'force of A on B = force of B on A.'

Of course, the first feature, should *not* be present as action/reaction pairs act on different objects, but these four features were *all* present in what, for these students, was the situation used to introduce Newton's third law and therefore likely to be, for them, the prototypical situation. As well, the second and third features are emphasised in the traditional formulation 'for every action there is an equal and opposite reaction,' and the fourth feature was emphasised in the wording of the questions.

In so far as action/reaction pairs never act on the same object, it was surely unfortunate that the situation chosen to introduce Newton's third law was one where a rope was being treated as an idealised method of transmitting force. For the teacher, presumably, the force is to be implicitly understood as being 'transmitted' by a (light inextensible) string from A to B and from B to A, but any misunderstanding by the students is understandable. For the students, both A and B were pulling on the rope, and the forces due to A and B were identified as action and reaction. For the students therefore, the reaction force to the pull of A on the rope would very likely be understood as the pull of B on the rope.

Assuming that the prototypical features for action/reaction pairs in the minds of these students were as hypothesised above, one can interpret their answers. In the first two questions, they were asked to identify action/reaction pairs. In each case the most common answers were pairs of forces present in the situation represented in the diagram. In the first situation, a person standing on the ground, the normal force due to the ground and the weight force had three of the four prototypical features (i.e. all except fitting the formula) and were identified as the action/reaction pair by 'over two thirds' of the 37 students who had answered incorrectly. In the second situation of the falling rock, the weight force and air resistance had two of the four prototypical features (i.e. acting on the same object, and in opposite directions) and were identified by as an action/reaction pair by 18 of the students. In so far as air resistance had so few of the prototypical features the other 21 students either did not consider it or rejected it. Of these, 17 were unable to identify the reaction force (which after all was not acting anywhere within the situation represented in the confines of the diagram), while 4 students (prompted perhaps by the phrasing of the question) were able to provide the correct answer.

Table 17: Prototype Features vs. Choices for Action/Reaction Pairs

Choice	Description	Both in situation represented by diagram?	No of prototypical features	Chosen by
1	Weight force/normal force	Yes	3 of 4	67% (approx.)
	Force on man by earth/force on earth by man	No	3 of 4	5%
2	Weight force/air resistance	Yes	2 of 4	46%
	Force on man by earth/force on earth by man	No	3 of 4	10%

In the next two questions, students were not asked to compare the size of forces and 'there was no evidence to indicate that the pupils interpreted the interactions in terms of a pair of forces' (p. 295). The remaining three questions did ask for forces to be compared.

Table 18: Prototype Features of Situations vs. Correct Use of Newton's Third Law. Data from Terry and Jones (1986)

Number of prototypical features	Question	Newton's third law used correctly by
0 of 5	3	10%
1 of 5 (2 people)	4	15%
1 of 5 (focus on pair of forces)	7	40%
4 of 5 (focus on pair of forces, 2 people, on skates, pulling on rope)	6	50%
5 of 5 (focus on pair of forces, 2 people, equal mass, on skates, pulling on rope)	5	90%

From this set of data three conclusions can be argued for. Firstly, that as the number of prototypical features increases, the situation is more likely to be recognised as a member of the category. Secondly, that not all features of the prototype carry equal weight: equality of mass (compare the responses for questions 6 and 5), and focus on a pair of forces (compare the responses for questions 4 and 7), are clearly of more importance than other features (compare the responses for questions 7 and 6). Thirdly, where a force is not present in the immediate situation it is unlikely to be thought of, even though it might have many prototypical features (compare the popularity of the two different responses for question 1, or for question 2).

Brown (1989) reported on three studies investigating aspects of students' understanding of Newton's third law. Brown interpreted the results of the first two studies, interviews with pre-physics high school students, as indicating that before instruction students viewed forces as properties of objects. This view clearly causes students difficulty in their understanding of Newton's third law. This is because such a view would entail that students would have to interpret Newton's third law as stating that what one might call the 'force-property' of two objects was equal in size and opposite in direction. Such an interpretation would make Newton's third law essentially a mystery: 'After all,' students might think, 'why should the force-property in an interaction be the same when, say, the mass- or colour-properties were not?' Nevertheless, this view throws light on the prototypical status of equal mass objects in the understanding of Newton's third law. When other properties are the same, then it may well seem logical that the forces are the same too.

Brown's third study compared the pre-instruction and post-instruction scores of high school students (N = 78) on a subset of six of the questions from a multiple-choice test on mechanics. In this subset there were three questions with diagrams involving boxes or blocks (*b*, *d* and *f* in Figure 31), and three involving drawings of lifeworld situations (*a*, *c* and *e* in Figure 31). None of these six situations were prototypical Newton's third law situations: none explicitly indicated that the masses of the two objects were equal (although some students might have assumed this in situations *a* and *c*, where the diagrams do not show any noticeable difference in size). For all of these situations, some students did and some did not state that the forces were equal, reflecting their status as non-central examples of the category 'situations in which Newton's third law should be used'.

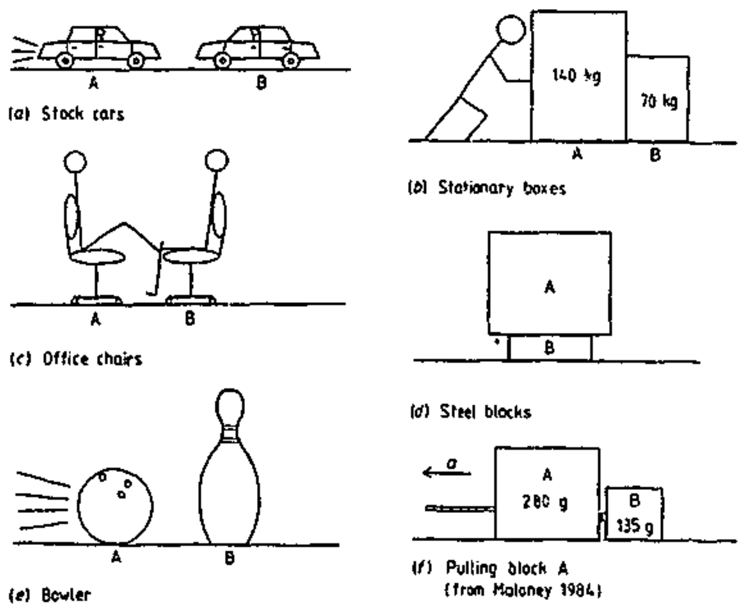


Figure 31: Diagrams from Brown (1989)

Overall, then one finds that where the work on Newton's third law is relevant it is clearly interpretable in terms of prototype theory.

Different features each separately contribute to the activation of the idea of Newton's third law for any given situation. Those situations that are most highly activated are the prototypes. These prototypical situations are typically the first learned, and they are also the core of the concept, shared by all who use it. Extensions to the category of "Newton's third law situations" are not adopted by fiat, or at random, but rather are motivated by regularities perceived by the students, and by regularities to which they are introduced by dialogue with others.

The next section looks at discussions between groups of students negotiating answers to questions which are designed to elicit misconceptions related to Newton's third law.

Study Two: The use of Newton's third law in student discourse

Evidence presented in the first half of this chapter showed that a prototypical Newton's third law situation is one where two identical objects at rest push against each other. Around this central idea are cases where the objects have different masses, are moving at constant speed or accelerating, and where only one is actively pushing, and various combinations of these.

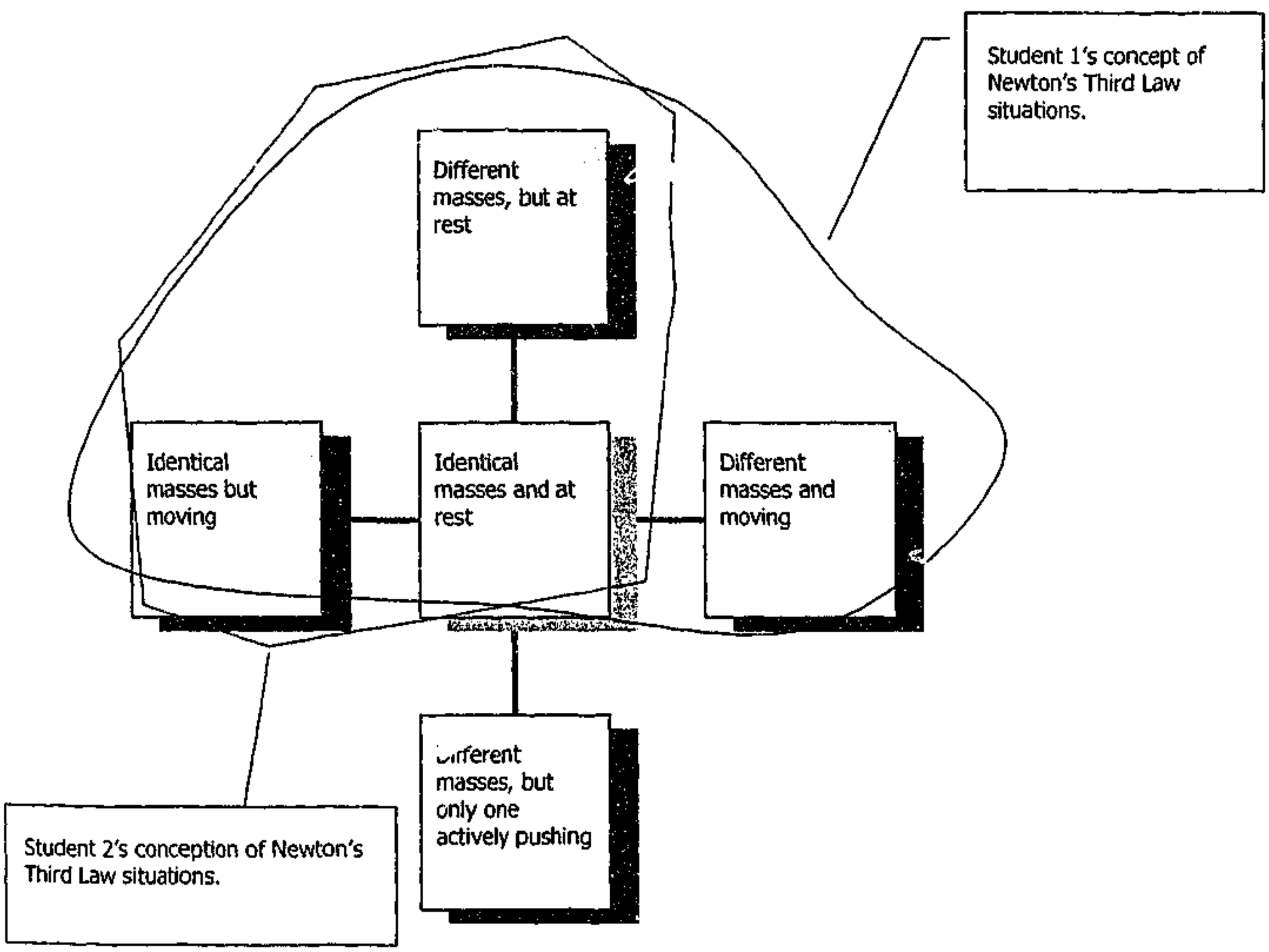


Figure 32: Two alternative conceptions of when Newton's 3rd Law is appropriately used.

In this second half of the chapter, students' discussions of problems involving non-prototypical situations are discussed. It is argued that students develop their ideas about the situations in which Newton's third law should be applied by focusing on the boundaries of their conceptions and arguing over reasons for extending the boundaries to include these situations, or contracting the boundaries to exclude them. Thus in Figure 32 the two students would agree in using Newton's third law in most circumstances, but disagree in cases where there were moving objects with different masses. They would argue about whether student 2 should extend the boundaries of his conception to include this case, or whether student 1 should contract the boundaries of his conception to exclude it.

Students discussed a number of problems from the *Force Concept Inventory* of Hestenes, Wells, & Swackhammer (1992) which could be solved by using Newton's third law. These were questions 2, 11, 13 and 14. Questions 2, 13 and 14 asked about forces between cars and trucks in different situations. Question 11 asked about forces between people. Each of the questions was about a non-prototypical situation.

The amount of time spent in discussions obviously varied from group to group and from question to question. However, a very noticeable pattern was that where the group members agreed on one of the options for the answer, the question was quickly dismissed. Students spent more time in discussion when there were a number of different options proposed as the answer, but the number of options did not appear to change the amount of discussion, so much as the presence or absence of disagreement (see Figure 33). Putting this in terms of the discussion above, where conceptions overlapped (as for three of the four situations for students 1 and 2 in Figure 32) there was little or no discussion, but where they didn't (as at the boundaries in Figure 32) the students needed to negotiate to determine whether it fell inside or outside.

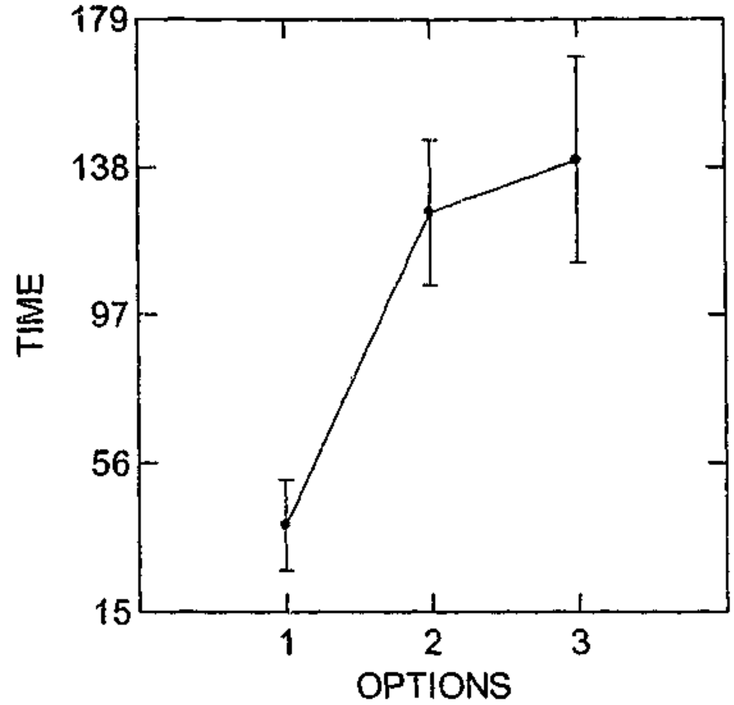


Figure 33: The average amount of time (in seconds) spent in discussion for groups who agreed (1 option) or disagreed (2 or 3 options) amongst themselves.

Discussions of Question 2

In answering this question, the details of which are given in Box 1, students were faced with a non-prototypical situation for Newton's third law. That is to say, the two objects were not of equal mass. The relevance of Newton's third law was noted by all groups (see Table 19) who had any discussion of the problem. However, an issue which arose repeatedly was the size and mass of the truck compared to the car. As shown in Figure 32 this was a sticking point for many students – they had to be argued out of a predilection to believe that the forces were different when the sizes or masses of the objects were different. It is also notable that not all of the students suggested that the larger object exerted the larger force – a speaker in Group 9752 suggests that the force by the car will be greater because the car is smaller (possibly due to thinking in terms of how concentrated the interaction will be if it occurs in a smaller region; or, in effect, confusing force with stress). In countering these arguments students argued in three ways:

Pointing out that one *can* apply Newton's third law (either explicitly mentioning it or referring to action/reaction pairs). (e.g. in groups 9751, 9752 and 9765)

Differentiating one aspect of the situation from the *size* of the force (e.g. in group 9765 the mass of the car is agreed to be different, but nevertheless the forces were still the same)

Pointing out that a situation even further out from the boundaries that the forces were known to be equal:

- hit the baby, the baby hits you with the same force
- punch on the nose, nose on the fist

In introducing the topic it was explicitly emphasized that, in the case of a punch on the nose, the forces by the fist and nose were identical in magnitude, but the example of the baby was original to the student.

Box 1: Question 2 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

2. Imagine a head-on collision between a large truck and a small compact car. During the collision,

   (A) the truck exerts a greater amount of force on the car than the car exerts on the truck.

   (B) the car exerts a greater amount of force on the truck than the truck exerts on the car.

   (C) neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.

   (D) the truck exerts a force on the car but the car doesn't exert a force on the truck.

   (E) the truck exerts the same amount of force on the car as the car exerts on the truck.

In some cases students accepted these arguments, in others they seemed only to accept that they were outnumbered (e.g. in group 9765 the discussion ends with "OK, when in Rome…").

Table 19: Brief Notes of Student Discussions for Question 2.

Group	Time (s)	Discussion
9753	7.6	What did you get E E E E yah
9711-3	22	I got D ah E, sorry all agreed?
9711-1	27	{Newton's third law immediately stated.}
9751	64	I think that's impossible I can't explain Newton's third law
9711-2	65	Is it the same amount of force matter and mass what answer you put E same why I don't know Newton's third law
9752	73	All except Andy says E because Newton's third law why did you say B I thought the car is smaller than the truck it will exert more force but Newton's third law applies to all objects so E
9762	81	Newton's third law but the mass of the truck is greater so there is more force from the truck but action/reaction punch on the nose, nose on the fist
9765	83	Ah! E, why! Newton's third law the force on the car and the truck is the same E! but you must imagine the size of the truck more force Newton's third law is in theory the car is less mass but the force is the same OK, when in Rome…
9764	112	{Newton's third law mentioned} "but greater mass" {but eventually agree on E}
9761	125	Action/reaction what about the mass Newton's third law hit the baby, the baby hits you with the same force I don't think so! either A or E

Study Three: Does teaching with prototypes help?

The very small amount of research devoted to seeing whether teaching with prototypes is useful has been restricted to the teaching of invented concepts (Tennyson, Chao, & Youngers, 1981). Given that there are clear prototype effects in the use of Newton's third law it was thought worthwhile to try basing the teaching sequence around them when an opportunity to do so arose, at short notice, soon after the writing of this chapter. It should be noted that this small scale and informal study is described only in its role as a pilot for a controlled and better designed study at a later date.

A ten week introductory science course for Year 10 students was being taught by a number of teachers. During the one week of this class scheduled for Physics, the topic to be taught was Newton's laws. At this time I was scheduled to teach a year 11 introductory science class where a week was to be devoted to ecology. A biology teacher who was taking the Year 10 class suggested that we swap classes for this week, so that each of us would be teaching from our specialty. We did so, and this provided me with the opportunity to try teaching Newton's third law based on a sequence developed about prototypes. This teaching sequence comprised just the four 50 minute lessons that were available for teaching, and one for a class test on the topics taught.

The first lesson was devoted to teaching Newton's third law, using prototypical cases (i.e. equal masses, not moving) to introduce it, and then stating that it always applied in all situations, and explicitly teaching that it applied when the masses were not equal, when the masses were moving, and whichever of the masses was the active cause of the event. In addition, Brown and Clement's method of teaching by means of anchoring conceptions was used to teach about passive reaction forces (Brown, 1994; Brown & Clement, 1989).

The second lesson was used to teach Newton's first law using demonstrations with a linear air track, and the third and fourth lessons were devoted to teaching the second law quantitatively (using $F = ma$) and qualitatively (using arrows to represent the direction of a force causing things to get faster, slower or change direction.

In each of the four lessons some time was devoted to each of problem work, class discussions, and small group discussions.

In the class test at the end of the week (See Appendix C) there were a number of questions on each of these topics. Three of the questions on Newton's third law were taken from the Force Concept Inventory (Hestenes et al., 1992), and six from Brown (1989). For each of these, the authors of the articles had provided figures for the percentage of students who answered these questions correctly.

In the case of the Force Concept Inventory, results were recorded for Arizona High school students in regular classes (both pre-test and post-test), for the honours High school students taught by Wells, and for end-of-first-year Physics undergraduates at Harvard University. A box graph comparison of these group's mean scores on these test items is shown in Figure 34. The results of the Year 10 students taught by means of prototypes are also shown.

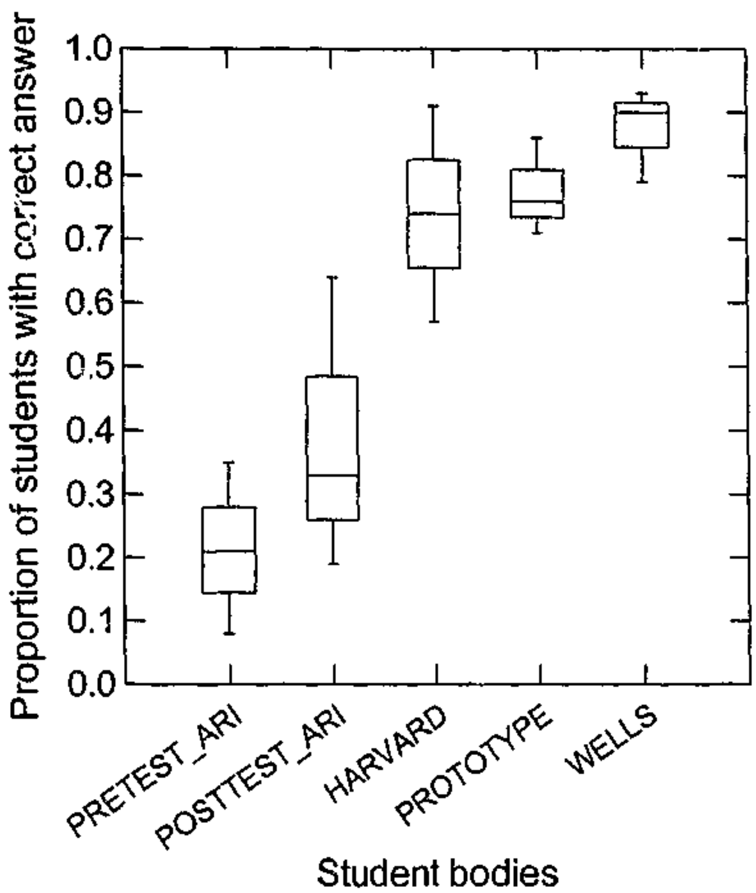


Figure 34: Proportions of students in classes who correctly answered questions 2, 13 and 14 from Hestenes, Wells and Swackhammer (1992).

In the case of the six questions from Brown (1989) figures were available for the percentage of students who correctly answered before and after instruction for a term in Mechanics. These can be compared, in Figure 35, to the results for the group of Year 10 students taught by means of prototypes.

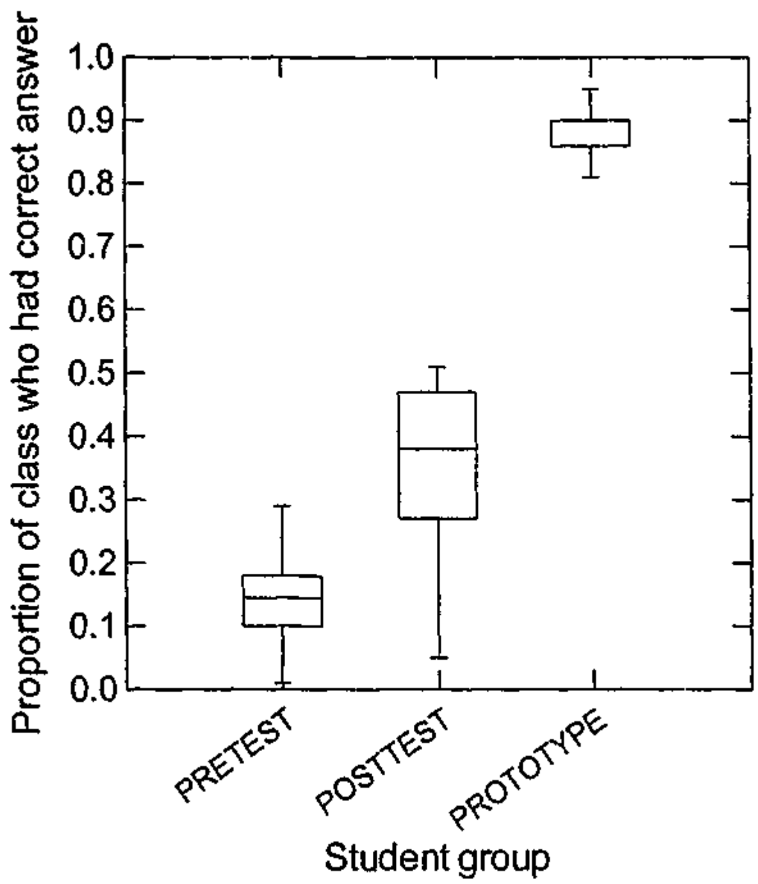


Figure 35: Proportions of students who correctly answered six questions about the sizes of action/reaction pairs. Pre-test and post-test data from Brown (1989).

It is worth noting that these Year 10 students were doing a compulsory unit in science, and were generally younger than the students in the other studies, who had elected to take courses in Physics. While the lack of a control group (perhaps these students were unusually adept, and would have performed just as well however taught) and the lack of a pretest and post-test comparison (perhaps they had already learnt this topic in earlier classes) mean that one cannot rely on these results as conclusive evidence for the efficacy of teaching Newton's third law by means of prototypical situations, they are nevertheless suggestive. They are certainly not inconsistent with there being benefits in doing so.

Conclusion

It has been argued in this chapter that just as there are prototypes for concepts such as acceleration and force, there are also prototypes of "situations which are covered by Newton's third law."

On the one hand, there is nothing unusual in this finding. Barsolou (1985) found that there were prototypes for "foods not to eat on a diet" and other categories describable by phrases rather than labeled by a single word. Furthermore, in Chapter Four it was shown that while concepts were prototypically describable by a single word, that nevertheless concepts described by short phrases were not uncommon.

On the other hand, thinking about the problem of "how to get students to generalise their knowledge so as to be able to apply a principle in new situations" as being equivalent to the problem of "how to get students to learn to recognise the category of situations covered by a principle" has three consequences. Firstly, it enables one to make sense of earlier research into student understanding of Newton's third law in the light of prototype theory. Secondly, it provides a useful framework for gaining insight into student discussions while problem solving. Thirdly, it suggests ways to apply prototype theory to course design. Whether these ways of using prototype theory in teaching are more efficient than standard methods remains an open question, but the results of an exploratory pilot study suggest that this is worth investigating.

Newton's first law

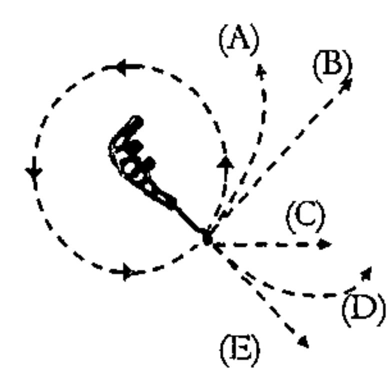
The previous chapter looked at Newton's third law. It argued that students build their understanding of Newton's third law upon the foundation of a prototypical situation, and then extend their understanding of the law step by step, to other less prototypical situations. This process is quite different from that presented in the conceptual change literature based upon the work by Posner and others (1982), and therefore constitutes an implicit argument against it. In this chapter it is explicitly argued that in developing their Physics conceptions students do not proceed in the overtly rational manner suggested in the conceptual change literature. It will be recalled from Chapter Two that, according to Posner and his colleagues, various conceptions are measured against criteria until ultimately a single conception outranks all others and triumphs. It is argued here, on the contrary, that students maintain a set of conceptions, and from these are selected those which will be used on particular occasions. This is done by a subconscious process similar to that familiar to psycholinguists who study text comprehension (e.g. Gernsbacher, 1990).

In order to study the way that their different conceptions developed, students were asked to discuss their answers to the *Force Concept Inventory* (Hestenes et al., 1992). In order that all students would be prepared for the discussion they were first required to complete the inventory individually. Most of the students in this study, like most students in the world, started with a non-Newtonian world-view, as evidenced by their low scores on this instrument. During the next lesson, normally the following day, they were asked to decide upon the correct answers in groups, and their discussions were audio-taped and transcribed in detail.

A notable difference to the usual way of answering a multiple-choice test was that instead of being asked to give the single correct answer, students were encouraged to indicate which answers they found appealing. In this modified scheme, instead of selecting a single correct answer for each question in the *Force Concept Inventory* (hereafter *FCI*), they were encouraged to rate each option with a score between 0 (indicating that they were certain that this choice was wrong) and 10 (indicating that they were certain that this answer was definitely correct). The total number of points to be allocated for any question was 10, and these could be spread amongst the choices in any way that the student decided. The purpose of this answering scheme was to give credit for partially correct answers. As an example of the sorts of responses that were found, consider question 4 of the *FCI*, reproduced in Box 2.

Box 2: Question 4 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

4. A heavy ball is attached to a string and swung in a circular path in a horizontal plane as illustrated in the diagram to the right. At the point indicated in the diagram, the string suddenly breaks at the ball. If these events were observed from directly above, indicate the path of the ball after the string breaks.



For this question most of the students selected only one answer (mainly the correct one, B) but seven chose to distribute their points as shown in

Table 20.

Table 20: Allocation of points between multiple-choice responses to Question 4 of the *FCI*

Response:	A	B	C	D	E
Student 1		5		5	
Student 2	6	4			
Student 3	7	3			
Student 4	3	7			
Student 5	2	8			
Student 6		7			3
Student 7	1	9			

Further details of student responses can be found in Taylor and Gardner (1999).

An unexpected benefit of this method of answering was that it effectively provided a snapshot of the student's mind before a final answer was reached. Thus, in Table 20, Students 4 to 7 all clearly preferred B. From this it can be inferred that if they had been forced to select a single answer, it would have been B. However, in giving a single answer only, they would have given no indication that they had been considering other answers.

One concludes that the standard method of checking understanding of a concept in physics (i.e. ability to use it correctly to solve a problem) systematically precludes the observation of multiple simultaneously-available conceptions.

Evidence that these conceptions are activated during discussions is given below.

As mentioned in Chapter Two, other studies of transcribed problem solving sessions have been carried out, and have given limited support to the usefulness of the conceptual change framework of Posner. An example discussed was De Jong (1996). In this investigation, however, students were first taught the conceptual change framework and then asked to use it to structure their discussions.

By contrast, the students in the study reported here were not instructed in any theoretical approach, but were simply asked to come to a decision about the answers to a set of particular problems. Far from being conscious of their reasons for choosing a particular conception, students were often unaware that there were alternatives until others brought them up. And far from using generally applicable criteria, the reasons for the use of a particular conception were generally specific to the context and conception under consideration.

As evidence for these claims, some discussions of a problem dealing straightforwardly with Newton's first law are examined. This problem, number four, in the set of twenty-nine in the FCI was shown in Box 1, earlier.

As mentioned above, students had each done the entire *FCI*, including this question, before their group discussion took place. Each student had narrowed down their preferred answer to one or a few choices, which they had rated on a scale from 0 (impossible) to 10 (certainly correct). By the time they tackled this question, they had already discussed three other questions, and had settled into a commonly occurring pattern.

While there was a great deal of variety in the different groups' approaches to the discussions, this pattern seemed to fit all of them. The discussion would typically consist of five main parts. First there would be the *description* of the problem. Secondly, there would be a briefer or longer *pause*, depending on the group and the question, during which students considered their answers. Thirdly they would *compare* answers (and this would sometimes be repeated after some discussion had taken place).

Next, fourthly, they would *argue* over the correct answer. This itself took place in stages. Initially, there would be a proposal, often that the group should settle on a particular option, but sometimes that a particular principle should be applied, or that the discussion should move in a certain way, or that the group should direct its attention to a particular point. The others would respond to this, agreeing, asking for clarification, or disagreeing (either by flatly contradicting the proposal, or by challenging the proposer to provide a justification). This fourth stage, the argument, could sometimes be repeated when different issues were raised.

Finally one or more students would *summarise* the discussion, stating what answer they felt that the group had settled upon.

In Chapter Two a five-stage pattern to spoken narrative talk in groups, identified by Labov and Waletzky (quoted in Chafe, 1995, p. 128) was mentioned. They called these stages *orientation*, *complication*, *climax*, *denouement*, and *coda*. However, telling a story and arguing about the answer to a physics problem are very different situations, and the middle sections of these two patterns are quite different: one does not expect, nor does one find, specialised narrative functions like *complications*, *climax* and *denouement* in the discussions analysed here. Nevertheless, there is some correspondence between the first and last stages of the two patterns.

For Labov and Waletzky's speakers, the *orientation* of a spoken narrative identifies

- who or what is being discussed,
- the viewpoint of the speaker,
- where the action happened and

- what was being done at the time.

Here the same features were put forward in the part of the discussion I have labelled the *description*. This occurred at the beginning of each problem, since students were asked to identify the problem that they were discussing. How they chose to do this gives some idea of how they conceived the context of their discussion, and also shows some parallels to the *orientation* section of a spoken narrative.

Some examples are given in Table 21.

Table 21: Introductions to the discussions of question 4

Group A	Male student 2:	okay number four
	Male student 3:	Number four
	Male student 2:	Number four is about er heavy ball sort of attached to a string and swing circular path
Group B	Female student:	What's number four Oh number four is about the heavy ball that attached to a string being swung in a circular path
Group C	Male Student:	Number four. A heavy ba- heavy ball is attached to a string and swung in a s- circular path
Group D	Male Student:	Question 4 is the one about the .. heavy ball attached to the string and its swung in a circular path.
Group E	Male Student:	Question 4. RD say B.

Clearly they all identify the question by its number on the test paper: equally clearly, most feel that this in itself will not be enough to bring the question to mind. Given that this question is the only one on the test paper about "a heavy ball" they could have said something like: "Number four is the one about the heavy ball." However, no groups chose to do this. All who elaborated at all chose to mention both that the heavy ball was attached to a string, and that it was swung in a circle. This was not just a matter of reading out the first sentence, since this continued with further information: namely that the circle was in a horizontal plane, and that it was illustrated in a diagram.

Like Labov's speakers of narratives in the *orientation* section, the students' descriptions clearly identified

- the referents (the ball, the string),
- the viewpoint of the speaker (onlooker, not involved – not "my" ball, but "a" ball), and
- what was happening (swinging in a circle).

As well as this similarity at the beginnings, there is some similarity at the ends of the two types of discussion. At the end of a narrative, Labov's speakers finished off with a *coda*, which served to highlight the key point of the narrative, and signalled to others that the narrative was complete. At the end of the discussions here, one or more of the students summarise the group's decision, whether this is to choose a particular answer, or a decision to abandon the question without having resolved it. In both cases what is central to the discussion is emphasised and the discussion can then move on to a new topic.

From the resemblance between the features of these Physics students' problem *description* and final *summary* and Labov's *orientation* and *coda*, one can infer that they may have similar functions — something like left and right parentheses in a mathematical expression, serving to bracket off the discussion as a unit. Or, in Gernsbacher's terms (1990) the *description* reflects the process of starting a foundation, and the *summary* signals that a shift will be made to a new structure. Further evidence that each problem serves as a structured unit in the students' memories will be adduced in the analysis of student discussions of question 10.

It will be recalled that the second common stage observed in these student discussions was a *pause* where students considered their answers: a pause is commonly observed at topic boundaries in discussions (Gernsbacher, 1990) and so its presence here is natural, and provides further evidence that each problem discussion forms a single unit. (However, somewhat weakening this argument, in some cases there was no significant pause between stage one, the question description, and stage three, the comparison of answers. Nevertheless, this may simply be because the time taken to describe the question aloud was sufficient for the other participants to switch their attention from one topic to another.)

While this feature of the discussions has been called a pause, it is not intended to imply that this is a period of silence. In some cases, certainly, the pause was silent, but in others it was filled with meaningless vocalisations like these from Group A (xxx indicates something that could not be transcribed because of lack of volume or clarity):

Male Student 2: xxx
Female Student 1: Hmm?
Male Student 3: Ahh
Female Student 1: Mmhmhm

The third stage, where students compared their answers, was no doubt related to the way in which the discussions were set up, whereby all students had done the test before they began. Nevertheless, its universality within this study is worth noting, in so far as no instructions were given to proceed in this way, and logically one might have expected the discussion to begin with premises and build up to the conclusion. In fact invariably all discussions began with conclusions! It may well be that this section consisted of an extension of the essential orientation section: each speaker felt the need to provide their viewpoint on the correct answer. Some examples are given in Table 22.

Table 22: Examples of comparisons of answers (A vertical bar indicates where overlap of speech begins. FS1 = Female student 1, etc.)

Group A	FS1:	B		
	MS3:	B		
	MS2:	Okay B is xxx		
	FS1:	B for boy		
	MS1:	Yah		
	FS1:	B B B		
	FS1:	All B?		
Group B	MS2:	What the answer?		
	MS1:	B		
	MS2:	| B		
	MS3:	| Of course	| B	
	MS4:		| B	
Group E	MS1:	Question 4. Ardy say B.		
	FS1:	B.		
	FS2:	B or E		
	MS2:	E		
	MS3:	B		
	MS1:	So we have again some		| conflict.
	MS2:			| Oh
	FS2:			| {Laughter}

With the first three parts of the discussion out of the way, the stage is set for the discussion proper: the argument. This fourth stage of the discussion is the most interesting from the point of view of the way students develop their understanding of mechanics. One can clearly see the lack of any evidence for the use of general criteria of the sort proposed by Posner et al. (e.g. plausibility, intelligibility, fruitfulness) by students in evaluating conceptions. On the contrary, decisions are made "on the fly" in response to highly specific factors.

Before going into the details of these, it is worthwhile looking at the overall pattern that can be discerned in the arguments transcribed here. What is called an argument here, seems itself to be divided into three parts: the proposal, an evaluation, and a justification. However, with multiple speakers many of the arguments were interrupted and so not finished, or different arguments would overlap, or sub-arguments would develop within the main argument: and these in turn could overlap with others, be stopped short, or interrupted by requests for clarification (for example: Group E, MS2: "Do you say D or B?") and so on.

In spite of the fragmentary nature of some of the discussion, it is still possible to discern some regularity. Typically what I am calling an *argument* would begin with someone proposing that a particular answer (or answers, bearing in mind that the option of distributing points between different choices had been available when they did the test individually) was correct, for example:

Group A	MS:	I suggest..
		I thinks it's A and B

or:

Group E	MS1:	It's E!

The next stage would be an evaluation by one or more of the others in the group: either assent (an example from Group B is in Table 22), or dissent. Dissent can be an explicit contradiction (whether politely as in the example from Group B, or emphatically as in the example from Group D):

Group B	**MS1**:	B
	MS2:	I think A is closer
Group D	**MS1**:	No no no no no

Alternatively a negative evaluation might be implied, rather than explicit, being put into the form of a request for further information.

Group B **MS1**: Ah, why?

Group E **MS1**: So why do you say E?

Group E **FS1**: You have to say more.

At this point the speaker, often jointly with others, as in the extract from Group A below, would attempt to justify their proposed answer. Occasionally, the speaker would provide some extended sequence of points in their argument (e.g. Group A below), but often their argument consisted of simply highlighting some aspect of the situation (e.g. in Group B below).

Group A

MS2:B because

FS1:Because it

like this point

the force is

actually going this way

so I say xxx

go there

unless there's another force

Group B MS1: Because this is the tangent, right?

Interpreting the arguments

The discussion by Group A, immediately above, raises an important point for this investigation, because it can be interpreted in (at least) two different ways, depending on what direction was intended by "this way". Because the students were talking to one another face to face they had access to contextual information like this, which is no longer available in the audiotape. Some of this context might have been retrievable had there been the resources available to videotape all the discussions, but even so there would have been times that the camera was filming from the wrong viewpoint, or had insufficient resolution, to see things like which part of a diagram was being pointed to. Of course this could in turn be remedied by having multiple cameras, or by having skilled camera-operators focussing on relevant details minute by minute. However, quite apart from the expense of such procedures, they would be far more intrusive than the inconspicuous audiotape deck and far more likely to make students self-conscious, thus complicating the interpretation of data in a different way. As discussed in the section on hermeneutics in Chapter Two, there will always be multiple interpretations to choose between for any form of evidence.

Since the data obtained in this investigation includes audiotapes, not videotapes, the limitations of this form of evidence must simply be taken into account.

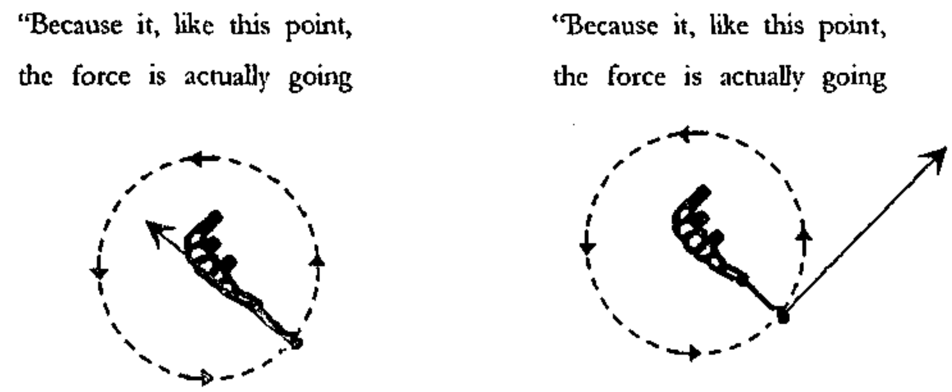


Figure 36: Two possible interpretations of an utterance.

In the discussion by Group A above, all the group members had agreed immediately that B was the correct answer. Then two of them participated in summing up, the first beginning, the second completing the explanation.

On one interpretation, this second student argues that at the point on the diagram, just before the string breaks, there is a centripetal force acting, and just a moment later, after the string has broken, there is nothing to stop the ball from continuing in a straight line, unless another force operates. This is the correct Newtonian explanation.

However, another interpretation exists. The student could be saying that, since there is a force on the ball in the direction of the motion, with the string broken there is nothing to stop the ball from continuing in the direction of the force acting upon it, unless another force operates. This is the common "motion implies force" misconception (Gunstone & Watts, 1985).

Which interpretation is correct? If, by the correct interpretation one means that which was intended by the speaker (Hirsch, 1967), then there is no way of knowing. We do not know which direction the speaker indicated, and she was absent on the occasion of a subsequent test (the Force Conceptual Evaluation test described in Thornton & Sokoloff, 1998), which might have established whether she was more likely to use the Newtonian conception or the "motion implies force" misconception.

If, however, one accepts that the interpretation is in the mind of the listeners (Mueller-Volmer, 1985; Palmer, 1969), then it can be argued that both interpretations are "correct". It is a truism that we hear what we want to hear, and it is a well-established fact that physics teachers who teach Newtonian physics find that they have students who nevertheless have non-Newtonian misconceptions. In this case, either of the two interpretations results in the same (correct) answer to the question. Those of the students who were using the non-Newtonian conception may well have "heard" the corresponding version of the argument. Irrespective of which direction may have been indicated, students may well have "seen" the direction they expected. Utterances are part of the evidence for a conceptual view, but they do not exist in a one-to-one correspondence with views.

Thus, it is argued here, not only can different conceptions can co-exist within the mind of the individual student (as reflected, for instance, in the choices shown in Table 20), but different conceptions may co-exist within apparently unanimous groups.

In other groups, the use of different conceptions by different students was completely obvious. For example, in the discussion of Question 4 by Group 97-63, the student called here MS2 (i.e. male student two) consistently argued for answer B using the Newtonian conception, but was opposed, and supported, by others who used different misconceptions. The discussion exhibited the usual stages described above, and is reproduced below in its entirety to illustrate the rapidity with which arguments were made, and the fragmentary nature of these arguments. Furthermore, the transcript again illustrates the need for some interpretation (mine is indicated in italics) to make sense of the discussion.

A discussion about circular motion

The group consisted of four males. Names have been replaced by abbreviations: MS1 = male student 1, etc. (MSU = male student unidentified). Curly brackets enclose uncertain transcriptions Unintelligible utterances are noted as xxx or yyy. Italicised sections indicate possible interpretations.

Stage of discussion	Speaker	Utterances	Comments and interpretations
Description			
	MS1:	OK ah..	
		question 4 is the one about the ..	
		heavy ball attached to the string	
		and its swung in a circular path.	
		Comparison	
		All of us got ah..	
	MS2:	Not all of us	
	MS1:	B..	
		except for *MS3..	
		who got E.	
	MSU:	{everyone}	
Argument			
	MS3:	it's the force	
		anticlockwise	*Argues using "motion implies force"*
	MS4:	and considering the	
		centriii...petal force	*Possibly hesitating between "centripetal" and "centrifugal".*
		which acts ah	
		directly	
		perpendicular to the ..	*Argument seems compatible with Newtonian conception of force.*
		distance	
		which is..	
		swinging,	
		so its	
	MS1:	that was perpendicular?	*Queries direction of force, which does not fit with using "motion implies force"*
		Why?	
	MS4:	{Yeah ah totally}	
	MS2:	Now actually its not the circular	
	MS4:	xxx	
	MS1:	yyy	
	MS4:	It will fly as the force	*Argues using "motion implies force"*
	MS3:	Maybe I	
		So ah..	
		but it doesn't	
		you have to go with	
		what is written	*Argues for combining everyone's answers: so many points for E, so many for B.*
		make it short	
	MS2:	All right the three of us got	*Argues for answering by voting.*
	MS3:	stay constant	*Refers to Newton's first law*
	MS1:	So ah B	
		Ya so we	
	MSU:	B	
	MS2:	All right so its	
	MS3:	E	*Insists on his own answer.*
	MS2:	But after	*Starts to refer to the situation after the string breaks.*
	MS3:	We we don't care	*Refuses to accept discussion.*
	MS4:	No no no no no	*Insists on his own answer.*
	MS2:	You must	
		You must	
		accept that	*Insists on discussion of reason for answer.*
	MSU:	Its not the value	*The ball doesn't have a force…*
		because its acting on it	*… a force acts on the ball.*
	MS3:	Yeah I'll give the ball	*The ball can be given a force…*

Stage of discussion	Speaker	Utterances	Comments and interpretations
	MS2:	No. No.	*The ball cannot be given a force…*
		When when when its ah..	
		no force acting	
		it just ah	*Trying to state Newton's first law as applied to this situation.*
	MS1:	flying	*Supplies word for "continues moving" to finish off sentence.*
	MS4:	when I've got something that's no force acting it will swing here..	
		in a circle	*Explicitly states that circular motion requires no force to maintain it.*
	MS1:	but we're releasing that	*Directs attention to the fact that the ball is released from the string.*
	MS4:	if..	
		you know that	
		releasing what	*Asks what is referred to by 'that' in 'releasing that.'*
	MS3:	{Anyhow taking any route}	
	MS4:	suddenly break	*Directs attention to the fact that the string breaks.*
	MS1:	Where ..	
		is the opposite?	*Directs attention to idea of "equal and opposite forces"*
	MS2:	No no.	*Rejects discussion of "equal and opposite forces" as irrelevant*
		No force.	*Insists that there is no force acting.*
		No force ..	
		ah when ..when it	
	MS3:	{Many I'll take the ball }	
	MS1:	This guy here	*Attempts to isolate opponent from peer group.*
	MS3:	Its E!	*Refuses to cave in to peer pressure.*
		{laughs}	
Summary			
	MS1:	Okay we we have um ah	
	MS4:	We haven't decided which answer	
	MS1:	{laughs}	
	MS2:	The three of us got B..	*Argues for answering by voting.*
		so its B	
	MSU:	{laughs}	
	MS4:	Ah OK	
	MS1:	Lets move to five.	
	MS3:	Good luck.	*Dissociates himself from final decision.*

What becomes increasingly obvious as one listens to such discussions is that there is a wide variety of approaches which students can use to answer these questions without using Newtonian physics. Of the four students in the above discussion group, only one consistently held to the necessity of using Newton's first law, yet three of the four had the correct answer.

The evidence suggests that rather than students *having* either the Newtonian conceptions or "motion implies force" misconceptions and therefore employing them in solving problems, that they instead solve problems by utilising whatever comes to mind. What comes to mind may indeed be the correct Newtonian conception, as it was for student MS2 above, or it may be any of the widely attested misconceptions, such as "motion implies force" (MS4 above) or "conservation of circular motion" (MS3 above). But it may be something as strictly speaking irrelevant (and yet as effective!) as choosing the correct answer by voting (MS1 above), or recalling from previous experience that in this sort of problem the tangent has a special role:

Group 97-51: MS1: Because this is the tangent, right?

Or it may be personal experience outside of the classroom that is called upon:

Group	Speaker	Utterances	Comments and interpretations
97-91	MS1:	Mana boteh A! A...Wei. You know when people throwing hammer in the in the.. in the sports, they they they spin around and when they...	*{It cannot be A!}*
	MS2:	Yeah, that's why it goes straight.	

Thus, instead of more or less extended sequences of reasoning, much of the discussion consists of fragmentary phrases directing others attention to the same points that one finds compelling. A quick glance at the Group 97-63 discussion shows many examples of such fragments, for example:

Group	Speaker	Examples of utterances which direct attention to a particular point
97-63	MS3:	it's the force anticlockwise
	MS4:	and considering the centriii...petal force
	MS4:	It will fly as the force
	MS1:	but we're releasing that
	MS1:	Where .. is the opposite
	MS2:	No force

Utterances like these are simply uninterpretable in terms of the theory of conceptual change put forward by Posner et al. They cannot be considered as rational arguments, and they do not have any straightforward relationship to dissatisfaction with current conceptions, intelligibility of proposed conceptions, fruitfulness of new conceptions or their plausibility. Yet utterances like these are very common in these transcripts, and the students are clearly highly involved in their production, comprehension, and in responding to them.

These utterances, which are so fragmentary in appearance, can be given a complete form when considered within their context. The groups are co-constructing in their dialogue a mental model of the problem solution, and these utterances serve as prompts that guide students to contested sites in that model. When a site in the mental model is highlighted, the students who are participating in the dialogue momentarily focus their attention at that point. In effect, speakers in mentioning a topic are able to modify the activation of that particular in their auditors, a pre-rational process. Because the students are actively involved in the same task, it is not necessary for the speaker to provide fully structured arguments: the other participants see (or think they see) where the speaker is headed and respond accordingly.

Summary

Evidence gathered from a 'probability format' test of students working on questions on the *FCI* showed that multiple conceptions were available to students as they attempted the problem solutions.

Transcripts of discussions showed that many fragmentary contributions to the arguments were best interpreted as attempts to highlight, or activate, particular ideas in their listeners.

The way that students organised their discussions was argued to be due to the way each new problem situation required that the students mentally establish a foundation, and map new information onto that foundation. Further, when a new problem was encountered, the previous problem representation had to be mentally filed away while attention was shifted to constructing the next problem representation.

Student discussions of Newton's first law problems showed that in developing their solutions students were opportunists: any method which led to an answer would be welcomed. Personal experiences, or an appealing schema like "More of A – More of B," were as likely to be activated as Newtonian mechanics. When one student disagreed with another they would direct each other's attention to key points rather than put forward logical arguments.

The process of building an understanding was jointly undertaken and does not proceed by means of a rational weighing up of fully expressed arguments: rationality in the use of concepts is, here, an achievement of the process of discussion, not the precondition for it as implied by the Posner et al. theory.

Newton's second law

Coming to an understanding of Newton's second law of motion is a complicated process for students.

Most often this law is presented to students in the form $F = ma$, and is quickly learned in this mathematical form. However, testing which is restricted to substitution into this formula will not uncover the serious conceptual difficulties students face in understanding it.

These difficulties can be subdivided into four main areas:

- The concept of net force.
- Interpretation of the equation.
- The distinction between mass, size, and weight.
- What counts as acceleration.

One could say that the easy part is $F = ma$, but the difficulties are

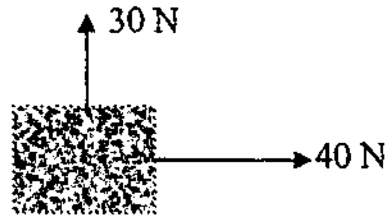
- F
- $=$
- m
- a

Substitution into the formula $F = ma$

Newton's second law is usually introduced as the equation $F = ma$, and there can be no doubt that students soon become adept in its use in this form. For example, the chapter on Newton's third law discussed a year 10 class which had had four lessons on Newton's laws. In the final test (Appendix C), every student in this class was able to answer a question requiring substitution into the equation for Newton's second law, question 20 (see Box 3), by dividing their answer to question 19 by five (not all were able to answer question 19 correctly).

Box 3: Questions 19 and 20 from Year 10 test (Appendix C)

19. What is the size of the net force on this box, which is being pulled along a frictionless surface by two people pulling on ropes with forces of 30 N and 40 N?



20. The box in question 19 has a mass of 5.0 kg. How much does it accelerate?

Furthermore, the senior students who were discussing problems from the Force Concept Inventory would refer to the equation to bolster their arguments, confident that their partners would understand what they meant. For example, in discussions of many different questions (the details of the questions do not matter for the current point, but they can be found in Appendix 1 if desired), students would spontaneously quote the formula F = *ma*:

Question	Group	Speech
3	97 6-3:	... because because this depends on the mass *F* is equal to *ma*
5	97 5-1:	Now if we're talking about force, Force is mass times acceleration, right,
9	97 5-3:	... because because this depends on the mass *F* is equal to *ma*
28	97 5-3:	This force plus this force xxx, no force, net force zero, net force zero, *F* equals to *ma*, net force is zero, *a* equals zero

It is clear, therefore, that these students are familiar with and confident in their knowledge of this formula.

Interpretation of the formula

It has often been reported that students have difficulty with relationships between three or more quantities. In particular, where a relationship involves three quantities, students tend to ignore one of them, and deal with only two at a time. For example, given the formula $v = f\lambda$, and asked what happens to the speed of a sound if its frequency is doubled, students often incorrectly answer that the speed is doubled. This reflects competence at arithmetic and algebra, unsupported by understanding of physics: since the speed of sound is a constant for a given medium at a given temperature, an increase in frequency is counterbalanced by a corresponding decrease in wavelength. The student error may be because, in answering the problem, the student has activated a learnt procedure for dealing with an algebraic relation without activating the required physical data needed to use the procedure correctly.

When using Newton's second law, students will often assume that since $F = ma$ an increase in mass will mean an increase in force.

On the one hand, when the acceleration is constant, reasoning in this way can lead to initially correct conclusions: the force of gravity *is* larger on larger masses.

Box 4: Question 1 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

1. Two metal balls are the same size, but one weighs twice as much as the other. The balls are dropped from the top of a two story building at the same instant of time. The time it takes the balls to reach the ground below will be:

(A) about half as long for the heavier ball.
(B) about half as long for the lighter ball.
(C) about the same time for both balls.
(D) considerably less for the heavier ball, but not necessarily half as long.
(E) considerably less for the lighter ball, but not necessarily half as long.

On the other hand, this can lead to immediate errors: as seen in the chapter on Newton's third law, and in the examples from groups 97 6-4 and 6-5 on the next page, students can argue that in a collision between a truck and a car, the truck exerts a larger force because of its larger mass.

Furthermore, while reasoning in this way can lead to initially correct conclusions, it can then lead to further conclusions that are incorrect: for example that an increase in mass results in an increase in force, and an increase in force results in an increase in acceleration, so that heavier things will fall faster. This exact chain of reasoning is seen in one of the student discussions of the answers to question 1 from the Force Concept Inventory (see Box 4). Female speaker 4 (FS4) in group 97 11-1 says:

My answer's D because.. the heavier the mass .. the mass.. that means the weight is greater.. the gravitational field is greater.. that means the time it takes to drop down from the building must be less.

The key steps in her argument are: "the heavier the mass .. the mass.. that means the weight is greater" – This is true, since the weight force = mg, an increase in m results in an increase in the weight force, since g is constant.

"The weight is greater.. the gravitational field is greater" – This is false: since the weight force is equal to mg, an increase in the weight is the result of an increase in m, not in the gravitational field strength which is a constant $g = 9.8\ ms^{-2}$.

Rather than deal with three variables that are related to each other, students simplify it into a relation between two variables at a time, where whenever one increases the other will also.

Box 5: Question 3 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

3. Two steel balls, one of which weighs twice as much as the other, roll off of a horizontal table with the same speeds. In this situation:

(A) both balls impact the floor at approximately the same horizontal distance from the base of the table.
(B) the heavier ball impacts the floor at about half the horizontal distance from the base of the table than does the lighter.
(C) the lighter ball impacts the floor at about half the horizontal distance from the base of the table than does the heavier.
(D) the heavier ball hits considerably closer to the base of the table than the lighter, but not necessarily half the horizontal distance.
(E) the lighter ball hits considerably closer to the base of the table than the heavier, but not necessarily half the horizontal distance.

In fact, this form of reasoning, dubbed "More A – More B" by Stavy and Tirosh (1996; 2000) was very noticeable throughout all the student discussions. According to Stavy and Tirosh, students naturally work with an intuitive rule: more of one quantity will lead to more of another quantity. So an important difficulty that students must overcome, before they can understand Newton's second law, is that they must somehow overcome this natural tendency to interpret an equation as an example of the schema "More A – More B."

That students do regularly use this schema is evident from the transcripts of their discussions. Examples include the following, from the discussions of the first three questions. These examples show that this schema is widely utilised, and is not restricted to the (mis)interpretation of Newton's second law.

Question 1: more weight means more effect

Group 97 5-2:		But it will affect the the one with the bigger weight more than

Question 2 (see Box 1 previous chapter): more mass/size means more force

Group 97 6-4:		but the mass of the truck is greater so there is more force from the truck
Group 97 6-5:		E! but you must imagine the size of the truck more force

Question 3 (see Box 5) more mass means more distance

Group 97 11-1:		You see right the lighter b- the the heavier b- I mean some of you studied inertia before right so that the heavier ball it takes more force to stop it right and you'll agree it'll go further right ... It's just like you are you are you are swinging ah the bell what is hard to stop
Group 97 5-2:	MS3:	the the bigger the ... in the the the more mass the ... the ball ... go faster and go further goes further
Group 97 5-2:	MS3:	um for the bigger m bigger mass well we got them more distance right
Group 97 6-2:	MS1:	Yeah. I'm I'm I mean I think the heavier ball heavier the heavier one will be ah go further than the lighter one ...

than the one
than the heavier one
the heavy one will go
further. I mean
a ball leaves a table and
then on the ground the further
ah

MS3: the mass
MS1: the mass

The transcripts of the discussions of these three questions record a total of 949 utterances, and 73 of these, or 7.7%, are involved with putting forward examples of this 'More A – More B' schema.

Given that assimilating Newton's second law to this 'More A – More B' schema is naturally appealing, how can students argue against it? Five types of counter-arguments were present in the transcribed discussions:

- use of an alternative schema: 'Same A – Same B'
- arguing that a quantity was irrelevant
- reference to equations
- appeal to experience
- appeal to 'fudge-factors'

Alternative schema: 'Same A – Same B'

Stavy and Tirosh (2000, pp. 42-63) show that across many different subject areas, students appeal to a schema which suggests that whenever quantity A is the same in two instances, then quantity B will be the same also. A few students made appeals to this schema in discussions of these questions.

Question 1: 'Same A – Same B'
Group 97 5-2

MS1: Same size. Everything the same — *i.e. 'Same size – Same everything'*

Question 3: 'Same A – Same B'
Group 97 6-2

MS3: I say both are
because they're the same speed
you know — *i.e. 'Same speed – Same distance'*

MS1: Yeah. Yeah the same distance I
think

Group 97 11-1

MS2: they are the same time
to ...
same height
at the same time — *i.e. 'Same height – Same time'*

Effectively, these statements work in the argument by offering as a focus of attention quantities which are the same in the two instances (here, *size*, *speed* and *height*) instead of differences (*mass*), and appealing to the 'Same A – Same B,' schema. (While these assertions were in fact correct, they were put forward without further justification and were presumably reliant on their appeal to this schema.)

Irrelevance

When students argued that some other quantity depended on the mass, a common counter-argument was to claim that mass was simply irrelevant. Again, this shifted the focus of attention away from the differing quantity *mass*, and weakened the appeal of the 'More A – More B' schema.

Question 1: mass is irrelevant

Group 97 11-1

MS3:	Because it doesn't depend on the	*Second student completes first student's statement, clarifying what is referred to by 'it'*
MS2:	mass gravity	

Group 97 5-1

MS1:	the time does not depend on the weight	*Two students complete each other's sentences to argue that the weight is irrelevant. While correct in this, they are incorrect in conflating gravitational acceleration and gravitational force.*
FS2:	Why?	
FS1:	... because weight	
MS1:	xxx the gravity ...	
FS1:	gravity they are both being pulled at the same ... force	
FS2:	... it doesn't depend on the mass of the ball so	*The student who asked 'why' is convinced.*

Group 97 5-2

MS1:	Never consider the mass during the calculation for gravity

Question 3: mass is irrelevant
Group 97 6-4

No I think it will be the same
because ah the mass is not
calculated

Equations

The last two examples from the previous section are closely related to the category of argument via equations. Rather than simply state that mass is irrelevant, these students argued that the problems could be solved by equations without needing to know the mass. The arguments were not necessarily valid. Of course, the same style of argument could be utilised by those arguing that mass was important, and their arguments, too, might be invalid.

Question 1: equations
Group 97 5-2

MS2:	No the formula, right xxx *u v*	*Three students cooperate to dredge up the appropriate formula from the recesses of their memories.*
MS1:	Vertical calculation s equals to what?	
MS3:	No ah *ut* plus half *a t* squared	

Group 97 5-4

	Okay. My reason is that umm we're using the equations of motion: *v* is *u* plus *at* and um acceleration is a constant for whatever whatever whatever the mass is so since they both started with zero, they should both accelerate with the same speed they should both have the same speed when they reach the ground so the time taken should be the same so that's why I said C	*One student presents a complete well-formed solution.*

Question 3: equations
Group 97 6-3

MS2:	but the speed is different for that one you see	*Starting from the fact that potential energy refers to mass, these students incorrectly argue that the velocity will be dependent on it. They have neglected the mass dependence of kinetic energy.*
MS1:	Yeah, it's different because of because of their mass From here the *mgh* right and when you're over here it become all kinetic energy so the *mgh*	

MS2: oh

MS1: xxx the heaviest mass
the larger

MS2: velocity

Group 97 6-3

MS1: the heavy one will go
further. I mean
a ball leaves a table and
then on the ground the further
ah

MS3: the mass

MS1: the mass

MS2: yep

MS3: it must here from here

MS1: Hold on. How to express it. Hold
on
Yeah from here to
about here and
xxx
... xxx

MS4: well

MS1: When you push
you give the same push on the
ss .. the the ball
it's the same

MS3: but the acceleration's different

MS2: ...
yep

MS1: ...
because
because this
depends on the mass
F is equal to *ma*

Three students combine to argue against the fourth (MS4, whose responses are not included in this extract) that the heavier ball will go further. However, they unconsciously use the 'More A – More B' schema repeatedly:
MS3 focuses on the bigger mass as the explanation for the greater distance, and is echoed by MS1 and supported by MS2.
MS1 argues that even though 'you give the same push' that 'the acceleration's different'. He appeals to Newton's second law, for support. However, while it is true that the acceleration is different for different masses when the same force is applied, the difference is in the opposite direction to his statement: the larger mass accelerates less.

Experience

Another way students argued for or against the use of the 'More A – More B' schema was to appeal to experience, either personal or shared. In the discussions of question 1, only one reference to personal experience was found, but in the discussions of question 3 there were many references to a shared classroom experience, a particular practical exercise.

Question 1: arguing from experience against 'More A – More B'

Group 97 6-2

Ah
But where
does whatever I
I
I experience like
with xxx...
another dish
when I

While presenting neither a coherent nor a clearly expressed argument, this student is nevertheless employing the strategy of referring to experience.

its drop it
get down
he said that was the same time so
I think its pretty much the same

This practical exercise (which was designed to teach about the role of uncertainty, see Appendix D for details) was referred to repeatedly in the discussions of question 3. Briefly, in this practical exercise, balls of different mass, density and radius were launched horizontally from the same ramp on a table top, and the students had to determine whether or not they were launched with the same speed. The students had to do this by recording the positions at which the balls landed on the floor below. For each particular type of ball the landing points clustered close together and each of these clusters of landing points was at a distance from the other clusters. By comparing the within-cluster variation with the between-cluster variation students were expected to conclude that the balls were indeed launched with different speeds. In their practical reports they did so. However, by and large, they did not recall that the balls were launched with different speeds in these discussions. Overwhelmingly, they believed that the results of the experiment confirmed their answer to question 3: that is, that balls of different mass launched with the same speed would travel different distances.

Question 3: arguing from experience for 'More A – More B'

Group 96 B

F1
It was just like .. the last time we did
that roll-off thing

Group 97 9-1

MS4:
No. No. Not if you go from the
practical. {laughs}

Group 97 5-1

FS1:
But why
FS3:
Because from the experiment
you remember like the ball one

Group 97 5-2

MS4:
...
I don't think its because of the mass
because if the size the same
the three balls are
MS3:
Yeah
but we got answer depend
different for the big
MS1:
The
...
We got bigger distance for the

smaller ball
even though

Group 97 6-1

FS1:
remember that experiment that we
are …

Group 97 6-5

MS3:
My answers are
.. because
xxx
remember we did last time

Group 97 11-1

MS5:
Is it based on the experiment

Group 97 11-3

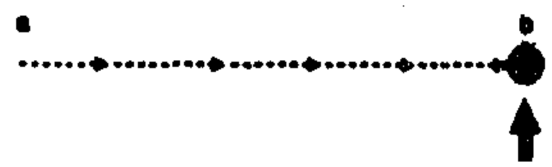
MS1:
As for question number three
I think the answer is
because of the experiment we did

Effectively, the surface features of the question (two balls rolling off a table, and landing some distance away on the floor) often overcame the structure of the problem (comparing launch speeds for different distances vs. comparing distances for identical launch speeds) in the contest for attention in the students' minds, as is typically the case with novice as opposed to expert problem solvers (Chi, Feltovich, & Glaser, 1981; Chi, Glaser, & Rees, 1982). As a result students often initially based their problem solution on trying to recall the result of the experiment without reference to the physics involved.

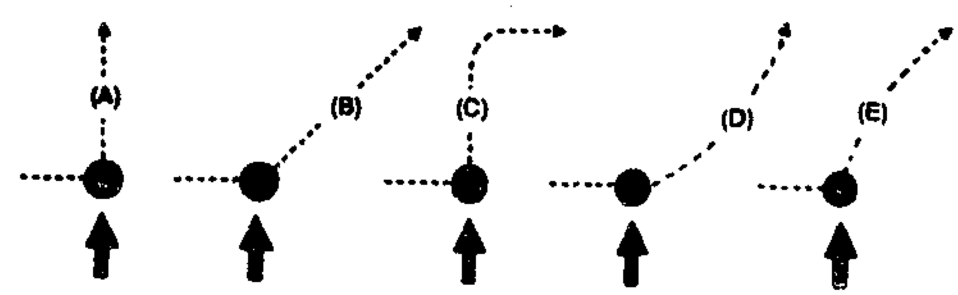
While question 3 had by far the largest number of discussions of experiences, the strategy of arguing from experience was not restricted to question 3. It is clearly a general strategy used in answering any question which, for whatever reason, activates memories which seem relevant to its solution. For example, in discussion of question 6 (see Box 6), several students appealed to their experience of sports. The question itself was about hockey, and at least one student drew on his experience in this game, but others were reminded of their experiences in playing ball games, such as soccer.

Box 6: Question 6 of the Force Concept Inventory
(Hestenes, Wells & Swackhammer, 1992).

* Use the statement and diagram below to answer the next four questions:
* The diagram depicts a hockey puck sliding, with a constant velocity, from point "a" to point 'b" along a **frictionless horizontal surface.** When the puck reaches point "b", it receives an instantaneous horizontal "kick" in the direction of the heavy print arrow



6. Along which of the paths below will the hockey puck move after receiving the "kick"?



Question 6: the effect of a force at right angles to the motion

Group 95 2-4:

MS2:
Why did you say B?
MS1:
Because I play hockey, and ... the thing will go like this.

{*and again, much later*}
MS1:
Ah ..so change it. Because.. Yeah because you know ..it it must be like that. Because when.. if you playing ..at the wing when they hit the ball ..they will go straight.. to the goal, not, not, not perpendicular

This group discussed the question twice, at first inconclusively, then revisiting it. Student MS1 correctly argued that the ball would change direction to follow path B (see Box 6*), basing his argument on personal experience.*

Group 97 6-2

MS2:
number six is about um
MS3:
hockey puck
MS2:
a hockey puck
yep um
I I choose A because, like
when when you're playing soccer ah right
when when you see a ball and you tackle it and it go

MS2 argues, incorrectly, for alternative A, based on his experience of soccer. Other students disagree with the conclusion drawn from the experience, and persuade him.

to ah
perpendicular di direction if you
tackle it from xxx
MS1:
No no
{*Some discussion omitted…*}
MS2:
that happens because like
umm
it goes straight right xxx because of
friction or sometimes when you
pass horizontally or
MS1:
xxx the angle
MS3:
Maybe you are not you are not
kicking at that angle, right
MS2:
yeah xxx

Group 97 5-1

FS3: In reality, if the ball come here and I kick the ball it will go straight here lah
FS1: In reality. But in reality there's friction.

Again the problem with the hockey puck activated memories of kicking a ball. Here FS3 is arguing, incorrectly, for alternative A. Again, the conclusion is disputed.

For each of the students quoted in this section experience seemed to offer a quick answer, making it unnecessary to use the physics theory they had learnt. However, because they were involved in discussions with other students who had taken a different route to the solution of the problem, they were often forced to confront the discrepancy between their experience-based solution and others' theory-based solutions.

'Fudge-factors'

The last two examples above illustrate that there is a certain amount of 'slippage' between real life experience and physics. Students need not accept real world knowledge as a transparent view of physical reality, nor do they necessarily require that physics-theory correspond straightforwardly with personal experience. When confronted with someone who claimed that their experience led to a particular answer, others were quick to point out that there is always the possibility of complicating factors that make a straightforward application of the experience to the problem solution incorrect. In these two instances the complications cited were friction (both times), and the possibility that in the speaker's experience the kick may not have been perpendicular to the ball's path.

Of these two suggestions, the second is almost certainly correct. In order that the ball should move in direction A, the net force on the ball would need to be in the direction shown in Figure 37.

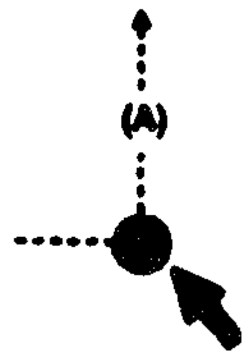


Figure 37: The direction of the net force on an object which turns 90º

The other suggestion, that friction might account for the difference between experience and theory, while more popular (appearing in these instances twice, but also in many other discussions of other questions) does not in fact account for the discrepancy. Friction might slow a hockey puck or a soccer ball along either its initial or final path, but it would not affect the direction at which the ball moved.

Box 7: Question 8 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

> Along the frictionless path you have chosen, how does the speed of the puck vary after receiving the "kick"?
> (A) No change.
> (B) Continuously increasing.
> (C) Continuously decreasing.
> (D) Increasing for a while, and decreasing thereafter.
> (E) Constant for a while, and decreasing thereafter.

This seems to imply that these students are treating the effects of friction very loosely, not considering the actual effects of this force, but instead treating it as one of the set of all-purpose explanations (air-resistance and air-pressure sometimes fill a similar role) to be called upon whenever there is an apparent difference between the predictions of physics and the facts of real life. That is to say, it is treated as a 'fudge-factor.' That friction, air resistance and other factors are grouped together in the minds of at least some students is reflected in the following quotation.

Question 3

Group 97 11-1

MS1:
but xxx
you don't consider all those
frictional forces and air resistance
and shit like that you know
MS2:
{=laughs}...

An example where air resistance is introduced without considering its actual effect:

Question 8 (**see** Box 7)

Group 97 5-4

Maybe if you take into account air
resistance – maybe constant for a
while then decreasing

This student is suggesting that taking into account air resistance would result in answer C: that the hockey puck after being hit and while travelling on a frictionless surface would continue for a short time with a constant velocity and then decelerate. Clearly this is incorrect, since by Newton's second law a resultant force will cause acceleration immediately, not after some delay. Air resistance is not being taken into account as an additional force: it is being used as a 'fudge-factor.'

That the forces of air resistance and friction are not always considered on a par with other forces, leads to the question of how students use the concept of force in Newton's second law.

Interpretation of *Force* in Newton's second law

In the equation for Newton's second law, $F = ma$, the F refers to the net (i.e. resultant) force, that is to say, to the vector sum of all forces acting on the body of mass m. However, this is not necessarily how the students interpret it. Other ways of interpreting it include regarding force

- as a property of the body
- as a continuing influence or as a transient action
- as a single force acting on the body

Force as a property

It has been suggested that force can be seen by students as a property of a body (Brown, 1989). Students, Brown argued, worked with the idea of force as if it was a property of a single body, rather than an interaction between two bodies.

There are, however, two ways, at least of interpreting this suggestion. 'Force' might be seen as a semipermanent attribute, like a thing's colour. On this interpretation, a strong man, or a heavy truck, might be thought of as possessing the property of having force irrespective of the situation in which he or it was involved. Strong men or heavy trucks might be thought of as 'forceful.'

On the other hand, the idea of force, as a property of a body, could be interpreted as a more or less temporary attribute like temperature, which is dependent on recent history. A cup of tea is hot when it is freshly poured, but one wouldn't think of warmth as being a long-term property of it. Similarly, a ball might be thought to have a property of force temporarily: if it had just been bowled, for example.

From the discussions looked at here, there is no evidence suggesting that the first of these interpretations is plausible. One place one can look for such evidence is in the language used to refer to forces. Expressions such as 'force of gravity' and 'force of 10 newtons' are commonly used so when one hears expressions referring to the force of a body it is usual to accept them as natural. However, when students refer to the force of a body, they might be thinking in terms of force as a semi-permanent property of a body. In these transcripts expressions like 'force of the man,' 'the force of the boy,' 'the force of the car,' or 'the force of the truck' all occur. Even so, examining the contexts where these expressions occur, one sees that the students are not referring to a property of a body, but rather to forces which are identifiable by their agents. For example, in discussions of question 19 of the Force Concept Inventory (see Box 8) the following exchange occurred:

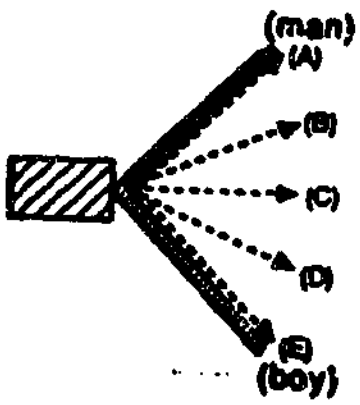
Group 95 2-4

MS1: {Reads:} "Two people, a large man and a boy, are pulling as hard as they can on two r" ...Blah-blah-blah blah blah blah blaaaah, blah blah blah blah blaah... The answer is...
FS1: Why do you say its B?
MS1: Yes B. Because the force is not the same. The man is...
MS2: The force of the man is greater

MS1: greater than the force of the boy.. so the...the...
MS2: the force of the boy
MS1: so the... the thing will go like this.

Box 8: Question 19 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

19. Two people, a large man and a boy, are pulling as hard as they can on two ropes attached to a crate as illustrated in the diagram to the right. Which of the indicated paths (A-E) would most likely correspond to the path of the crate as they pull it along?



It is clear that the students are aware that the forces referred to are the force by the man pulling on the rope and the boy pulling on the rope. Male student 1 (MS1) starts to explain his answer by saying 'The man is..,' presumably meaning to complete the sentence by some phrase like 'stronger' or 'pulling harder.' MS2 interrupts with the alternative phrasing 'The force of the man is greater' and this phrasing is accepted by MS1 who incorporates it into his explanation. There is nothing here to suggest that these students are thinking of the man as having a semi-permanent property of force.

Box 9: Question 13 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

Refer to the following statement and diagram while answering the next two questions.

A large truck breaks down out on the road and receives a push back into town by a small compact car.

13. While the car, still pushing the truck, is **speeding up** to get up to cruising speed:

(A) the amount of force of the car pushing against the truck is equal to that of the truck pushing back against the car.

(B) the amount of force of the car pushing against the truck is less than that of the truck pushing back against the car.

(C) the amount of force of the car pushing against the truck is greater than that of the truck pushing against the car.

(D) the car's engine is running so it applies a force as it pushes against the truck but the trucks engine is not running so it can't push back against the car, the truck is pushed forward simply because it is in the way of the car.

(E) neither the car nor the truck exert any force on the other, the truck is pushed forward simply because it is in the way of the car.

Similarly, in a discussion of question 13, which describes a completely impractical situation in which a car is supposed to push a truck along the road, (see Box 9,) the student referred to 'the force of the car':

Group 97 6-5

but if the force of the car is pushing with ten newtons the truck is also pushing back with ten newtons

Again, the phrasing 'the force of the car' alternates with the 'the truck is also pushing' showing that these are seen as different ways of expressing the same idea and there is no reason to believe that this student thinks of the car as having a property called force.

However, in the examples so far discussed, the students have been discussing questions where the forces are fairly prototypical. They each exhibit three or four of the four features which were found to characterise prototypical forces in Chapter Five.

There are clearly two objects interacting (car and truck, or man and box, or boy and box), the forces have observable effects (the box is pulled along the floor, the truck speeds up), the forces occur at a particular time, and involve effort (although this, in the case of the inanimate car, is only figuratively so).

In other questions, particularly those involving projectiles, one of the two interacting objects (the ball) is salient but the other (the Earth) is not. It is in these questions that evidence for the force as property conception can be seen. However, it will be seen that this evidence is clearly related to the second interpretation of this idea: force is seen as a temporary attribute that things get, and then lose. That some students are using the idea that force can be transferred from one thing to another may be inferred from discussions such as these:

Question 5 (see Box 10 p.197)

Group 97 6-3

MS1:

But you see, like, if you throw a ball, you give a force to the ball, is it. ... Ohh

The key point illustrated here is that for this student 'force' is a temporary attribute which can be given to the ball.

This conception of force treats 'force' in this situation as if it can be transferred from one object to another, and can be used up. It is sometimes referred to as the 'impetus theory,' so-called because of its similarity to a medieval theory of motion (Franklin, 1976). It has also been pointed out that this conception is similar to the scientific concept of momentum (Grayson, 1994). In terms of prototype theory as presented here, students who have a prototypical conception of force which is scientifically acceptable (a hit or punch is a perfectly acceptable example of a Newtonian force) can nevertheless include within the boundaries of their concept situations which are not included in the scientifically valid conception.

Figure 38 attempts to illustrate this point: the extension of the Newtonian concept of force does not include all that is included in the extension of student's concept of force. In particular, with respect to the everyday use of the concept of throwing, there are aspects which are not covered by the Newtonian concept of force. This occurs because, in common with many other actions which take place over a time interval, the action of throwing has what have been called a **source**, a **path**, and a **goal** (Jackendoff, 1983). When one throws something one may be referring to the initial action of the throw (as in 'He didn't drop it, he threw it.'), the path followed (as in 'She threw it over the fence.'), or to the endpoint (as in 'She threw me a sweet.'). Only the first of these would be a situation in which the throw would be a force in the Newtonian sense.

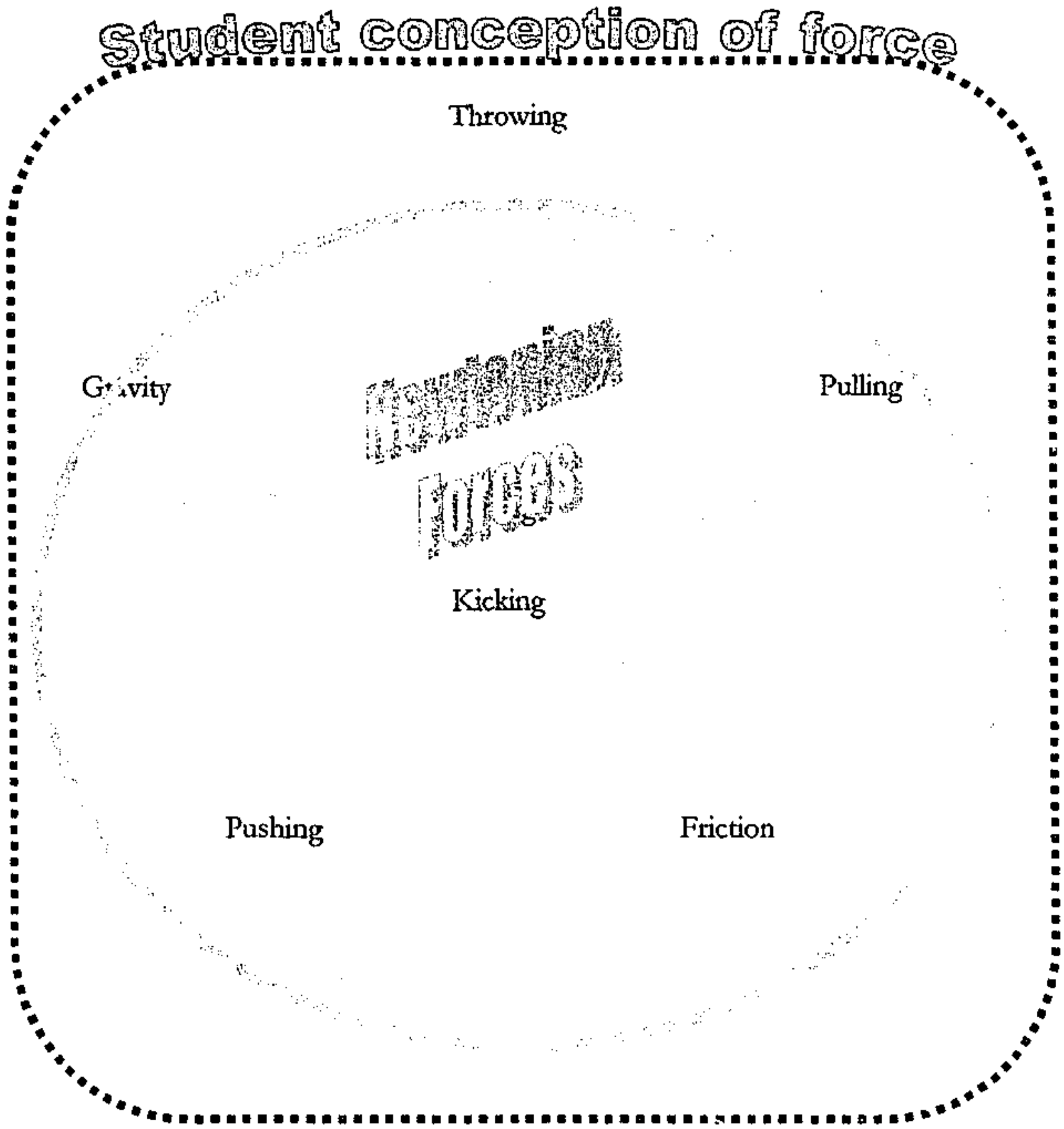


Figure 38: Relationship between the extension of the Newtonian concept Force (solid line) and a hypothetical student's concept Force (dotted line).

In her research, Grayson found that by accepting that there was a common factor in the situations such as that of a thrown ball, or a rolling car but relabelling this as 'momentum' and carefully distinguishing it from the Newtonian concept of force, that her students were able to accept Newtonian conceptions more readily. Grayson suggests that the success what she calls 'concept substitution' depends upon accepting the students' intuitions that there is 'something' which is transferred from one object to another in an interaction, but then relabelling this 'something' as 'momentum' rather than as 'force'. It could be argued that in addition, this process accepts the students' basic understanding of force, clarifying the boundaries of the Newtonian concept of force, without rejecting the student's own understanding in its entirety.

Force as a continuous influence vs as a transitory effect

Similar to the idea of force as a property of an object, but distinct from it, is the idea that a temporary force has a continuous influence. In particular, some students think that a throw can result in a force that continues throughout a projectile's flight. This is illustrated in the following extract from a discussion of question 5:

Question 5 (see Box 10, p. 217)

Group 97 5-3

MS2: upward force um 'cause you're throwing that ***MS3:*** xxx ***MS1:*** Yeah	*The student refers to an upward force which is present, but does not refer to it as a property of the ball.*
MS2: So there is upward force but it decreases huh I think we should talk there	*This force, he proposes, gets weaker.*
MS1: Why why is it up? ***MS3:*** By the constant downward	*This view is challenged.*
MS2: Because if there were no upward force then xxx	*While his words are not all audible, he is clearly justifying his view by appeal to the consequences which would arise should this force not be present.*
xxx like that ***MS4:*** So what is the upward force ***MS2:*** The throw up, man ***MS4:*** Throw it ***MS2:*** Yeah throw and it decreases as it goes up Something like that I think	*Unsure of how to rebut this, another challenger switches tactics and asks what this force could be. The reply, identifying this continuing force with the throw satisfies the group.*

These students accept that the force of the throw is a continuing influence on the motion of the ball. One can see that for some forces, such as gravity, this idea is correct: the force of gravity is a continuing influence on the motion of the ball. However, for other forces, such as a hit or kick or, as here, a throw, this is plainly incorrect.

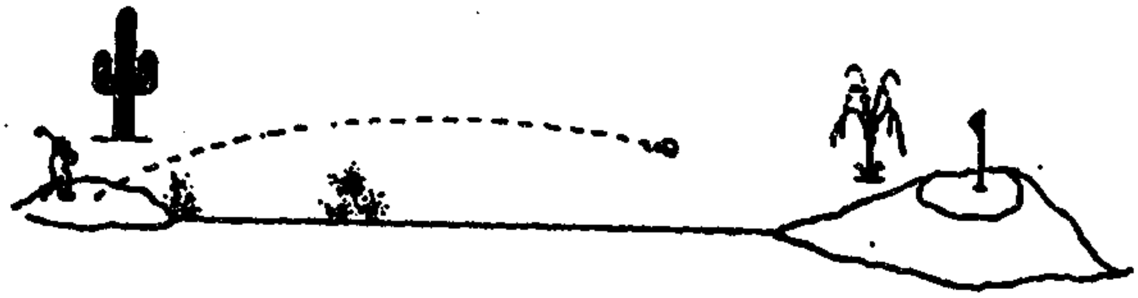
Such incorrect ideas are also found in discussions of question 22. However, one should be aware that in these discussions the use of the phrase 'force of the "hit"' is not in itself sufficient to indicate that this idea is being used by the students, as this phrase is present in the question.

Box 10: Question 5 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

5. A boy throws a steel ball straight up. **Disregarding any effects of air resistance,** the force(s) acting on the ball until it returns to the ground is (are): (A) its weight vertically downward along with a steadily decreasing upward force. (B) a steadily decreasing upward force from the moment it leaves the hand until it reaches its highest point beyond which there is a steadily increasing downward force of gravity as the object gets closer to the earth. (C) a constant downward force of gravity along with an upward force that steadily decreases until the ball reaches its highest point, after which there is only the constant downward force of gravity. (D) a constant downward force of gravity only. (E) none of the above, the ball falls back down to the earth simply because that is its natural action.

Box 11: Question 22 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

22.A golf ball driven down a fairway is observed to travel through the air with a trajector
Which following force(s) is(are) acting on the golf ball during its entire flight?



1. the force of gravity
2. the force of the "hit"
3. the force of air resistance

(A) 1 only
(B) 1 and 2
(C) 1, 2 and 3
(D) 1 and 3
(E) 2 and 3

In the discussion below it can be seen that the crucial distinction that is made by those who wish to persuade the others that there is no 'force of the "hit"' during the flight is its prototypical property of being localised to a particular time, so that it is not a continuously present influence like gravity.

Question 22 (see Box 11)

Group 97 5-1

FS3: It's D ...xxx Because after hitting the force gone	*Prototypical property of a 'hit' – localised in time.*
MS1: ... It's not a continuing force ***FS1:*** But in fact ***FS3:*** Like just now we all we all conclude already after after moving xxx the force isn't there any more	
When it's during the kick there is a force	*Similarly with a 'kick' – again this is localised in time.*
FS4: so the xxx ***MS1:*** ... So we want to go with D mmh ***FS1:*** Agree with ***MS1:*** D ***FS1:*** D ***MS1:*** Okay	

In contrast to the above discussion which was soon resolved, the long transcript (almost two minutes) of the discussion by Group 97 5-3 shows that they eventually reached an impasse as the two sides were unable to see their way to the solution. As in the discussion by group 97 5-1 some noted that the prototypical hit is localised to a brief instant of time. Opposed to this, others argue that the force of the hit has the prototypical property of an "effect", that is to say, that the projectile continues to move.

Group 97 5-3

MS2: Okay twenty-two	
A golf ball is driven down a fairway	*Describing the problem situation*
MS3: ...	
xxx Tiger Wood I guess	*The situation activates real world associations*
MS2: How about here ***MS3:*** xxx I don't understand ***MS2:*** xxx ***MS4:*** ... What one are you doing	*The group organises its discussion.*
MS2: ... force of gravity ***MS3:*** the force of the hit ***MS1:*** force of the hit	*Reading from the options*

one two three right	*The first to speak finds it reasonable that all three forces are present: gravity, air resistance and the original hit.*
MS3: one two three yah	*A second person agrees*
C yah	
MS4: This here	*Others hesitate*
MS3: C	
MS2: ...	
Yeah	
MS3: ...	
C yah	
MS2: ...	*After a pause, the key difficulty is raised:*
force of the hit still got?	*Can the hit still be present: the prototypical 'hit' occurs at an instant of time.*
MS4: force of	*Others are now less certain*
MS1: Oh	
MS2: xxx	
MS1: xxx	
MS2: I say one and two	*Counter proposal: only gravity and the hit. Possibly, thinking of all the problems where one is told to ignore air resistance.*
MS1: Yeah xxx	*Seconded*
MS3: xxx	
MS1: ...	
one and three are there	*Countering one suggestion with another. Possibly, thinking that in this problem there is no instruction to ignore air-resistance.*
MS2: ...	
yeah	
MS1: How about two	*Focusing again on force 2, 'the force of the hit'*
Two	
If you	
If there's no force of hit	
it will xxx go forward	*The indistinct word (i.e.* xxx *in the transcription) is probably something like 'not' from the context.*
MS4: Yeah	
it will not go forward	*Agrees with the viewpoint that the 'hit' is necessary for the motion – but the statements made so far are ambiguous between two interpretations:* the force of the hit is a necessary **precondition** for the ball to start moving (true) the force of the hit is a **continuously present** influence on the ball's motion (false)

yeah	
MS1: It will not have horizontal air velocity	*The phrase "horizontal air velocity" doesn't actually mean anything. The speaker is attempting to bolster his viewpoint with the prestige attached to the use of technical vocabulary, probably intending to echo the sentiment that "it will not go forward".*
MS4: The gravity will only pull down	*This statement is, of course, true. However, it only makes sense here in this argument if the direction of the force is relevant to the direction of the velocity, which it is not.*
MS3: Also it stop the hit from going further	*This is true on the interpretation that because of gravity the ball will follow a parabolic path which will eventually bring it into contact with the ground and therefore gravity will – indirectly – slow it to a stop. It is false on the interpretation that gravity will exert a force that will directly slow the ball's horizontal flight.*
MS2: ...	
Cannot be	*Rejecting the previous statement, presumably on the second interpretation.*
MS3: ...	
push of the hit	*Focusing attention on the key point again.*
MS2: ...	
but it will only act during the hit	*Focusing attention on the prototypical property of occurring at an instant.*
but after that it won't act at all really	*Restating the same point, but this time as a negative, as* **not** *occurring over an interval of time.*
...	
MS3: But the xxx	*Starts to reply*
MS2: That's why That's why	*Interrupts to point out the logical consequence of Newton's first law: that it is the lack of force which ensures a constant velocity.*
you can talk about	
...	
constant	*Emphasising the word by pausing before and after.*
...	
what you call it speed lah	*Downplaying the importance of the technical vocabulary by emphasising his lack of attention to it.*
MS1: horizontal speed	
MS3: There's a hit that's why there's a horizontal speed	*The same ambiguity between* ***hit*** *as "precondition" and "continuously present"*

MS1: Yeah but xxx to air resistance it will slow down	*The prototypical effect of air resistance.*
MS3: And due to gravity it will come down	*The prototypical effect of gravity.*
MS2: But ... Yeah but	
there's no force	*Here,* MS2 *means 'no force of the hit' not zero net force.*
that's why it will slow down	*I.e. in the absence of a continuous forward force, the air resistance has this effect.*
you know	
MS1: The hit upward force	*This speaker starts to reply,*
ahh	*...but can't formulate his response*
MS3: But	*His ally completes the response, and for the first time states unambiguously*
but the force is still there	*that the force is continuously present.*
MS2: ...	
I think	*Now that this statement is out in the open, that is to say the focus of the group's attention, it can be directly addressed*
during the flight yah ... if the force is still there then	
MS3: xxx ***MS2:*** it will go faster ... yeah it going faster all the time right	*MS2 proposes a hypothetical argument – **If** there is a constantly present force, then the logical consequence of* Newton's *second law follows: a force forward will result in acceleration forward.*
... xxx	
MS3: xxx resistance	MS3 *seems to be calling upon 'air resistance' as a fix-all.*
which xxx	
MS2: No no no no no	*But* MS2 *will have none of this.*
MS1: ... but its during the entire flight	MS1 *now focuses attention on the key phrase of the question...*
so no hit I think	*and now seems to agree with* MS2's *main contention, that there is no continuous presence of the hit.* *However, at this point in the discussion,* MS1 *is in fact arguing against* MS2's *hypothetical argument by denying the antecedent.*
MS4: Yeah	*Another agrees.*
MS2: No no ...	*But now MS2 seems to disagree with what has been his own contention.*

Because ... if really acting upwards all the time	*An interpretation which makes sense of what he is saying, however, is that he is rejecting their right to make the above counter argument.*
that means it should have faster and faster and faster	*The others have been arguing that there is a force of the bit acting throughout the flight. Now, when a logical consequence of this view is shown to be absurd, they temporarily abandon it. He argues that they can't have it both ways. If, counterfactually, there was a continuously present 'force of the bit' then there would be continuous acceleration.*
faster faster faster faster faster	*As he speaks, his speech becomes more and more rapid, echoing his meaning. Also, one gets the impression, reflecting his impatience.*
MS3: C's xxx ***MS1:*** Yah man? ***MS4:*** Ah yah Think so ***MS1:*** Okay	*The others try to end the discussion by making a decision. The first proposal is that all three forces, including the force of the bit are present throughout the entire flight.*
MS2: D	MS2 *is certain he is right, and won't accept the first proposal. He suggests the answer which does not include the force of the bit.*
MS3: Is it all right if I put D yah ***MS4:*** No ***MS2:*** xxx	MS3 *consults the others.*
xxx ***MS3:*** Anything wrong ***MS2:*** I think we can't discuss this also right We can't discuss this also ...	*There are irreconcilable differences which they cannot address within the time constraints of the group discussion.*
No time to discuss also xxx answer	*The discussion ends inconclusively.*

Single force vs total force

The use of the word force in normal technical usage is polysemous between the meanings *total force* and *a particular force*. For example: in the statement "Force is mass time acceleration" one is referring to the total (i.e. net or resultant) of all the individual forces acting. When one refers to "the force *mg*," one refers to a single force, the force of gravity. Similarly when one says that there is no force on an object, this could be taken to mean that there is no force at all acting, or that there are multiple forces acting, but that these cancel out.

Students need to sort out this ambiguity. The members of Group 95 2-4 can be seen doing so in the discussion of Question 12 of the FCI (see Box 12). The discussion was very animated with a great deal of overlapping speech, and sometimes three speaking at once.

Box 12: Question 12 of the *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992).

> 12. A book is at rest on a table top. Which of the following forces is(are) acting on the book?
>
> 1. A downward force due to gravity.
> 2. The upward force by the table.
> 3. A net downward force due to air pressure.
> 4. A net upward force due to air pressure.
>
> (A) 1 only
> (B) 1 and 2
> (C) 1, 2, and 3
> (D) 1, 2, and 4
> (E) none of these, since the book is at rest there are no forces acting on it.

Group 95 2-4

MS1: Why do y...Why do you say E?
MS2: Each force is act cancelling each other.. so there is.. the answer is E.
MS1: Cancel out no but the force is still acting
FS1: yes
MS1:on on the book..
FS1: Yeah but you're not talking about the net force
MS1: It is cancelling but.. but the force, the force is there...the force is still there.. although it is cancelling but it is still there
MS2: It is acting yet it is cancelling ..so the force is .. none.
FS1: No it's acting, so it's not talking about net force.
MS1: Yes. .. It's not talking about net force. It's talking about the force. If there is no force, there is no good.

While this question was very successful in prompting lively discussion, it was, in some ways, poorly designed. Getting the correct answer not only depends on distinguishing when 'force' does or does not mean 'net force', but also confounds this with the unrelated issue of whether students know of the presence of a small upthrust due to the fact that air pressure at the bottom of an object is greater than it is at the top – a topic which is rarely included in physics curricula at this level. In the most recent version of the FCI this question has been rewritten to eliminate this second issue.

Furthermore, as well as the ambiguity of force there is also an ambiguity in the word net: it can be used with different scopes. When one speaks of the net force one usually refers to the total of all forces acting, but in this question the "net … force due to air pressure" restricts the scope to adding together just the forces due to air pressure.

Students can interpret the scope wrongly: for example in **Group 97 5-3** a student says: "See here it says net you know, it means total, there's no net downward force."

In this instance the student seems to infer correctly from Newton's first law that there is no net force, hence no net downward force, using the wider scope of net (i.e meaning the total of all forces acting). However, the only net downward force referred to in question 12 of the *FCI* is the "net downward force due to air pressure," and Newton's first law is irrelevant here: it would not rule out such a force if there were one or more forces to counterbalance it.

The argument is therefore invalid (although the conclusion in this case is, by luck, correct: the net force of air pressure is upward, not downward).

It has been shown above that far from being deployed as a simple mathematically defined technical term, the use of the concept of force is in fact complex. While students are able to work with prototypical instances of forces quite comfortably, they are not aware of just where the boundaries between instances and non-instances of forces occur. This, it has been argued accounts for much of the difficulty that students have with using Newtonian Physics.

Interpretation of *mass* in Newton's second law

That there are subtleties involved in the concept of *mass* and especially in the distinction between *mass, weight* and *density* (Galili, 1994) and in more advanced students between *inertial mass* and *gravitational mass*, (Domenech et al., 1993) has been established in the research literature. However, in the transcripts of the discussions examined here there was no mention of density, and all 31 uses of the terms *mass* and 23 of *weight* were appropriate. This was not really surprising as Galili's work was done with junior high school students, while the students here were senior high school. Presumably either they had resolved the issues which junior students had not done, or possibly, the instrument used here did not probe these difficulties. Certainly, the issue of *inertial* vs. *gravitational mass* did not arise at all in the test.

Interpretation of *acceleration* in Newton's second law

It was shown in Chapter Four of this thesis that change of direction was considered a highly dubious example of acceleration, with some students rating it as not an example of acceleration at all. For these students motion at a constant speed would be considered unaccelerated motion. Thus Newton's second law would simply not be activated when these students encountered problems involving circular motion.

Conclusions

Overview of the thesis

An extensive review of the science education research literature in Chapter Two showed that there has been a great deal of inconsistency in the terminology used to describe researchers' interpretations of students' ideas. When this was compared with work in linguistics it was seen that there were systematic differences in the data collected by science education researchers which were not being attended to: concepts were not being differentiated theoretically from larger scale mental entities.

Furthermore it was seen that the prototype structure of the concepts in physics had not been investigated at all.

Chapter Three examined the four main types of research methods to be used to investigate the development of concepts in mechanics. The reasons for their choice were given and their limitations were discussed.

In Chapter Four it was shown, in spite of the terminological and theoretical disputes about what concepts were, that when authors of science education research papers gave examples of *concepts* they were remarkably consistent with each other. A *concept*, it was shown for these authors, was prototypically something that could represent the meaning of a single word and this usage was found to be consistent with the usage of the word in general, non-technical contexts.

Chapter Five established that the technical concept of *acceleration* has a prototype structure both in its general usage and in its technical usage within physics (the evidence shows this is certainly the case for the students investigated here, and statistical analysis, discussed below, strongly suggests it is the case also for the wider population represented by this sample of students). The structure of the concept *acceleration* was examined, firstly, in terms of its lifeworld sense. The concept's definition in terms of the fundamental concepts of the Natural Semantic Metalanguage was determined by a careful examination of a large corpus of usage. Secondly, the technical concept of *acceleration* amongst a group of students was examined in terms of the concept's reference: its prototypical exemplars, how well other exemplars rate, and the extent to which the concept's boundaries vary for different people, and in different contexts. In addition, it was shown that the prototype for the technical usage was clearly the same as it was for the general usage: something getting faster. The difference was that the boundaries of the technical concept stretched further away from the prototype. The sequence of conceptual development was argued to consist of two extensions of the boundary of the lifeworld concept: the first extended the boundaries of the concept to include things slowing down, and the second extended it to include things changing direction.

While Chapter Five dealt with a technical concept where the boundaries of the lifeworld concept had to be extended, Chapter Six presented four studies which converged in showing that the denotation of the concept of *Newtonian force* could be described as being a subset of lifeworld forces. It is based on the lifeworld concept of *force*, but is restricted to those forces which are like the prototypical actions of pushing and hitting.

This was reflected in the need to have a longer NSM definition of Newtonian force than was needed to define the verb 'to force': more restrictions in the definition reduce the number of cases which fall under it.

Chapter Seven extended the notion of prototype to deal with prototypical Newton's third law situations: it was shown that Newton's third law was activated and used far more easily when the two interacting objects were at rest and of equal mass. It was argued that learning to use Newton's third law correctly required that the class of situations that were seen to be relevant needed to be extended to include unequal masses and masses that were moving.

Chapters Four to Seven considered together showed that the process of conceptual development could be modelled by a process which involved establishing a prototype to serve as the foundation of the concept, and then adjusting and fine-tuning the border between instances and non-instances of the concept. There was no need to invoke the conceptual change theory of Posner et al. (1982): there was no evidence of rational consideration of intelligibility, plausibility, or fruitfulness. The variations between concepts that these students had were the results of largely unconscious decisions they made about how close to the prototype instances at boundaries of their concepts were.

It was, however, the evidence presented in Chapter Eight that brought out most clearly the role of activation and suppression for these students in the building of mental models of problems and their solutions. Three sets of evidence pointed toward this conclusion. Firstly, 'probability format' tests showed that multiple conceptions were simultaneously available to students. Secondly transcripts of discussions showed that in developing their solutions students were opportunists, adapting any method which led to an answer, whether personal experiences, intuitive rules or even, sometimes, Newtonian mechanics. Thirdly, the transcripts also showed many fragmentary contributions to the arguments that could best be interpreted as attempts to highlight, or activate, particular ideas in their listeners.

Rather than the rational process of weighing up criteria as implied by the theory of Posner et al., for these students the joint process of building an understanding was a messy affair involving interruptions and occasionally feelings running high.

Chapter Nine emphasises the messiness of the process of coming to understand Newton's second law. There are a large number of ways in which the formula $F = ma$ can be understood. These are due, partly, to the different concepts of *force* and *acceleration* which students were shown to have in Chapters Four and Five.

Conceptual change

Taken together, the chapters provide evidence for two main types of conceptual change.

In the first, changes in the understanding of what is meant by a word are reflected in extensions or restrictions of the denotation of its concept, but the prototype itself remains unchanged, only the boundaries moving. This is exemplified by the technical concept of *acceleration* which differs from the lifeworld concept only in being more inclusive.

In the second type, one establishes a new prototype: the technical definition of *force* is of this sort. The lifeworld meanings of 'force' are quite varied, but many have in common the idea of "imposing one's will." One small subset of these involved in pushing things or hitting them becomes the prototype for the development of a new technical sense of *force*.

Scope of conclusions

To what extent can these results, and the other conclusions of this thesis, be generalised?

Physics concepts exhibit prototype effects

Consider, first, the conclusion that mechanics concepts have a prototype structure. In Chapter Two the vast literature on prototypicality developed from the pioneering work of Rosch was reviewed. The concepts that have been investigated in this literature demonstrate the features of prototype structure, and this is true cross-culturally. What sorts of concepts have been investigated? There are concrete concepts both for non-living objects (including specific, like *cup*, and general, like *furniture*) and living ones (like *tree*, *fruit*, *animal*, or *bird*,). There are abstract concepts (for qualities like *colour*, or actions like *lie*). (Coleman & Kay, 1981; Labov, 1973; Lakoff, 1987; Lakoff & Johnson, 1980; Rosch, 1978; Rosch & Mervis, 1975; Rosch et al., 1976) Prototypicality effects have been demonstrated for concepts which one would not expect to observe graded membership: *odd numbers*, *females*, *plane geometric figures* (Armstrong et al., 1983). Prototypicality effects have been observed for concepts which were introduced anew to experimental subjects: *things to take on a picnic*, *foods not to eat on a diet.* (Barsolou, 1985). There is ongoing controversy over whether prototype effects can be observed in categorisation by other species such as pigeons (eg. White, Alsop, & Williams, 1993)!

Meanwhile, in science education research, prototype effects in concepts have not been observed because, as shown in the literature review, nobody has looked for them (I speak here of the English language literature, and such other language literature as has been incorporated, in translation, in the journals, monographs, and electronic literature and citation indices to which I have had access.) Nevertheless, the first time I looked for evidence of prototype structure of a concept (*acceleration*), in a very rough-and-ready way, asking students to raise their hands in class, the effect was so strong it passed the admittedly rough-and-ready "intra-ocular" statistical test: it hit one between the eye-balls! More formal tests, detailed in Chapter Five, replicated the result in three different classes. Furthermore, in spite of the fact that the students had very varied first-language backgrounds, as reported in Chapter Five, statistical testing did not support the hypothesis that their backgrounds made any difference to the results.

The conclusion that is being proposed is that physics concepts have a prototype structure. For the purpose of being specific, let us examine the assertion that "a car getting faster" is a more prototypical example of *acceleration* (i.e. is rated more highly) than "a car getting slower." How far can this result (in Table 6 of Chapter Five) be generalised? More precisely, the claim that "getting faster" is more prototypical than "getting slower" is equivalent to the assertion that the first will be rated more highly than the second by any sufficiently large sample from the population. Can this be claimed from the evidence presented in this thesis?

To begin with we note that the standard error in the estimate of the mean is $\frac{\sigma}{\sqrt{N-1}}$ where σ is the standard deviation and $N\text{-}1$ is the number of degrees of freedom (where N is the number of data points). One notes, in passing, that the total number of the population being sampled does not enter into the equation, only the size of the sample. Consulting Table 6 of Chapter Five, one finds that for a car getting faster the mean is 9.2, $\sigma = 2.1$ and $N = 46$, so that the error is $\frac{2.1}{\sqrt{46-1}} = 0.31$. Similarly, for a car getting slower, the mean is 7.3 and the error 0.40. The results are clearly different (the difference is somewhere between 0.51 and 3.07 at the 99% confidence level), and the difference is in the direction expected. For the population represented by these students, and this (given that their varied language backgrounds made no difference) can plausibly be taken to be final year high school physics students, it is clear that the rating for a "car getting faster" will be greater than for "a car getting slower." On the assumption that the students investigated are a random sample, then the number of students in the sample is indeed sufficient to support the general statement that the concept of *acceleration* exhibits prototype effects. Similar arguments apply for the interpretation of the data for the concept *force.*

Given the overwhelming evidence for prototype effects in all concepts, noted above, it could hardly be otherwise.

Students first learn prototypical, then less prototypical instances of concepts

Consider now the claim that students learn the concepts involved in Mechanics first with prototypical examples, and that only later do they learn to include cases which are less prototypical. One might ask whether this can be asserted without qualification based only upon the evidence presented in this thesis.

Clearly if the evidence of the thesis was restricted to the students who filled out the various questionnaires, undertook tests, and discussed their answers then there would be insufficient basis for this claim.

This however, would be to miss the point of the examination of the linguistic corpora. Take the concept of *acceleration*, for example. A very large sample of language use, carefully designed to be representative, was examined for the uses of the word 'acceleration' and showed that it is used by the general public only in the sense of "getting faster." (See Chapter Five.)

This, of course, is the prototype for *acceleration* amongst these students, as argued above and in Chapter Five. It was also shown in that chapter that these students accepted that other situations including getting slower, and for many, changing direction were also cases of *acceleration*, although not such good examples.

It follows, therefore, on the assumption that these students (indeed any students) were members of the general public before they became physics students, that they first learnt the prototypical sense and later added the less prototypical senses. (As stated in Chapter Five, however, the evidence adduced in this thesis is insufficient to determine the order in which the less prototypical senses are developed. The order in which they are shown as being added in Figure 10 in Chapter Five is conjectural.)

'Acceleration' is a comparatively infrequent word, and unsurprisingly was not observed in the transcripts in the CHILDES corpus (MacWhinney, 2000a, 2000b). However, because it is a comparatively common word, three stages in the development of the concept of force can be glimpsed in evidence independent of that obtained from the students investigated in this study (which constituted a fourth stage).

In the first stage, before the concept itself begins to develop, one can see that the prototypical cases for *force* (which were argued in Chapter Six, on the basis of the investigation of the students in this study, to be pushing, hitting, and kicking) are shown by the CHILDES corpus to all be in use by children before the age of two (see Figure 21 in Chapter Six).

A second stage occurs after early childhood, but before physics instruction when people encounter the lifeworld senses of the word 'force.' The analysis of the use of the word 'force' by the general public, using the British National Corpus (BNC, 1994), while showing the complexity of its multiple senses, also showed that in general usage the sense most nearly aligned with the physicist's sense of *force* covered cases of pushing and various kinds of impact.

A third stage occurs as students begin their study of mechanics. Hart's investigation of the denotation of the concept of force for students who were just beginning their study of physics uncovered evidence of the inclusion of more general lifeworld notions (especially in terms of agency and goal-directedness) not yet fully differentiated from the correct Newtonian notion: "… my students still included a whole variety of interactions in their conception of force that were not part of the concept as I intended them to understand it." (Hart, 2002, p. 237)

The students in the investigations reported in this thesis were at a fourth stage, having completed their study of high school mechanics. Thus, although this thesis does not contain a longitudinal study, one can reasonably make plausible inferences about the development of the concept of *force*. As was the case with *acceleration*, the prototypical senses are learnt before students enter the classroom, and less prototypical senses (in the case of *force*, these might be friction, or the normal reaction force, for example) learnt later, if at all (Twigger et al., 1994).

Can one state, then, that prototypical cases of physics concepts are always learnt first? Clearly this has only been established to be the case for *acceleration* and *force*. For other concepts like *velocity*, *momentum* or *electric current* the question remains open, but one can hypothesise that they too will follow this pattern.

Prototypical applications of Newton's Laws

Similar patterns of development were argued to occur at the level of conceptions, such as Newton's laws, which are built using concepts. It was shown that much of the research into Newton's Third Law could be explained by assuming that certain situations will be prototypical applications. For students to recognise that other situations also fall under the same law requires that they recognise that the situation falls into the same category as the prototype. In particular, it was shown by a re-analysis of these earlier investigations that their results were consistent with a process of first learning a prototypical situation and then expanding it to include less prototypical examples.

It should be noted that this way of looking at the problem of generalisation of laws, brings the problem of transfer of learning into the same theoretical domain as prototypes.

Activation and suppression

Evidence from the analysis of transcripts of discussions in Chapters Eight and Nine suggests that, when larger scale conceptions are built up, the concepts which are used in this process are not consciously selected. Rather, just as in normal speech comprehension, one unconsciously selects the most highly activated concept to incorporate, and this then suppresses competing concepts. Contexts which spread activation to Newtonian concepts will therefore result in different conceptions than those which do not: in the latter case, whichever other concepts are available and most highly activated will be selected and these will suppress the Newtonian concepts.

The ACT-R theory of cognition argues that not only the contents of declarative memory, such as concepts are activated and suppressed in this way, but also the contents of procedural memory, the production rules which code our methods for dealing with problems.

This therefore offers a principled account of the often noted inconsistency of student use of scientific conceptions and misconceptions in problem solving.

Possible further investigations

There are four main areas of investigation which lead on naturally from the work reported in this thesis. The first would look to find further evidence of the prototype structure of the concepts of *force* and *acceleration*. The second would seek to investigate how the prototype structure of the concept of *force*, as described in this thesis, is relevant to linguistics, in so far as the results described here are inconsistent with some work done in this discipline. The third would look to find evidence of prototype structure in other concepts related to physics and science education. The fourth would aim to discover if it is helpful to follow a teaching sequence where one aims first to teach the prototypical examples before moving on to less prototypical cases. These will now be discussed.

Other measures of prototype effects

Firstly, in so far as the concepts of acceleration and force have been argued here to have a prototype structure, it would be worthwhile checking whether they exhibited other prototype effects. Reaction time measurements are predicted to be less, and accuracy greater, for judging statements as instances or non-instances of a concept when they are more prototypical (Lakoff, 1987, p. 41). Thus "A car is getting faster" should be recognised as an instance of *acceleration* more quickly than "A car turns a corner." Similar results should be obtained in judging instances of *force*.

In designing an experiment to determine whether this was the case or not, it would be necessary to control for confounding variables which have been shown to be relevant to reaction times (Gernsbacher, 1990; Gernsbacher, 1994). One of these is the time taken to comprehend the sentences being tested. Some relevant factors are:

- The grammatical form of the sentence (for example, "A car is driven faster" is passive, while "A car turns a corner" is active).

- The relative frequency of the words in common usage (for example "A yacht tacks" uses less frequent vocabulary than "A boat slows down").

- The length of the sentence (for example "A four-wheel drive pulling a caravan slows down for a red light" vs. "A car swerves").

- The 'expectedness' of content (consider for example the contrast between "A ball rolls downhill" and the relatively incongruous "A snail speeds up").

A second confounding variable is the pre-existing level of activation of the concepts being tested. For instance, one expects increasingly faster reaction times and more accurate responses if one asks a series of questions related to the same concept: the first question activates the concept, and it is more easily and quickly accessed for the second and subsequent questions.

In order to determine experimentally the difference in reaction times in a task where situations to be classified as exemplifying the concept of *acceleration*, one would therefore need to measure comprehension times for the sentences describing these situations (using methods drawn from Gernsbacher, 1990). It would also be necessary to use mixed orders of presentation, and to include other classification tasks so that the gain in speed of response due to the priming of the concept of *acceleration* would not swamp the speed difference being sought.

While the predicted results can be quite simply stated, and the techniques mentioned above address issues that are relevant to measuring reaction times, nevertheless, such an investigation would require a great deal of preparation and a significant investment of time and effort on the part of the researcher.

Applying the knowledge of the prototype structure of force

Apart from applications in physics pedagogy, which will be discussed in a separate section, there is an implication for the linguistic study of causative constructions.

As discussed in Chapter Two, Talmy (1976; 1985) sought to explain a number of cross-linguistic regularities with his widely noted theory describing causation in terms of force dynamics. According to this theory, causative (and related) constructions in languages are modelled on the regularities observed in the action of physical forces on objects. However, the evidence presented in this thesis shows that, on the contrary, the idea of physical force is based upon the lifeworld conception of force, and that it is developed after lifeworld senses that involve "imposing one's will" (a sense which necessarily assumes the existence of *causes*, *agents* and *goals*: roles explained by Talmy in terms of physical force). Thus, Talmy's theory seems to have things the wrong way around: the evidence suggests that the sense of physical force is based upon the lifeworld conception of force, not vice versa. Given that the cross-linguistic regularities nevertheless exist, an alternative explanation for them needs to be sought.

It was conjectured in the discussion of Talmy's theory in Chapter Two that the cognitive processes of *activation* and *suppression* would serve better as models for these phenomena than physical forces and resistances. In order to argue for or against this hypothesis one would need, at the minimum, to re-examine in its light the linguistic data adduced by Talmy and by others who have utilised his theory as a framework for their own work. While the implementation of an investigation into this alternative explanation for the data unearthed by Talmy would require a significant amount of effort and expertise, the conclusions of this thesis are in conflict with Talmy's theory as it stands, and so provide justification for such an investigation.

For which concepts would a prototype analysis be useful?

The third area of investigation which would follow as a natural extension of the work in this thesis is the examination of the structure of other concepts of physics. Possible candidates for this type of investigation include *velocity, speed, distance, displacement, mass, weight, electricity, voltage, current, power, heat, temperature, cold, radiation,* and *radioactivity.* Other concepts of importance to different disciplines in science education – for example *respiration*, or *animal* from biology, *intelligence* from psychology, or *acid* from chemistry – could also be examined.

But, one might ask, which of these is worth investigating? Students have been learning physics and the other sciences for many years without any attention being given to the structure of the concepts used. Why should we be concerned with identifying concept prototypes, identifying the boundaries beyond which the students do not accept that an exemplar belongs to a concept, and seeking the reasons which students accept for extending or contracting these boundaries?

There are three reasons. The first is to do with ease of teaching and learning: prototypical cases form the foundation of the concept and are the easiest to grasp. If we know what is prototypical then we may be able to structure our introductory courses more effectively.

The second is to do with communication: we need to know whether our students understand what we say. For example, when we say "an orbiting satellite is accelerating" do our students realise that we are talking about the changing direction of the satellite's velocity, or do they think we mean the prototypical sense, which is that it is getting faster? When we speak of animals do our students realise that we include jellyfish and spiders, or do they think we speak only of prototypical animals like cows or dogs? Without a knowledge of the prototype structure of these concepts we do not know.

The third is to do with logical inference: in order to reason effectively it is necessary to know just where the boundaries of the concepts' denotations occur. To clarify this third point it is necessary to examine the role of concept boundaries further.

As discussed in Chapter Six, by utilising concept-prototypes people are able to get the gist of a meaning and work with it usefully without needing to know exactly where the concept boundaries lie. Given this natural ability, how important is a knowledge of the exact boundaries of a concept? For example, for a high school student completing practical work or end-of-chapter exercises on Newton's Laws, is it important, one might ask, to know where one stops using the concept *inertial mass*, and starts using the concept *gravitational mass*?

Consider an analogy with the concept *cup*. On the one hand, when we want to drink a coffee at home it is sufficient for our purposes that we know how to find a cup in the kitchen, and we simply look for something like the prototypical cup. It is not necessary to know exactly the width to height ratio at which we cease to use the concept *cup* and start to use *mug*, or at the other extreme, *bowl*. The exact boundaries between these are not important for many purposes, and, in any case, different people set the boundaries at different points (Labov, 1973; Wierzbicka, 1985). On the other hand, when ordering coffee at a shop, one may be asked "Cup or mug?" In this context, one does need to know the difference: the mug (as a consequence of its shape) holds a greater quantity and hence will cost more. This analogy suggests that where recognition is required, knowledge of concept boundaries is less important, while in cases where a logical consequence is to be drawn, the boundaries are more important.

An examination of a graphical representation of logical deduction can help make this point clear. It is common to represent the logical relationship of material implication (if p then q) by means of a Venn diagram, as shown on the right of Figure 39. In the case of the classical representation of concepts it is clear that the denotation of concept p (say, *people*) is wholly contained within that of concept q (say, *mortal*). Hence if we know that Socrates belongs to the set of *people*, then we can deduce that Socrates is *mortal*. However, on the left hand side, which represents prototypical cases using black and less prototypical cases using shades of grey, the concept denotation is not sharply defined and it is not clear whether or not all cases of p are also subsumed by q.

It is plain that the use of deductive logic relies upon concepts being defined – and one notes that the etymological derivation of 'define' refers literally to setting boundaries.

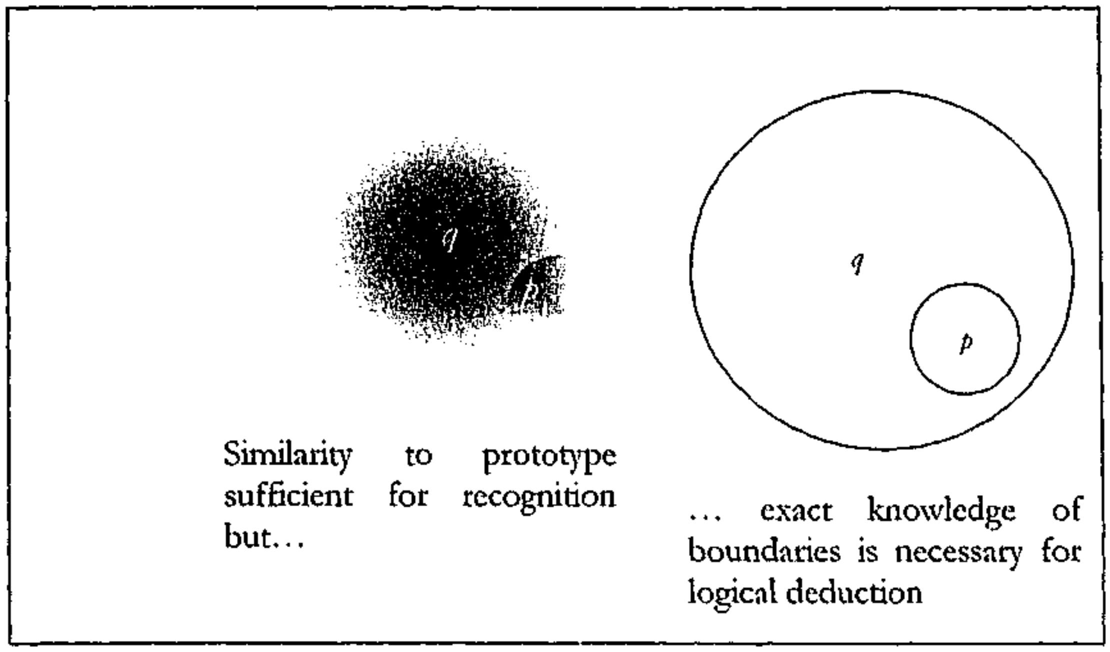


Figure 39: Concepts with graded exemplars (represented by degree of shade) on left, and classical representation of concepts (defined by boundaries) on right.

Given the evidence of prototype effects in all concepts, one might question how it is possible for people to use logical deduction at all. The way out of this impasse is to realise that when we reason deductively we do use clear cut boundaries: technical definitions build upon the foundation of prototypes, but have boundaries that are determined by agreement within the relevant community. This can be illustrated as shown in Figure 40. While prototype effects are still present (and are represented as before by the shading), it is nevertheless clear that all instances of p are subsumed within the denotation of q. For instance, while the number 3 may be a prototypical member and 437 may not be prototypical, they are nevertheless both clearly within the set of odd numbers (Armstrong et al., 1983).

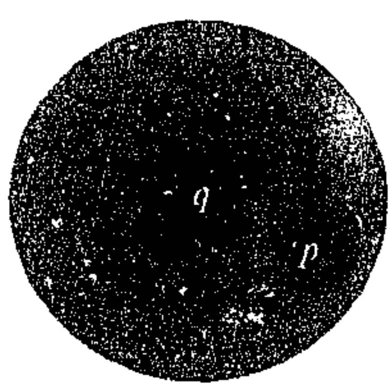


Figure 40: Technical concepts illustrating both degree of 'goodness' (by shading) and clear boundaries.

To summarise, then, and to return to the question with which this section began, it is worthwhile investigating the prototype structure of terms when they play a role in logical inferences in the theory. To take the example of the concepts investigated in this thesis, *acceleration* and *force* play a key role in Newtonian physics because of their role in inferences: if something is accelerating, we know it must have a net force acting on it. From the fact that planets orbit the sun Newton deduced the presence of the gravitational force, for instance. But the acceleration of these planets is not the prototypical acceleration of getting faster: if we wish our students to understand this inference (and be able to make similar inferences), then we must ensure that they have a full understanding of the concept of *acceleration*, not one limited to the prototypical cases.

Where the concepts play no role in logical inference, as is the case for *gravitational mass* vs. *inertial mass*, which are discussed in some introductory mechanics texts, there may be no need to introduce the distinction: one can work and communicate with students because they can grasp the prototypical cases. Knowledge of a concept does not begin and end, as in the classical model of concepts, with students knowing the technical definition with its clear-cut boundaries: the prototype is sufficient. Of course, when students reach the stage of studying relativity and the distinction between these concepts does play a role in logical inferences the technical definition becomes necessary, but not until then.

Summarising this section: the investigation of concepts which play an important role in the logical structure of a discipline using the methods utilised in this thesis (investigation of dictionaries, analysis of data from linguistic corpora, analyses of answers to open response items asking for examples, and of questionnaires asking for ratings) is clearly justified on three grounds. Firstly, it could serve as an aid in the structuring and sequencing of courses. Secondly, it would alert teachers to possible areas of miscommunication with students. Thirdly, it would aid teachers in getting students to understand the logical structure of the discipline being sought.

Effectiveness of teaching using prototypes

The fourth extension of the work reported in this thesis, would be to replicate the pilot study reported in Chapter Six using a quasi-experimental design (given that intact classes will almost certainly be the only ones available).

Four classes would be required. A pretest would be administered to two of these. This would be followed by teaching one class in the manner described in Chapter Six, while the other followed a traditional sequence. This would be followed by a post-test. An additional two classes would be needed to enable testing for the possible sensitising effect of the pretest on the results. These two groups, which would not receive the pretest, would otherwise parallel the first two classes.

A comparison, using two way ANOVA, of the groups which did and did not receive the pretest, and did or did not receive the teaching using prototypes would allow the contribution of the pretest, and interactions between the pre-testing and teaching using prototypes to be examined. If this showed that the interaction between presence or absence of pre-testing with teaching method was not significant, then a comparison of the first two class results using ANCOVA, with the results of the pre-test as a covariate would then provide some firm evidence as to the usefulness or otherwise of this approach.

Threats to the validity of this experimental design would include the possibility of a "Hawthorne effect," where students who were the subjects of the "experimental treatment" – teaching using prototypes – might do better, not because of the teaching, but because it was different from what they expected and therefore motivating, or because they felt more valued and therefore put in a greater effort to learn. If this appeared to be a significant worry then the experimental design could be extended. A third pair of groups would be introduced which used a different innovation – for example concept-mapping or computer based multimedia training. The motivational effect, if it existed, would apply to both experimental treatments, and could, again, be tested for using ANOVA.

A more subtle threat to validity would be in differential self-selection of the teachers. If teachers who volunteered to teach using a sequence based on prototypes were for some reason (because they were more adventurous, say) better at teaching than the control group, a difference between the two sets would not necessarily be due to the difference in teaching methods, but rather due to the teaching talent. To avoid this, one would need a sufficiently large group of teachers willing to be assigned at random to teaching using one method or the other. This would, of course, increase the scale of the study and the effort and preparation required to conduct it considerably, but would correspondingly increase one's confidence in the validity of the results.

Before such an experiment could be carried out, an extensive in-service would need to be made available to interested teachers. This training would need to show how to run group discussions which highlight conceptual aspects of physics. The notion of prototypicality would need to be clarified and examples given. Strategies for implementing this approach would also need to be developed.

Implications for teaching

Given that the consequences of the existence of prototype effects in concepts have not previously been subject to investigation, there is a question as to the usefulness of this approach in different areas of teaching. What sort of pedagogical problems can be solved by this approach? Three applications were mentioned above, and will be discussed in turn: increasing communicative efficacy, course design, and teaching the logical structure of a discipline.

Communicating and the use of concepts

Investigation can show whether students know the extent of the coverage of a concept. Class discussions and small group discussions of non-prototypical cases – and of non-cases similar to those covered by the concept – can clarify where the scientific community sets the boundaries.

Where the concept is an explicit one, for example *acceleration*, it is clear that this will aid in communication. In terms of hermeneutics, one might express this by saying that the horizons of the concepts of the student and the teacher are merged.

Where the concept is not explicitly named – for example the idea of situations which are subject to Newton's third law – an investigation of its structure also aids in communication. Specifically, it enables one to communicate it by describing it in terms that are significant to students. Taking the example of Newton's third law, for example, the evidence presented in Study One of Chapter Six, suggests that one needs to tell students that it applies not only to situations where the two masses interacting are not moving, and equal in mass (the prototype) but also applies when the two objects are not equal in mass, and when they are moving with a constant velocity, and when they are accelerating, and that it applies whichever of the two objects is the agent. Exercises covering each of these instances need to be worked through. (In terms of the theory of activation and suppression, simply stating that Newton's Law applies to all cases will not activate in the minds of the students anything beyond the prototypical case, and in order for a link to be set up both ideas must be in working memory, that is to say attended to, or focussed upon.) By explicitly addressing each of these cases one helps student build the ability to recognise the full extent of relevant situations. This in turn enables students to transfer their learning about the prototypical cases to other cases: in other words, to generalise their knowledge.

Sequencing of instruction

If the results obtained in the pilot study of teaching Newton's third law prove to be robust, it may be useful to base teaching strategies on the teaching of the prototype, and then moving out to less prototypical instances so as to delineate the boundaries of the concept. It seems reasonable to believe that one can build confidence by staying close to the prototypical situations in early teaching, and then in later work raise interest by looking at boundary instances. By the later stages each student will have an opinion as to what does and what does not fall under a concept, but each person's opinion will differ from others. This is a useful precondition for lively class discussions and group work.

A further implication for course design is related to the pacing of teaching: for the sake of specificity consider the concept of *acceleration*.

If key concepts, such as *acceleration*, are thought to be adequately introduced by means of a formula and a few substitution exercises there is no justification for spending much class time in doing this. One introduces the concept once and for all and then proceeds immediately to the applications: in the case of *acceleration* this would involve procedures involving the formulae for uniformly accelerated motion, looking at the slopes of tangents on velocity time graphs, and so on.

However, when the structure of the concept is understood, the need for a planned sequence of activities becomes evident. The amount of time devoted to elucidating the concept itself becomes greater.

Again we can use the concept of acceleration to illustrate this. A possible instructional sequence might begin with the idea of acceleration as getting faster. It would then, as is usual, look at procedures for calculating the value of the measure of acceleration, and look at its use in graph interpretation, and in the equations of accelerated motion, applying these amongst other things to cases of objects which are dropped or thrown downward and accelerate due to gravity. However, after this introduction, one would explicitly address the expansion of student understanding of the concept by directing attention to cases where things get slower. Revisiting the concept now sets the stage for clarifying the use of positive or negative signs in equations as shorthand for indicating directions. This is an appropriate point to look at the direction of the acceleration of an object thrown vertically upward into the air: the constant downward acceleration due to gravity can be seen to manifest itself as a decrease in speed when the velocity is directed upwards, and as an increase in speed on the way down. The constancy of gravity can serve as one motivation for including cases of *getting slower* together with cases of *getting faster* within the denotation of the concept *acceleration.* As well as introducing the idea, one would set exercises using the equations of uniform acceleration and graphical analysis related to getting slower. At this stage, in many syllabi, one might move on from kinematics to other topics, such as dynamics in one dimensional motion. However, when one returned to the topic of acceleration in two dimensional motion (for example projectiles, circular motion, and the motion of objects in orbits), rather than simply assuming that students know what *acceleration* is, one would explicitly direct attention to how and why changing direction is included within the denotation of *acceleration.* Firstly, the constant direction of gravity in projectile motion can be contrasted with the changes in direction and speed of the projectile: the same value of acceleration manifests itself as either a change of speed, a change of direction, or a combination of these. Secondly, the way that all three cases (faster, slower, changing direction) can be measured using the single definition as the vector rate of change of velocity, i.e. $a = \frac{\Delta \mathbf{v}}{t}$ can be used to justify grouping them together. Thirdly, the relativity of reference frames can be used: for a person in a train moving at a constant speed an object that is dropped appears to simply move down in a straight line and get faster, but for an observer viewing from a fixed position in the frame of reference of the ground it is seen to be following a curved projectile path and therefore changing direction. After this has been done the use of the measure of acceleration in formulae and so forth would continue.

The example of a sequence for the teaching of *acceleration* has been looked at in some detail: however, the point being made is a general one. It is an implication of the conclusions of this thesis that more time be devoted to explicating concepts than is currently allowed for.

Logical structure of mechanics

It has often been noted that the logical structure of the disciplines of the sciences is not sufficiently addressed in teaching (eg. Brincones & Otero, 1994; Gardner, 1975). This pedagogical problem is, for instance, identified in an overview of high school physics texts (Entwistle et al., 1999, p. 299):

> A more general flaw is the way that physics is presented as a series of disconnected laws and rules. The books dutifully present Newton's three laws, then proceed to ignore them in developing new concepts. For example, few of the books effectively relate conservation of momentum to the Third Law. Some of them derive centripetal acceleration but fail to connect it to satellite motion with the Second Law.

However, before students can understand the linkage between Newton's second law and satellite motion they must recognise that circling at a constant speed is an example of acceleration. They must recognise that "centripetal acceleration" is simply acceleration that happens to be directed toward a particular point: too often students think of it as a new concept, linked only in its vocabulary to (prototypical) acceleration. That this is not an unreasonable supposition on the part of students can be seen by the fact that the terms 'angular momentum' and (linear) 'momentum' do in fact represent different concepts in mechanics.

It is argued, therefore, that directing attention toward the understanding of concepts can aid in the promotion of coherency in the understanding of a scientific discipline.

Limitations

It is not claimed that the approaches advocated above are in any sense a panacea for all problems students have in understanding mechanics.

The structure of concepts is only one aspect of the discipline of mechanics. Students also must learn many procedures. Some of these are explicitly taught: how to interpret graphs in terms of slopes, areas and values; how to substitute into and solve various equations. Others must be learnt by practising them: how to convert a problem stated in words into a problem which can be solved using the techniques that have been explicitly taught; the skills of manipulation of experimental apparatus. There is also a certain degree of rote learning: knowing that the value of g at the Earth's surface is approximately 9.8 ms^{-2}, or that the speed of light is approximately 3.00×10^{8} ms^{-1}, knowing the names and prefixes for SI units, and knowing and being able to recall various formulae all form part of the subject. This all involves effort, and students may become discouraged by initial failures: those who feel that achievement is outside their control will not persist as well as those who view their achievements as falling within their locus of control (Marton & Booth, 1997). There is also widespread acknowledgement of the disproportionately low enrolment of women in the study of physics. The problems associated with mastery of explicit procedures, with skill acquisition, rote learning, self-image, and gender all fall outside the topic addressed in this thesis and all have repercussions which affect the learning of mechanics.

Summary

Earlier theories of conceptual development have been criticised in this thesis, and an alternative theory involving the prototype structures of concepts, and the activation and suppression of these concepts has been presented in detail as a more satisfactory account of the evidence uncovered by the investigations which were undertaken. The methods of investigation used to uncover this evidence were shown to be justified, and the results were argued to be generalisable. The theory presented has been shown to have implications for wider fields within science education than the study of mechanics, which was the focus of attention for this thesis. The theory puts forward specific and falsifiable predictions as required by Popper's (1969) account of scientific theories. It also has implications for the teaching of mechanics.

There is much scope for further work evaluating the application to the classroom of the ideas developed in this thesis.

Appendix A: *Force Concept Inventory* (Hestenes et al., 1992)

Force Concept Inventory

1. Two metal balls are the same size, but one weighs twice as much as the other. The balls are dropped from the top of a two story building at the same instant of time. The time it takes the balls to reach the ground below will be:

(A) about half as long for the heavier ball.

(B) about half as long for the lighter ball.

(C) about the same time for both balls.

(D) considerably less for the heavier ball, but not necessarily half as long.

(E) considerably less for the lighter ball, but not necessarily half as long.

2. Imagine a head-on collision between a large truck and a small compact car. During the collision:

(A) the truck exerts a greater amount of force on the car than the car exerts on the truck.

(B) the car exerts a greater amount of force on the truck than the truck exerts on the car.

(C) neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.

(D) the truck exerts a force on the car but the car doesn't exert a force on the truck.

(E) the truck exerts the same amount of force on the car as the car exerts on the truck.

3. Two steel balls, one of which weighs twice as much as the other, roll off of a horizontal table with the same speeds. In this situation:

(A) both balls impact the floor at approximately the same horizontal distance from the base of the table.

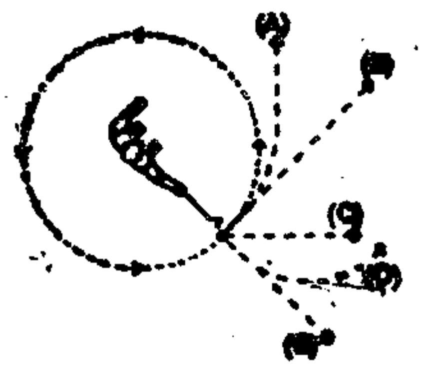
(B) the heavier ball impacts the floor at about half the horizontal distance from the base of the table than does the lighter.

(C) the lighter ball impacts the floor at about half the horizontal distance from the base of the table than does the heavier.

(D) the heavier ball hits considerably closer to the base of the table than the lighter, but not necessarily half the horizontal distance.

(E) the lighter ball hits considerably closer to the base of the table than the heavier, but not necessarily half the horizontal distance.

4. A heavy ball is attached to a string and swung in a circular path in a horizontal plane as illustrated in the diagram to the right. At the point indicated in the diagram, the string suddenly breaks at the ball. If these events were observed from directly above, indicate the path ofthe ball after the string breaks.



5. A boy throws a steel ball straight up. **Disregarding any effects of air resistance,** the force(s) acting on the ball until it returns to the ground is (are):

(A) its weight vertically downward along with a steadily decreasing upward force.

(B) a steadily decreasing upward force from the moment it leaves the hand until it reaches its highest point beyond which there is a steadily increasing downward force of gravity as the object gets closer to the earth.

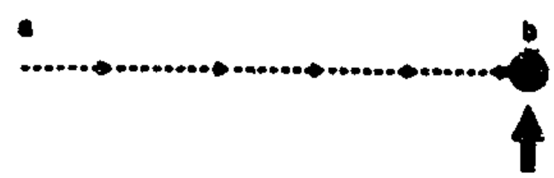
(C) a constant downward force of gravity along with an upward force that steadily decreases until the ball reaches its highest point, after which there is only the constant downward force of gravity.

(D) a constant downward force of gravity only.

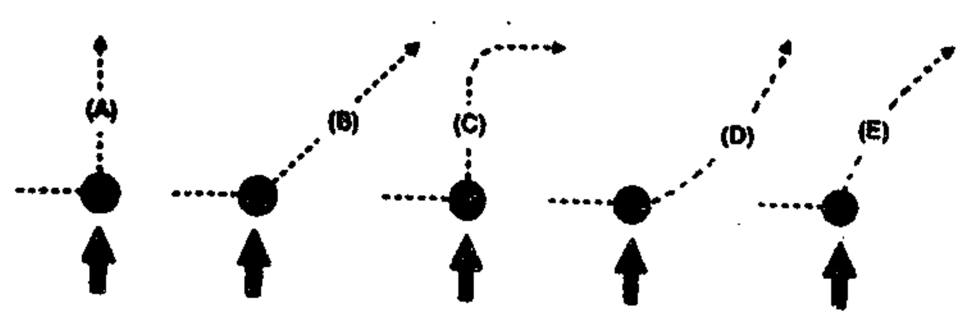
(E) none of the above, the ball falls back down to the earth simply because that is its natural action.

* Use the statement and diagam below to answer the next four questions:

* The diagram depicts a hockey puck sliding, with a constant velocity, from point "a" to point "b" along a frictionless horizontal surface. When the puck reaches point "b" , it receives an instantaneous horizontal "kick" in the direction of the heavy print arrow



6. Along which of the paths below will the hockey puck move after receiving the "kick"?



7. The speed of the puck just after it receives the "kick"?

(A) Equal to the speed "v_o" it had before it received the "kick"

(B) Equal to the speed "V" it acquires from the ""kick", and independent of the speed No".

(C) Equal to the arithmetic sum of speeds "v_o" and "V".

(D) Smaller than either of speeds "v_o" and "V".

(E) Greater than either of speeds "v_o" and "V," but smaller than the arithmetic sum of these two speeds.

8. Along the frictionless path you have chosen, how does the speed of the puck vary after receiving the "kick"?

(A) No change.

(B) Continuously increasing.

(C) Continuously decreasing.

(D) Increasing for a while, and decreasing thereafter.

(E) Constant for a while, and decreasing thereafter.

9. The main forces acting, after the "kick", on the puck along the path you have chosen are:

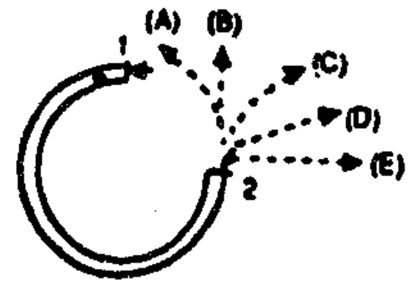
(A) the downward force due to gravity and the effect of air pressure.

(B) the downward force of gravity and the horizontal force of momentum in the direction of motion.

(C) the downward force of gravity, the upward force exerted by the table, and a horizontal force acting on the puck **in the direction of motion.**

(D) the downward force of gravity and an upward force exerted on the puck by the table.

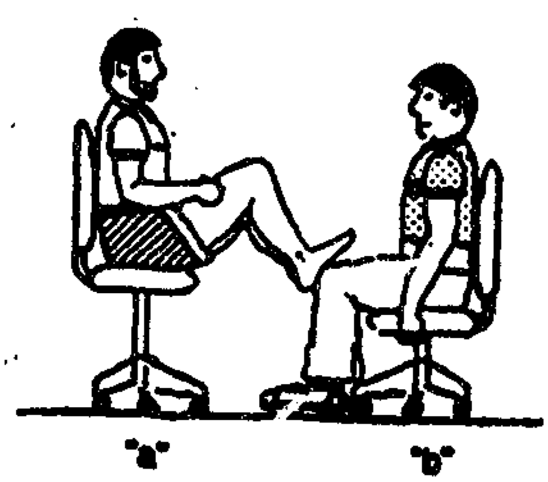
(E) gravity does not exert a force on the puck, it falls because of the intrinsic tendency of the object to fall to its natural place.



10. The accompanying diagram depicts a semicircular channel that has been securely attached, in a **horizontal plane,** to a table top. A ball enters the channel at "1" and exits at "2". Which of the path representations would most nearly correspond to the path of the ball as it exits the channel at "2"and rolls across the table top.

Two students, student "a" who has a mass of 95 kg and student "b" who has a mass of 77 kg sit in identical office chairs facing each other. Student "a" places his bare feet on student b's" knees, as shown below. Student "a" then suddenly pushes outward with his feet, causing both chairs to move.

11.. In this situation,



(A) neither student exerts a force on the other.

(B) student "a" exerts a force on "b", but "b" doesn't exert any force on "a".

(C) each student exerts a force on the other but "b" exerts the larger force.

(D) each student exerts a force on the other but "a" exerts the large force.

(E) each student exerts the same amount of force on the other.

12. A book is at rest on a table top. Which of the following force(s) is(are) acting on the book?

1. A downward force due to gravity.

2. The upward force by the table.

3. A net downward force due to air pressure.

4. A net upward force due to air pressure.

(A) 1 only

(B) 1 and 2

(C) 1, 2 and 3

(D) 1,2, and 4

(E) none of these, since the book is at rest there are no forces acting on it.

Refer to the following statement and diagram while answering the next two questions.

A large truck breaks down out on the road and receives a push back into town by a small compact car.

13. While the car, still pushing the truck, is **speeding up to get up** to cruising speed:

(A) the amount of force of the car pushing against the truck is equal to that of the truck pushing back against the car.

(B) the amount of force of the car pushing against the truck is less than that of the truck pushing back against the car.

(C) the amount of force of the car pushing against the truck is greater than that of the truck pushing against the car.

(D) the car's engine is running so it applies a force as it pushes against the truck but the trucks engine is not running so it can't push back against the car, the truck is pushed forward simply because it is in the way of the car.

(E) neither the car nor the truck exert any force on the other, the truck is pushed forward simply because it is in the way of the car.

14. After the person in the car, while pushing the truck, reaches the cruising speed at which he/she wishes to continue to travel at a constant speed:

(A) the amount of force of the car pushing against the truck is equal to that of the truck pushing back against the car.

(B) the amount of force of the car pushing against the truck is less than that of the truck pushing back against the car.

(C) the amount of force of the car pushing against the truck is greater than that of the truck pushing against the car.

(D) the car's engine is running so it applies a force as it pushes against the truck but the trucks engine is not running so it can't push back against the car, the truck is pushed forward simply because it is in the way of the car.

(E) neither the car nor the truck exert any force on the other, the truck is pushed forward simply because it is in the way of the car.

15. When a rubber ball dropped from rest bounces off the floor, its direction of motion is reversed because:

(A) energy of the ball is conserved.

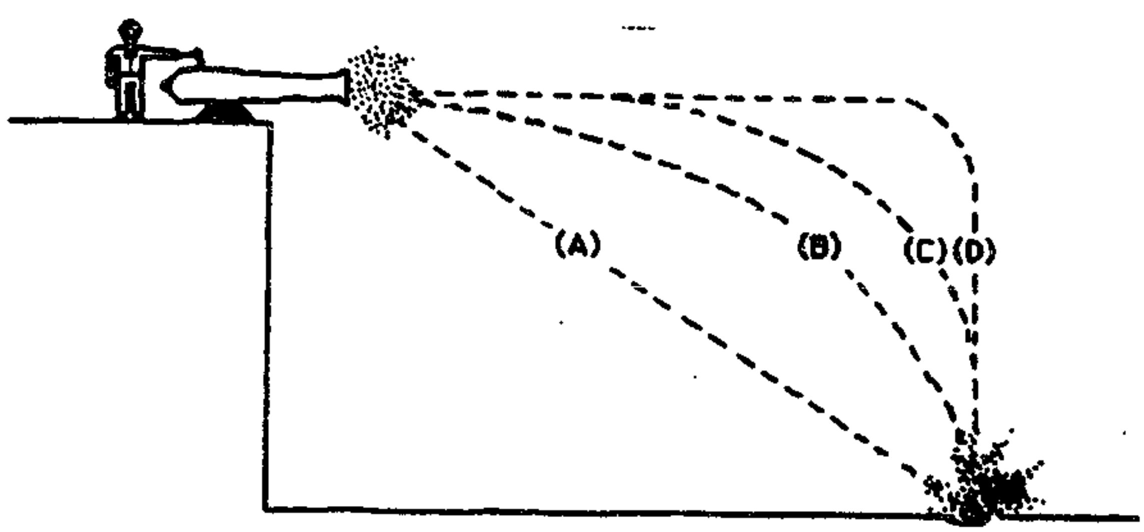
(B) momentum of the ball is conserved.

(C) the floor exerts a force on the ball that stops its fall and then drives it upward.

(D) the floor is in the way and the ball has to keep moving.

(E) none of the above.

16. Which of the paths in the diagram to the right best represents the path of the cannon ball?



17. A stone falling from the roof of a single story building to the surface of the earth:

(A) reaches its maximum speed quite soon after release and then falls at a constant speed thereafter.

(B) speeds up as it falls, primarily because the closer the stone gets to the earth, the stronger the gravitational attraction.

(C) speeds up because of the constant gravitational force acting on it.

(D) falls because of the intrinsic tendency of all objects to fall toward the earth.

(E) falls because of a combination of the force of gravity and the air pressure pushing it downward.

When responding to the following question, assume that any frictional forces due to air resistance are so small that they can be ignored.

18. An elevator, as illustrated, is being lifted up an elevator shaft by a steel cable. When the elevator is moving up the shaft at a constant velocity:

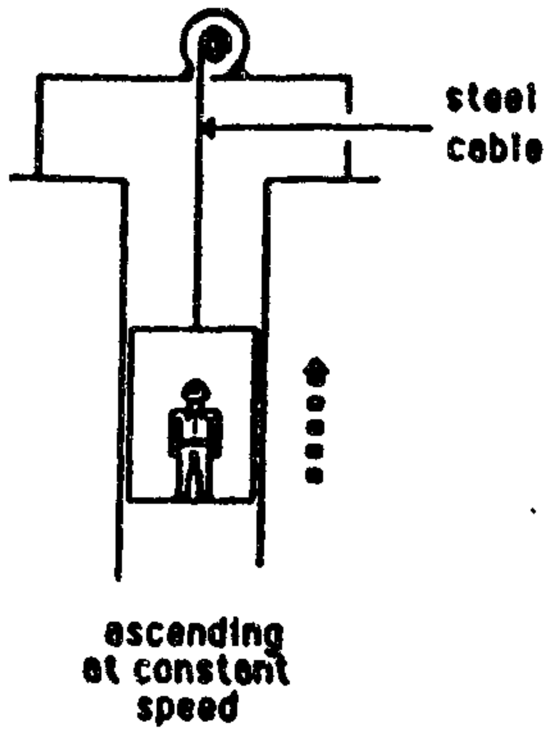
(A) the upward force on the elevator by the cable is greater than the downward force of gravity.

(B) the amount of upward force on the elevator by the cables equal to that of the downward force of gravity.

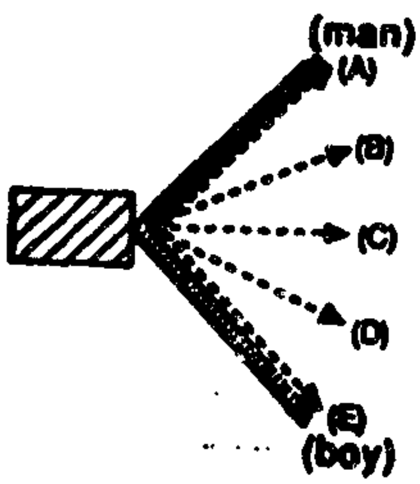
(C) the upward force on the elevator by the cable is less than the downward force of gravity.

(D) it goes up because the cable is being shortened, not because of the force being exerted on the elevator by the cable.

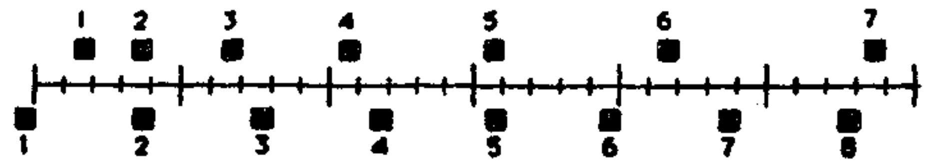
(E) the upward force on the elevator by the cable is greater than the downward force due to the combined effects of air pressure and the force of gravity.



19. Two people, a large man and a boy, are pulling as hard as they canon two ropes attached to a crate as illustrated in the diagram to the right. Which of the indicated paths (A-E) would most likely correspond to the path of the crate as they pull it along?



The positions of two blocks at successive 0.20 second time intervals are represented by the numbered squares in the diagram below. The blocks are moving toward the right.



20. Do the blocks ever have the same speed?

(A) No.

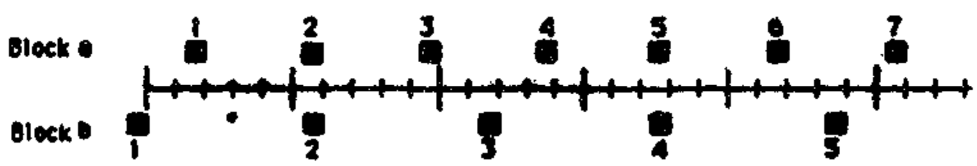
(B) Yes, at instant 2.

(C) Yes, at instant 5.

(D) Yes, at instant 2 and 5.

(E) Yes, at some time during interval 3 to 4.

The positions of two blocks at successive equal time intervals are represented by numbered squares in the diagram below. The blocks are moving toward the right.



21. The acceleration of the blocks are related as follows:

(A) acceleration of "a" > acceleration of "b"

(B) acceleration of "a" = acceleration "b" > 0

(C) acceleration of "b" > acceleration "a"

(D) acceleration of "a" = acceleration "b"= 0.

(E) not enough information to answer.

22. A golf ball driven down a fairway is observed to travel through the air with a trajectory (flight path) similar to that in the depiction below.

Which following force(s) is(are) acting on the golf ball during its entire flight?

1. the force of gravity
2. the force of the "hit"
3. the force of air resistance

(A) 1 only

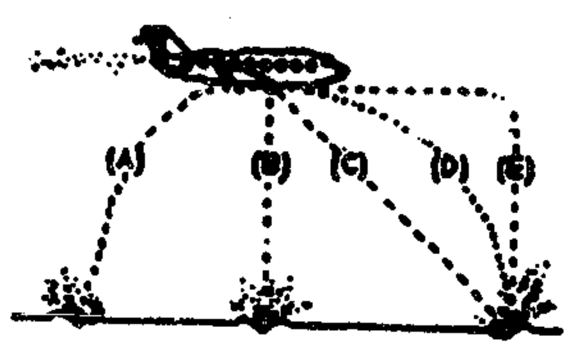
(B) 1 and 2

(C) 1, 2 and 3

(D) 1 and 3

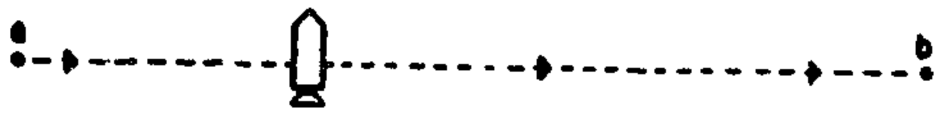
(E) 2 and 3

23. A bowling ball accidentally falls out of the cargo bay of an airliner as it flies along in a horizontal direction. As seen from the ground, which path would the bowling ball most closely follow after leaving the airplane?

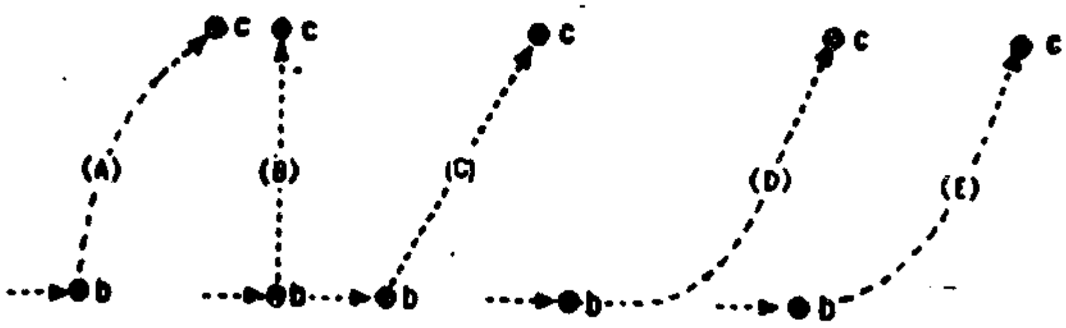


When answering the next four questions, refer to the following statement and diagram.

A rocket, drifting sideways in outer space from position "a" to position "b". Is subject to no outside forces. At "b", the rocket's engine starts to produce a constant thrust at right angles to the "ab". The engine turns off again as the rocket reaches some point "c".



24. Which path below best represents the path of the rocket between "b" and "c"?



25. As the rocket moves from "b" to "c", its speed is

(A) constant.

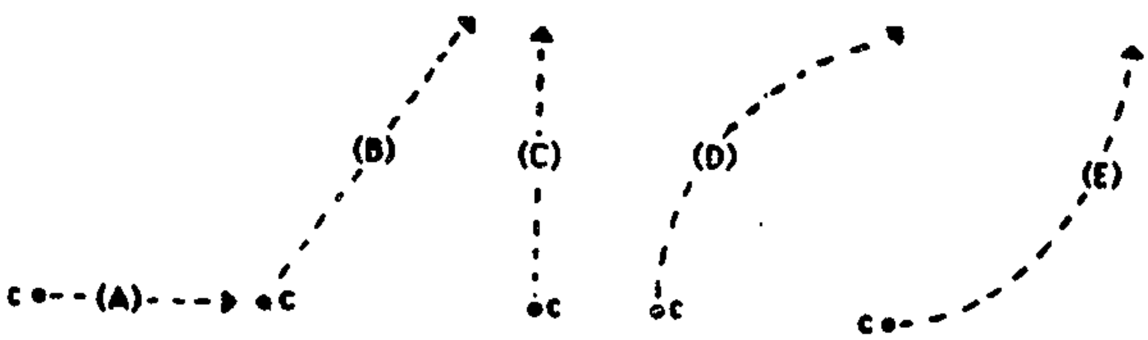
(B) continuously increasing.

(C) continuously decreasing.

(D) increasing for a while and constant thereafter.

(E) constant for a while and decreasing thereafter.

26. At "c" the rocket's engine is turned off. Which of the paths below will the rocket follow beyond "c"?



27. Beyond "c", the speed of the rocket is:

(A) constant.

(B) continuously increasing.

(C) continuously decreasing.

(D) increasing for a while and constant thereafter.

(E) constant for a while and decreasing thereafter.

28. A large box is being pushed across the floor at a constant speed of 4.0 m s^{-1} . What can you conclude about the forces acting on the box

(A) If the force applied to the box is doubled, the constant speed of the box will increase to 8.0 m s^{-1}.

(B) The amount of force applied to move the box at a constant speed must be more than its weight.

(C) The amount of force applied to move the box at a constant speed must be equal to the amount of the frictional forces that resist its motion.

(D) The amount of force applied to move the box at a constant speed must be more than the amount of the frictional forces that resist its motion.

(E) There is a force being applied to the box to make it move but the external forces such as friction are not "real" forces, they just resist motion.

29. If the force being applied to the box in the preceding problem is suddenly discontinued, the box will:

(A) stop immediately.

(B) continue at a constant speed for **a very short period of time and** then slow to a stop.

(C) immediately start slowing to a stop.

(D) continue at a constant velocity.

(E) increase its speed for a very short period of time, then start slowing to a stop.

Appendix B: Instruction sheet for FCI Probability Format

FRIDAY'S TEST

Two types of answer sheets will be distributed for the test on Friday. One type will be the usual multiple choice answer sheet. Those who get this sort just enter their answers as usual.

For example for question 30, if you decide that the answer is C then you enter C in the answer sheet as usual.

30.	C

The other type asks you to distribute the 10 points between the answers you think might be the correct answer.

FOR EXAMPLE ON QUESTION 30:

Jenny decides that C is definitely correct. She puts 10 in that box:

30	A	B	C	D	E
			10		

Kim cannot make up his mind which of C or E is correct, so he puts 5 points in each:

30	A	B	C	D	E
			5		5

Mei Mei believes that B is probably the correct answer, but she thinks maybe D is, just possibly. She might distribute the points like this:

30	A	B	C	D	E
		7		3	

Bob has no idea at all, so he distributes points like this:

30	A	B	C	D	E
	2	2	2	2	2

If the correct answer is A, then Bob gets two points, and Jenny, Kim and Mei Mei get no points.

If the correct answer is D, then Bob gets two points, Mei Mei gets three points and the others get no points.

Appendix C: Year 10 Class test on Newton's laws

Name:	
ID number:	

Year 10 Physics Test: Newton's Laws
Version 1.

1 (a)	
(b)	
(c)	
(d)	
(e)	
(f)	

2	
3	
4	
5	
6	
7	
8	

9	

10	
11	
12	

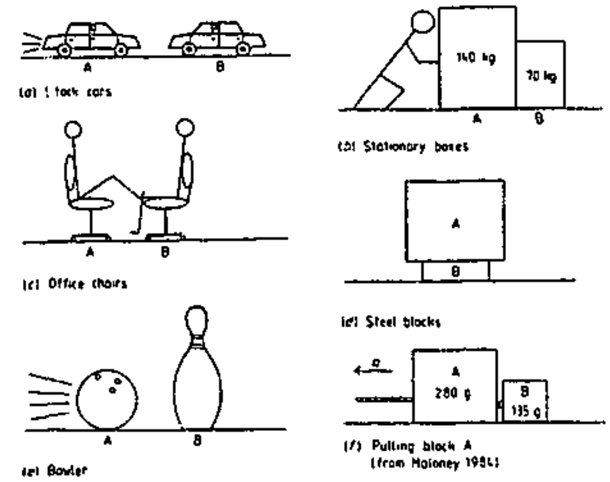
13	
14	

15	

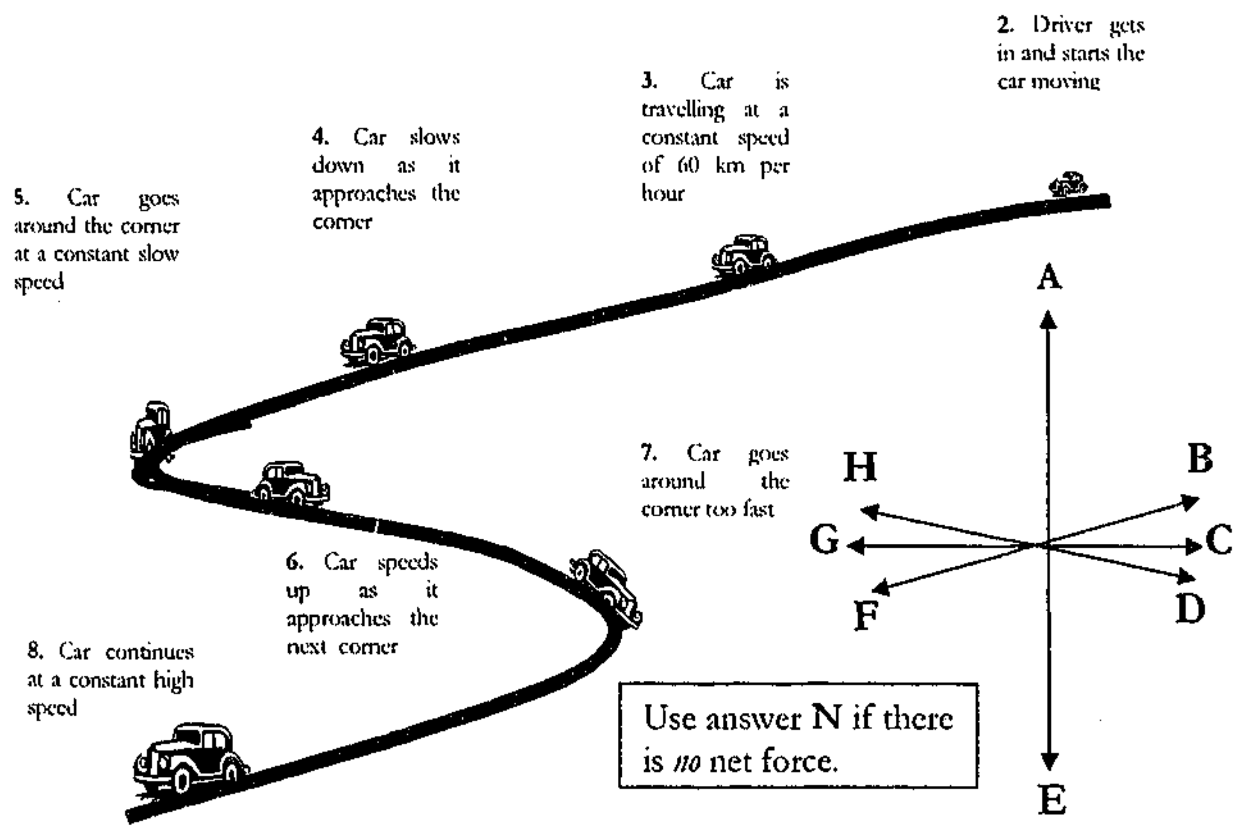
16	N
17	N
18	$m\ s^{-2}$
19	N
20	$m\ s^{-2}$

Year 10 Physics Test: Newton's Laws

1. For each of the situations (a) to (f) below, state which is the bigger force
   - A. the force exerted by A on B.
   - B. the force exerted by B on A.
   - C. they are the same.



A car is driven along a road as shown in the diagram below. For each position (2 to 8) of the car, use the following key to state the direction of the net force acting on the car at that point.



9. A large truck collides head-on with a small compact car. During the collision:

A. the truck exerts a greater amount of force on the car than the car exerts on the truck.

B. the car exerts a greater amount of force on the truck than the truck exerts on the car.

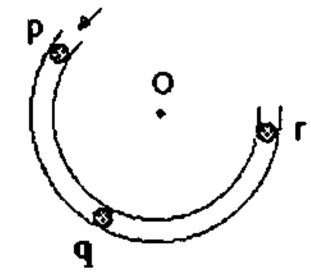
C. neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.

D. the truck exerts a force on the car but the car does not exert a force on the truck.

E. the truck exerts the same amount of force on the car as the car exerts on the truck.

USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT TWO QUESTIONS (10 and 11).

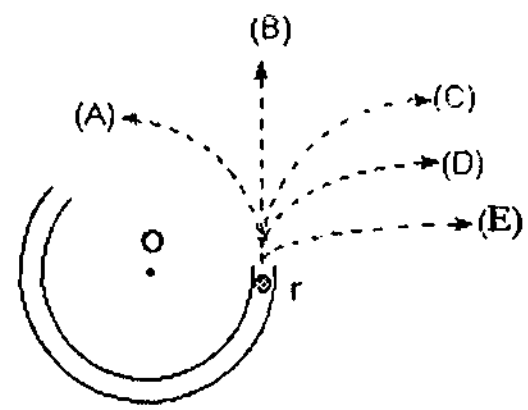
The accompanying figure shows a frictionless channel in the shape of a segment of a circle with center at "O". The channel has been anchored to a frictionless horizontal table top. You are looking down at the table. Forces exerted by the air are negligible. A ball is shot at high speed into the channel at "p" and exits at "r."



10. Consider the following different forces:
    1. A downward force of gravity.
    2. A force exerted by the channel pointing from q to O.
    3. A force in the direction of motion.
    4. A force pointing from O to q.

Which of the above forces is (are) acting on the ball when it is within the frictionless channel at position "q"?

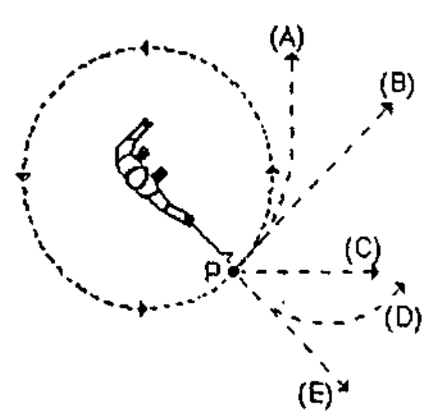
A. 1 only.
B. 1 and 2.
C. 1 and :1.
D. 1, 2, and 3.
E. 1, 3, and 4.



11. Which path in the figure at right would the ball most closely follow after it exits the channel at "r" and moves across the frictionless table top?

12. A steel ball is attached to a string and is swung in a circular path in a horizontal plane as illustrated in the accompanying figure.
At the point P indicated in the figure, the string suddenly breaks near the ball.

If these events are observed from directly above as in the figure, which path would the ball most closely follow after the string breaks?



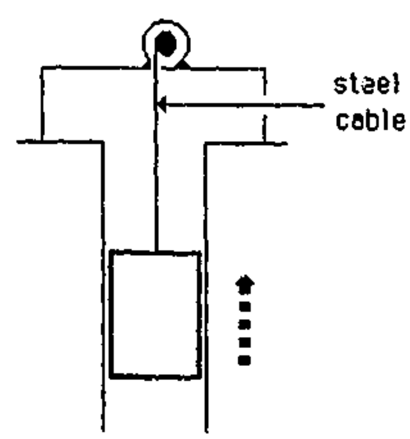
USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT TWO QUESTIONS (13 and 14).

A large truck breaks down out on the road and receives a push back into town by a small compact car as shown in the figure below.

13. While the car, still pushing the truck, is speeding up to get up to cruising speed:
    (A) the amount of force with which the car pushes on the truck is equal to that with which the truck pushes back on the car.
    (B) the amount of force with which the car pushes on the truck is smaller than that with which the truck pushes back on the car.
    (C) the amount of force with which the car pushes on the truck is greater than that with which the truck pushes back on the car.
    (D) the car's engine is running so the car pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.
    (E) neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

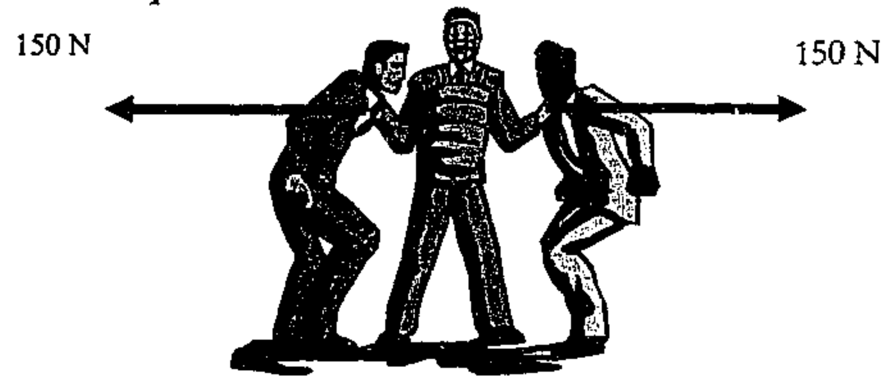
14. After the car reaches the constant cruising speed at which its driver wishes to push the truck:
    (A) the amount of force with which the car pushes on the truck is equal to that with which the truck pushes back on the car.
    (B) the amount of force with which the car pushes on the truck is smaller than that with which the truck pushes back on the car.
    (C) the amount of force with which the car pushes on the truck is greater than that with which the truck pushes back on the car.
    (D) the car's engine is running so the car pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.
    (E) neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

15. 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).



Elevator going up at constant speed

16. The umpire is pushing on the two men with forces as shown. But what is the net force on the ***umpire***?

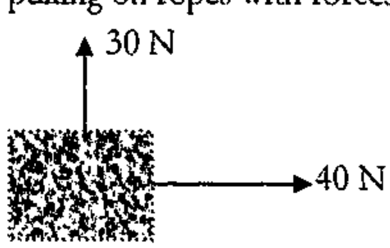


17. The child's foot is pushing back ***on the ground*** with a force of 60 N, but the wheels of the scooter experience a frictional force of 20 N. What is the net force ***on the child and scooter***?

Force of the child's foot pushing back ***on the ground*** = 60 N

18. The mass of a ball is 0.100 kg, and it is kicked with a force of 80 N. What is the ball's acceleration?

19. What is the size of the net force on this box, which is being pulled along a frictionless surface by two people pulling on ropes with forces of 30 N and 40 N?



20. The box in question 19 has a mass of 5.0 kg. How much does it accelerate?

Appendix D: Practical Exercise on Projectile Motion and Uncertainty of Measurement

Date: ....

Name: ....Roll No: .... Group: ....

Partners: ....

....

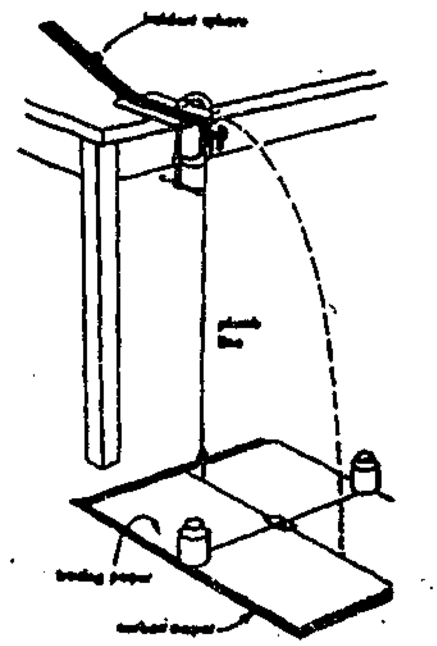
....

EXPERIMENT:

HORIZONTAL PROJECTILES

Aim:
To compare the speeds of three balls after they have rolled down a ramp, and use this to see if it is valid to ignore friction and rotational kinetic energy in calculations.

Apparatus:
Set up the equipment as shown.

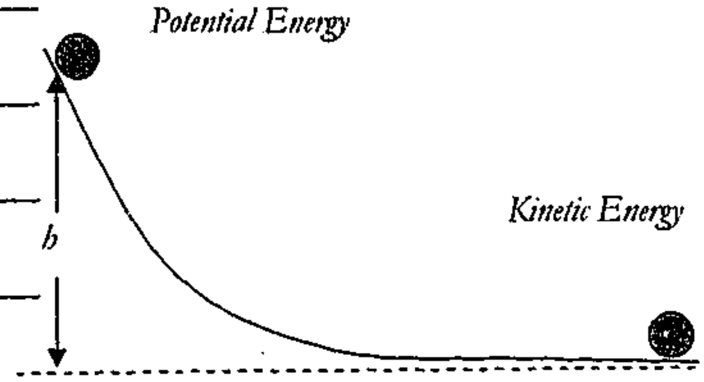


Method:
Measure the height h of the ramp above the bench, the height H of the bench top above the floor, and the distance d that the ball has moved as a projectile. From d and H calculate how fast each ball was moving after rolling down the ramp.
Roll balls of different mass and radius down the slope and measure the horizontal distance they travel.

Results:

	1	2	3	Mean (cm)
Ramp (h)				
Bench (H)				

1. From the height of the curved ramp calculate the speed of the ball as it leaves the ramp, assuming that there is negligible loss of energy due to friction. Also ignore the amount of energy required to set the ball into rotation.



2. Does your answer in question 1 depend upon the **mass** of the ball? Why (or why not)?

3. From the height above the floor of the point where the ball leaves the ramp calculate the time it takes the ball to reach the floor. Neglect air-resistance in your calculations.

t = ______ s

4. Does your calculation of the time it takes the ball to reach the floor (question 3) depend on the **mass** of the ball? Why (or why not)?

5. Does your calculation of the time it takes the ball to reach the floor (question 3) depend on the **speed** of the ball when it leaves the ramp? Why (or why not)?

6. From your experimental measurements calculate the horizontal speed of each ball:

Results:

Mass (g)				Radius (cm)				d (m)				v (m s^{-1})
1	2	3	Mean	1	2	3	Mean	1	2	3	Mean	

7. Which ball is moving fastest?

____

Note that the experimental value for the velocity is in each case lower than velocity in your calculations in question 1.

8. Which approximation in your calculations might be responsible for this difference?

____

____

____

Conclusion:

____

____

____

____

Appendix E: Acceleration questionnaire

The following was written on the blackboard:

Please rate the examples on the following scale.

0	1	2	3	4	5	6	7	8	9	10
Not an example	Very poor example				Average example					Excellent example

1 A car getting faster
2 A car turning a corner
3 A car that is parked
4 A car starting from traffic lights when they turn green
5 A car slows down
6 A car that falls as it goes over the edge of a cliff.

The above was then erased and the following written up:

There are three types of acceleration. Please list them in order from best to worst.

Appendix F: Force questionnaire

Examples of Forces: How good are they?

Please rate the examples on the following scale.

0	1	2	3	4	5	6	7	8	9	10
Not an example	Very poor example				Average example					Excellent example

Part One

Words are sometimes used differently in physics and in everyday life. Each of the following words is taken from a thesaurus under the heading "force". Rate how good an example of "a force" (as it is used in physics) each item is. The first set consists of nouns.

1	an army	0 1 2 3 4 5 6 7 8 9 10
2	a pressure	0 1 2 3 4 5 6 7 8 9 10
3	a propulsion	0 1 2 3 4 5 6 7 8 9 10
4	a punch	0 1 2 3 4 5 6 7 8 9 10
5	a push	0 1 2 3 4 5 6 7 8 9 10
6	a shame	0 1 2 3 4 5 6 7 8 9 10
7	a sinew	0 1 2 3 4 5 6 7 8 9 10
8	a squad	0 1 2 3 4 5 6 7 8 9 10
9	a squeeze	0 1 2 3 4 5 6 7 8 9 10
10	a strength	0 1 2 3 4 5 6 7 8 9 10

Each of the following words is taken from a thesaurus under the heading "force". Rate how good an example of "to force" (as it is used in physics) each item is.This set consists of verbs.

11	to compel	0 1 2 3 4 5 6 7 8 9 10
12	to cram	0 1 2 3 4 5 6 7 8 9 10
13	to make	0 1 2 3 4 5 6 7 8 9 10
14	to oblige	0 1 2 3 4 5 6 7 8 9 10
15	to pack	0 1 2 3 4 5 6 7 8 9 10
16	to press	0 1 2 3 4 5 6 7 8 9 10
17	to propel	0 1 2 3 4 5 6 7 8 9 10
18	to require	0 1 2 3 4 5 6 7 8 9 10
19	to shove	0 1 2 3 4 5 6 7 8 9 10
20	to thrust	0 1 2 3 4 5 6 7 8 9 10

Each of the following words is also taken from a thesaurus under the heading "force". Rate how good an example of "force" (as it is used in physics) each item is.This set consists of scientific terms.

21	drive	0 1 2 3 4 5 6 7 8 9 10
22	energy	0 1 2 3 4 5 6 7 8 9 10
23	impetus	0 1 2 3 4 5 6 7 8 9 10
24	impulse	0 1 2 3 4 5 6 7 8 9 10
25	inertia	0 1 2 3 4 5 6 7 8 9 10
26	momentum	0 1 2 3 4 5 6 7 8 9 10
27	power	0 1 2 3 4 5 6 7 8 9 10
28	resistance	0 1 2 3 4 5 6 7 8 9 10
29	strength	0 1 2 3 4 5 6 7 8 9 10
30	thrust	0 1 2 3 4 5 6 7 8 9 10

Before you turn over the page:

Write a sentence which describes a situation which you would consider a good example of a force.

Part Two

For each situation described below there is a different car and a different truck.In each of the following situations rate how good an example of a force is the force that the car exerts on the truck.

1	A car is on the back of a parked truck.	0 1 2 3 4 5 6 7 8 9 10
2	A car is driven past a parked truck, hits the truck's tail-light and breaks it. The car doesn't stop.	0 1 2 3 4 5 6 7 8 9 10
3	A car has a truck resting on its roof for an advertisement.	0 1 2 3 4 5 6 7 8 9 10
4	A car is driven into the back of a parked truck. The car is wrecked. The truck is badly damaged.	0 1 2 3 4 5 6 7 8 9 10
5	A car is on the back of a truck. The truck is getting faster.	0 1 2 3 4 5 6 7 8 9 10
6	A car pulls a truck out of a mudhole where it was stuck.	0 1 2 3 4 5 6 7 8 9 10
7	A car is on the back of a truck. The truck goes around a curve at a constant speed.	0 1 2 3 4 5 6 7 8 9 10
8	A car is driven into the back of a parked truck. The car is wrecked. This truck is not damaged.	0 1 2 3 4 5 6 7 8 9 10
9	A car is driven past a parked truck. They do not touch.	0 1 2 3 4 5 6 7 8 9 10
10	A car collides with a truck at an intersection.	0 1 2 3 4 5 6 7 8 9 10

Each situation described below there is a different person and a different object. In each of the following situations rate how good an example of a force is the force that the person exerts on the object.

11	A boy sucks a milkshake.	0 1 2 3 4 5 6 7 8 9 10
12	A boy picks his nose.	0 1 2 3 4 5 6 7 8 9 10
13	A cook chops up a cabbage.	0 1 2 3 4 5 6 7 8 9 10
14	A criminal squeezes a gun's trigger.	0 1 2 3 4 5 6 7 8 9 10
15	A dentist pulls out a tooth.	0 1 2 3 4 5 6 7 8 9 10
16	A girl presses a button on a camera to take a picture.	0 1 2 3 4 5 6 7 8 9 10
17	A soccer player kicks a ball.	0 1 2 3 4 5 6 7 8 9 10
18	A student spits out a watermelon seed.	0 1 2 3 4 5 6 7 8 9 10
19	A teenager chews gum.	0 1 2 3 4 5 6 7 8 9 10
20	A tired shopper sits on a park bench.	0 1 2 3 4 5 6 7 8 9 10
21	A weight lifter lifts 100 kg from the floor to shoulder height.	0 1 2 3 4 5 6 7 8 9 10
22	A boxer punches a boxing bag.	0 1 2 3 4 5 6 7 8 9 10

For each situation described below there is a different object.

In each of the following situations something is happening to the object. Rate how good an example of a force is the force on the object.

23	An avalanche of snow goes down a mountain side.	0 1 2 3 4 5 6 7 8 9 10
24	An iron bar is twisted out of shpae.	0 1 2 3 4 5 6 7 8 9 10
25	A cyclist rides around a circular cycle track.	0 1 2 3 4 5 6 7 8 9 10
26	A rubber band stretches.	0 1 2 3 4 5 6 7 8 9 10
27	A metal bar expands as the temperature rises.	0 1 2 3 4 5 6 7 8 9 10
28	A cat jumps into the air.	0 1 2 3 4 5 6 7 8 9 10
29	A ball rolls along a horizontal table.	0 1 2 3 4 5 6 7 8 9 10
30	A ball rolls up a ramp.	0 1 2 3 4 5 6 7 8 9 10
31	A ball rolls down a ramp.	0 1 2 3 4 5 6 7 8 9 10
32	A rock falls 100 m.	0 1 2 3 4 5 6 7 8 9 10
33	A string vibrates.	0 1 2 3 4 5 6 7 8 9 10
34	A pendulum bob reaches its highest point.	0 1 2 3 4 5 6 7 8 9 10
35	A car gets faster..	0 1 2 3 4 5 6 7 8 9 10
36	A roller coaster goes over the top of a "hill"	0 1 2 3 4 5 6 7 8 9 10
37	A rock drops 10 mm.	0 1 2 3 4 5 6 7 8 9 10
38	A car gets slower.	0 1 2 3 4 5 6 7 8 9 10
39	An iron bar is compressed.	0 1 2 3 4 5 6 7 8 9 10
40	A car goes around a curve.	0 1 2 3 4 5 6 7 8 9 10
41	Water comes out from a hose nozzle.	0 1 2 3 4 5 6 7 8 9 10
42	A person walks down the street.	0 1 2 3 4 5 6 7 8 9 10
43	A door is pushed gently but it doesn't open.	0 1 2 3 4 5 6 7 8 9 10
44	A door is pushed gently and it opens.	0 1 2 3 4 5 6 7 8 9 10
45	A door is pushed hard but it doesn't open.	0 1 2 3 4 5 6 7 8 9 10
46	A door is pushed hard and it opens.	0 1 2 3 4 5 6 7 8 9 10
47	A car starts when the lights turn green.	0 1 2 3 4 5 6 7 8 9 10
48	A bomb explodes.	0 1 2 3 4 5 6 7 8 9 10
49	An elephant sleeps.	0 1 2 3 4 5 6 7 8 9 10
50	The Moon orbits the Earth.	0 1 2 3 4 5 6 7 8 9 10

Appendix G: End of course questionnaire (adapted from Marsh, Roche, & Australia. Dept. of Employment Education and Training. Evaluations and Investigations Program., 1994)

The following questionnaire differs only in the introductory paragraph and in the Assignments/Reading section from the original version due to Marsh.

STUDENTS' EVALUATION OF EDUCATIONAL QUALITY (SEEQ)

Do NOT put your name on this survey. Please complete it as accurately and as candidly as possible. This is part of a larger project to improve teaching effectiveness in Physics. The purpose of this survey is to provide your teacher with feedback about his/her teaching effectiveness. For this reason you should base your responses on his/her teaching in this subject. If any items are not applicable, simply leave them blank.
Space has been provided on the back of this form for any written comments you may have.

Teacher:.. Subject:............................ Year Level: .... Date:............../............./............

Please indicate the EXTENT of your agreement/disagreement with the following statements as descriptions of this subject by using the following scale. Circle the number.

1	2	3	4	5	6	7	8	9
Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree

Learning / Academic Value	You found the class intellectually challenging and stimulating.	1	2	3	4	5	6	7	8	9
	You have learned something which you considered valuable.	1	2	3	4	5	6	7	8	9
	You have learned and understood the subject materials in this class.	1	2	3	4	5	6	7	8	9
Teacher Enthusiasm	Teacher was enthusiastic about teaching the class.	1	2	3	4	5	6	7	8	9
	Teacher was dynamic and energetic in conducting the class.	1	2	3	4	5	6	7	8	9
	Teacher enhanced presentations with the use of humour.	1	2	3	4	5	6	7	8	9
	Teacher's style of presentation held your interest during class.	1	2	3	4	5	6	7	8	9
Organisation / Clarity	Teacher's explanations were clear.	1	2	3	4	5	6	7	8	9
	Class materials were well prepared and carefully explained.	1	2	3	4	5	6	7	8	9
	Proposed objectives agreed with those actually taught so you knew where the class was going.	1	2	3	4	5	6	7	8	9
	Teacher gave presentations that facilitated taking notes.	1	2	3	4	5	6	7	8	9
Group Interaction	Students were encouraged to participate in class discussions.	1	2	3	4	5	6	7	8	9
	Students were invited to share their ideas and knowledge.	1	2	3	4	5	6	7	8	9
	Students were encouraged to ask questions and were given meaningful answers.	1	2	3	4	5	6	7	8	9
	Students were encouraged to express their own ideas and/or question the teacher.	1	2	3	4	5	6	7	8	9
Individual Rapport	Teacher was friendly towards individual students.	1	2	3	4	5	6	7	8	9
	Teacher had a genuine interest in individual students	1	2	3	4	5	6	7	8	9
	Teacher made students feel welcome in seeking help/advice in or outside of class.	1	2	3	4	5	6	7	8	9
	Teacher was adequately accessible to students during office hours or after class.	1	2	3	4	5	6	7	8	9
Breadth of Coverage	Teacher contrasted the implications of various theories.	1	2	3	4	5	6	7	8	9
	Teacher presented the background or origin of ideas/concepts developed in class.	1	2	3	4	5	6	7	8	9
	Teacher presented points of view other than his/her own when appropriate.	1	2	3	4	5	6	7	8	9
	Teacher adequately discussed current developments in the field.	1	2	3	4	5	6	7	8	9
Examinations / Grading	Feedback on examinations / graded material was valuable.	1	2	3	4	5	6	7	8	9
	Methods of evaluating student work were fair and appropriate.	1	2	3	4	5	6	7	8	9
	Examinations / graded materials tested class content as emphasised by the lecture.	1	2	3	4	5	6	7	8	9

Assignments / Reading	The textbook was helpful in understanding the theory.	1	2	3	4	5	6	7	8	9
	The study guides were helpful in understanding the theory.	1	2	3	4	5	6	7	8	9
	Practical work was helpful in understanding the theory.	1	2	3	4	5	6	7	8	9
	Writing a summary was helpful in understanding the theory.	1	2	3	4	5	6	7	8	9
	The technical writing project was helpful in understanding the theory.									
	Doing past years exam papers were helpful in understanding the theory.									
	Group discussions of the force-concept test were helpful in understanding the theory.									
Overall Rating	(1 = Very Poor...3= Poor...5 = Average...7 = Good... 9 = Very Good)									
	Overall, how does this class compare with other classes at this institution?	1	2	3	4	5	6	7	8	9
	Overall, how does this teacher compare with other teachers at this institution?	1	2	3	4	5	6	7	8	9

BACKGROUND SUBJECT / CLASS CHARACTERISTICS

Subject difficulty, relative to other subjects was:	(1 = Very Easy ... 5 = Medium, ... 9 = Very Hard)		1	2	3	4	5	6	7	8	9
Subject workload, relation to other subjects was:	(1 = Very Light ... 5 = Medium ...9 = Very Heavy)		1	2	3	4	5	6	7	8	9
Subject pace was:	(1 = Too Slow .... 5 = About right ... 9 = Too Fast)		1	2	3	4	5	6	7	8	9
Average number of hours per week required outside of class	(0 = None, 1 = 1 hr, 2 = 2 hrs ... 9 = 9 hrs)	0	1	2	3	4	5	6	7	8	9
Your sex:	(1 = Male, 2 = Female)		1	2							
Your expected subject mark:	(1 = F, 2 = E, 3 = D, 4 = C, 5 = B, 6 = A)		1	2	3	4	5	6			
In comparison with other subjects at this institution, how easy is it to get good marks in this subject?	(1 = Very Easy ... 3 = Easy ... 5 = About Average ... 7 = Difficult ... 9 = Very Difficult)		1	2	3	4	5	6	7	8	9
Reason for taking subject:	(1 = Course requirement, 2 = Chosen course major, 3 = Course elective, 4 = General elective, 5 = General Interest only, 6 = Other)		1	2	3	4	5	6			
Level of interest in the subject before the start of the class:	(1 = Very Low ... 5 = Medium ... 9 = Very High)		1	2	3	4	5	6	7	8	9

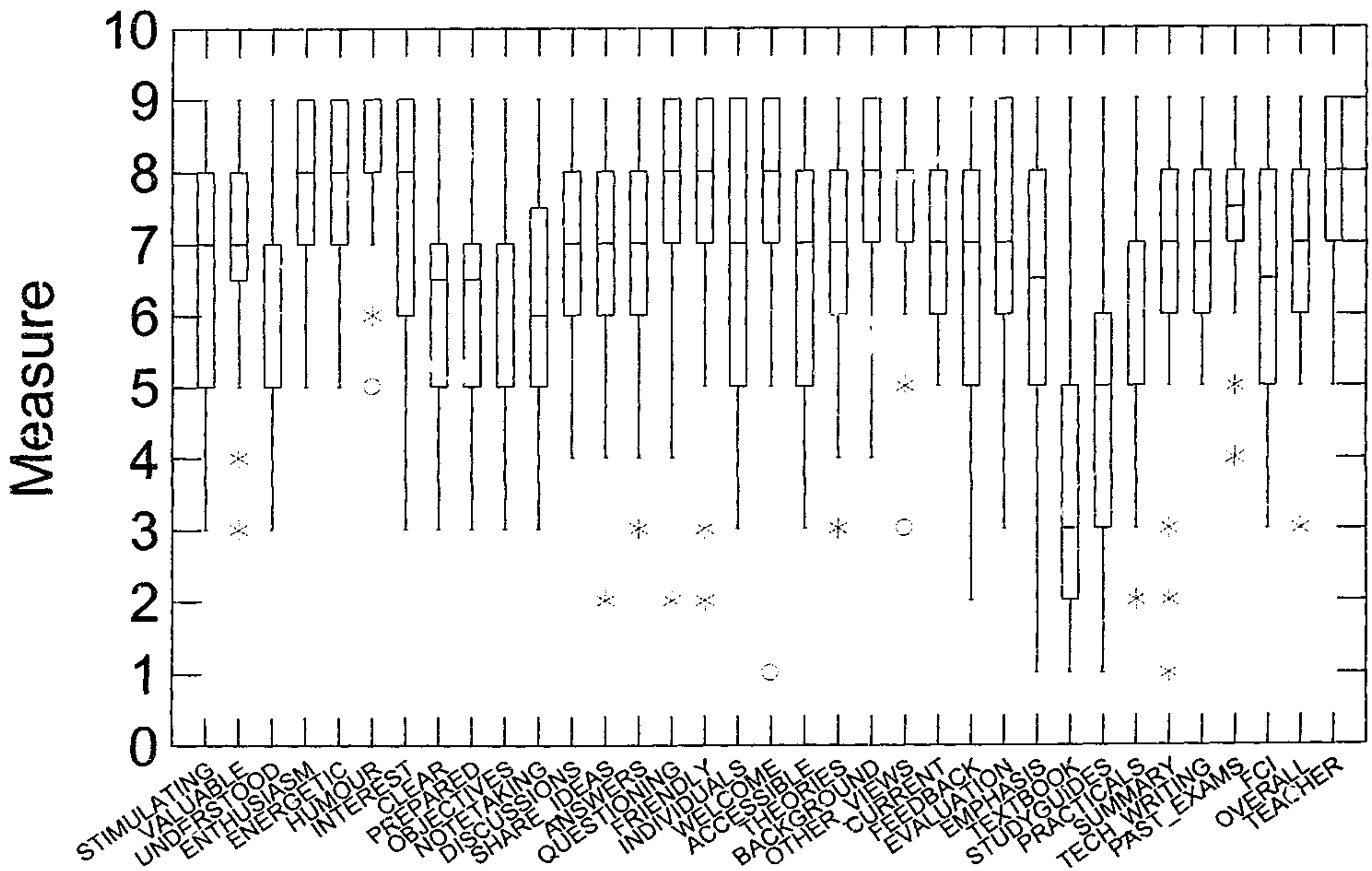
Any other comments:

____

____

____

____


Measure
10
9
8
7
6
5
4
3
2
1
0
STIMULATING
VALUABLE
UNDERSTOOD
ENTHUSIASM
ENERGETIC
HUMOUR
INTEREST
CLEAR
PREPARED
OBJECTIVES
NOTETAKING
DISCUSSIONS
SHARE_IDEAS
ANSWERS
QUESTIONING
FRIENDLY
INDIVIDUALS
WELCOME
ACCESSIBLE
THEORIES
BACKGROUND
OTHER_VIEWS
CURRENT
FEEDBACK
EVALUATION
EMPHASIS
TEXTBOOK
STUDYGUIDES
PRACTICALS
SUMMARY
TECH_WRITING
PAST_EXAMS
FCI
OVERALL
TEACHER

Agnes, M. (Ed.). (1996). *Websters new world dictionary and thesaurus.* NY: Macmillan General Reference.

Aitchison, J. (1994). *Words in the mind: an introduction to the mental lexicon* (2nd ed.). Oxford: Blackwell Publishers Ltd.

Alexopoulou, E., & Driver, R. (1996). Small group discussion in Physics: peer interaction modes in pairs and fours. *Journal of Research in Science Teaching, 33*(10), 1099-1114.

Alexopoulou, E., & Driver, R. (1997). Gender differences in small group discussion in physics. *International Journal of Science Education, 19*(4), 393-406.

Alsop, & Watts, M. (2000). Facts and feelings: exploring the affective domain in the learning of physics. *Physics Education, 35*(2), 132-138.

Ameh, C., & Gunstone, R. (1985). Teachers' concepts in science. *Research in Science Education, 15*, 151-157.

Anderson, J. (1996). *Communication theory: epistemological foundations.* New York and London: The Guilford Press.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, Mass.: Harvard University Press.

Anderson, J. R. (1996). *The architecture of cognition.* Mahwah, NJ: Erlbaum.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38*, 341-380.

Aristotle. (1899). *The organon: or logical treatises of Aristotle.* London: George Bell and Sons.

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. In E. Margolis & S. Laurence (Eds.), *Concepts* (pp. 225-259). Cambridge, Massachusetts: The MIT Press.

Ausubel, D. P. (1968). *Educational psychology: a cognitive view.* New York: Holt, Rinehart & Winston.

Bach, E. (1989). *Informal lectures on formal semantics.* New York: State Universty of New York.

Baddeley, A. (1997). *Human memory: theory and practice. Revised edition.* Hove, East Sussex: Psychology Press.

Bao, L., Zollman, D., Hogg, K., & Redish, E. F. (2000). *Model analysis of fine structures of student models: an example with Newton's Third Law.* Retrieved 3 October, 2001, from the World Wide Web: http://www.phys.ksu.edu/perg/papers/2000/Fine_Structure_Pre.pdf

Bar, V., Zinn, B., Goldmuntz, R., & Sneider, C. (1994). Childrens' concepts about weight and free fall. *Science Education, 78*(2), 149-169.

Bar, V., Zinn, B., & Rubin, E. (1997). Children's ideas about action at a distance. *International Journal of Science Education, 19*(10), 1137-1157.

Barnes, D. (1975). *From communication to curriculum.* Harmondsworth: Penguin.

Barnes, D., Britton, J., Rosen, H., & L.A.T.E., T. (1971). *Language, the learner and the school. Revised edition.* Harmondsworth: Penguin.

Barnes, D., & Todd, F. (1977). *Communication and learning in small groups.* London: Routledge and Kegan Paul.

Barsolou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory and Cognition, 11*, 629-654.

Bartlett, F. (1932). *Remembering: a study in experimental and social psychology.* Cambridge, UK: Cambridge University Press.

Basu, R. S. (Ed.). (1942). *Chalantika Adhunik Bangobhasar Abhidan (Modern Bengali Dictionary)* (4th ed.). Calcutta: M. C. Sorkar & Sons Ltd.

Bell, B. (Ed.). (1995). *Responses to 'Challenging NZ science education' by Michael R. Mathews.* Hamilton, NZ: Centre for Science, Mathematics and Technology Education Research.

Bellinger, D., & Gleason, J. (1982). Sex differences in parental directives to young children. *Journal of Sex Roles*(8), 1123 -1139.

Bennet, R. (1993). On the meanings of constructed response. In R.Bennet & W. Ward (Eds.), *Construction versus choice in cognitive measurement.* Hillsdale, NJ: Lawrence Erlbaum.

Berlin, B., & Kay, P. (1969). *Basic color terms and their evolution.* Berkeley: University of California Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: investigating language structure and use.* Cambridge: Cambridge University Press.

Biggs, J. (1987). *Student approaches to learning and studying.*

Bliss, L. (1988). The development of modals. *The Journal of Applied Developmental Psychology, 9,* 253-261.

Bloom, L. (1970). *Language development: Form and function in emerging grammars.* Cambridge,MA: MIT Press.

Bloom, L., Hood, L., & Lightbown, P. (1974). Imitation in language development: if,when and why. *Cognitive Psychology, 6,* 380-420.

Bloom, L., Lightbown, P., & Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development, 40*(160).

BNC. (1994). *British National Corpus.* Oxford University Computing Services. Retrieved 30 August, 2001, from the World Wide Web: http://www.hcu.ox.ac.uk/BNC/

Bohannon, J. N., & Marquis, A. L. (1977). Children's control of adult speech. *Child Development*(48), 1002 -1008.

Bolton, N. (1977). *Concept formation.* Oxford: Pergamon Press.

Bowden, J., Dall'Alba, G., Martin, E., Laurillard, D., Marton, F., Masters, G., Ramsden, P., Stephanou, A., & Walsh, E. (1992). Displacement, velocity and frames of reference: Phenomenographic studies of students understanding and some implications for teaching and assessment. *American Journal of Physics, 60*(3), 262-269.

Bower, G. (1992). *How might emotions affect learning?*

Braine, M. D. S. (1976). Children 's first word combinations. *Monographs of the Society for Research in Child Development, 41*(1).

Brincones, I., & Otero, J. (1994). Students' conceptions of the top-level structure of physics texts. *Science Education, 78*(2), 171-183.

Brouwer, W. (1984). Problem-posing physics: a conceptual approach. *American Journal of Physics, 52*(7), 602-607.

Brown, D. E. (1989). Students' concept of force: the importance of understanding Newton's third law. *Physics Education, 24,* 353-358.

Brown, D. E. (1992). Using examples and analogies to remediate misconceptions in physics: factors influencing conceptual change. *Journal of Research in Science Teaching, 29*(1), 17-34.

Brown, D. E. (1994). Facilitating conceptual change using analogies and explanatory models. *International Journal of Science Education, 16,* 201-214.

Brown, D. E., & Clement, J. (1989). Overcoming misconceptions via analogical reasoning: abstract transfer versus explanatory model construction. *Instructional Science, 18*, 237-261.

Brown, R. (1973). *A first language: the early stages.* Cambridge, MA: Harvard University Press.

Brun, W., & Teigen, K. H. (1988). Verbal probabilities: ambiguous, context-dependent, or both? *Organizational Behavior and Human Decision Processes, 41*, 405-425.

Bruner, J. S., Olver, R. R., & Greenfield, P. M. (1966). *Studies in Cognitive Growth.* New York: Wiley.

Bryan, R. A., & others, a. (1988). Illustrating Newton's Second Law with the Automobile Coast-Down Test. *Physics Teacher, 26*, 442-445.

Bunker, K. (1991). A new free-fall experiment to determine the acceleration due to gravity. *Physics Education, 26*(6), 386-390.

Carey, S. (1991). Knowledge acquisition: enrichment of conceptual change? In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: essays on biology and cognition.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Carruthers, M. (1990). *The book of memory: a study of memory in medieval culture.* Cambridge, UK: Cambridge University Press.

Carterette, E. C., & Jones, M. H. (1974). *Informal Speech: Alphabetic and Phonemic texts with statistical analyses and tables.* Berkeley, CA: University of California Press.

Carver, R. P. (1978). The case against statistical significance testing. *Harvard Educational Review, 48*(3), 378-399.

Chafe, W. (1980). The deployment of consciousness in the production of a narrative. In W. Chafe (Ed.), *The pear stories: cognitive, cultural, and linguistic aspects of narrative production.* Norwood, NJ: Ablex Publishing Corporation.

Chafe, W. (1995). *Discourse, consciousness and time.* Chicago and London: The University of Chicago Press.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 27-55). New York: Academic Press.

Chambers, S. K., & Andre, T. (1997). Gender, prior knowledge, interest, and experience in electricity and conceptual change text manipulations in learning about direct current. *Journal of Research in Science Teaching, 34*(2), 107-123.

Champagne, A., Gunstone, R., & Klopfer, L. E. (1985). Effecting changes in cognitive structure amongst physics students. In L. West & L. Pines (Eds.), *Cognitive structure and cognitive change* (pp. 163-187). Orlando: Academic Press.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorisation and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, NJ: Lawrence Erlbaum Associates.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*(1), 66-71.

Cobern, W. W. (1996). Worldview theory and conceptual change in science education. *Science Education, 80*(5), 579-610.

COBUILD. (2000). *CobuildDirect Corpus Sampler* [Internet Web Site]. Collins and Birmingham University. Retrieved, from the World Wide Web: http://titania.cobuild.collins.co.uk/form.html

COD. (1993). *Concise Oxford Dictionary* [CD ROM]. Oxford University Press. Retrieved Jan 18, 2000, from the World Wide Web:

Cole, S. (1992). *Making science: between science and nature.* Cambridge, Massachusetts: Harvard University Press.

Coleman, L., & Kay, P. (1981). Prototype semantics: the English word lie. *Language, 57*(1), 26-44.

Confrey, J. (1995). How compatible are radical constructivism, sociocultural approaches, and social constructivism? In L. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 185-225). Hillsdale, NJ: Lawrence Erlbaum.

Cruttenden, A. (1978). Assimilation in child language and elsewhere. *Journal of ChildLanguage*(5), 373 -378.

Dall'Alba, G., Walsh, E., Bowden, J., Martin, E., Masters, G., Ramsden, P., & Stephanou, A. (1993). Textbook treatments and students' understanding of acceleration. *Journal of Research in Science Teaching, 30*(7), 621-635.

Dawson, C., & Rowell, J. (1990). New data and prior belief: the two faces of scientific reasoning. *Research in Science Education, 20*(20), 48-56.

Dayhoff, J. (1990). *Neural network architectures: an introduction.* New York: Van Nostrand Reinhold.

de Finetti, B. (1972). *Probability, induction and statistics: the art of guessing.* Chichester: Wiley.

de Jong, E. (1996). *Conceptual change and reflection in physics learning.* Unpublished Ph. D., Monash University.

de Jong, M. L. (1988). What name should be used for the force required to move a mass in a circle? 1. Centripetal force, 2. Centrifugal force, 3. None of the above. *Physics Education, 26*, 470-471.

Deakin, M. A. B., & Troup, G. J. (1981). Centrifugal force: an elucidation. *The Australian Science Teachers Journal, 28*(3), 33-37.

Dekkers, P. J. J. M., & Thijs, G. D. (1998). Making productive use of students' initial conceptions in developing the concept of force. *Science Education.*

Delbridge, A., Bernard, J. R. L., Blair, D., Butler, S., Peters, P., & Yallop, C. (Eds.). (1997). *The Macquarie dictionary* (3rd ed.). Macquarie, NSW: The Macquarie Library Pty Ltd.

Derrida, J. (1976). *Of Grammatology* (G. C. Spivak, Trans.). Baltimore: Johns Hopkins University Press.

Dewey, J. (1963). *Experience and education.* New York: Collier Books.

di Sessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education, 20*(10), 1155-1191.

Dinsmore, J. (1991). *Partitioned representations.* Dordrecht: Kluwer Academic Publishers.

Domann, F. E. (1982). An Improved Newton's Second Law Experiment. *American Journal of Physics, 50*, 185-186.

Domenech, A., Casasus, E., & Domenech, M. (1993). The classical concept of mass: theoretical difficulties and students' definitions. *International Journal of Science Education, 15*(2), 163-173.

Domin, D. S. (1996). Comment: concept mapping and representational systems. *Journal of Research in Science Teaching, 33*(8), 935-936.

Dreyfus, H. (1991). *Being-in-the-world: a commentary on Heidegger's Being and Time, Division I.* Cambridge, Massachusetts: The MIT Press.

Dreyfus, H. (1993). *What computers still can't do: a critique of artificial reason.* Cambridge, Massachusetts: The MIT Press.

Driver, J. C. (1995). Constructivist approaches to science teaching. In L. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 385-400). Hillsdale, NJ: Lawrence Erlbaum.

Driver, R. (1983). *The pupil as scientist?* Milton Keynes and Philadelphia: Open University Press.

Driver, R., & Erickson, G. (1983). Theories-in-action: some theoretical and empirical issues in the study of students' conceptual frameworks in science. *Studies in Science Education, 10*, 37-60.

Driver, R., Guesne, E., & Tiberghien, A. (Eds.). (1985). *Children's ideas in science.* Milton Keynes and Philadelphia: Open University Press.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (Eds.). (1995). *Making sense of secondary science.* London: Routledge and Kegan Paul.

Duit, R. (1990). Preconceptions and misconceptions. In I. T. H. T. Posthelwaite (Ed.), *The international encyclopaedia of education: research and supplementary volume 2.* (pp. 4648-4652). Oxford: Pergamon Press.

Duit, R. (1993). Research on students' conceptions - developments and trends, *The proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics.* Ithaca, NY: Misconceptions Trust.

Duit, R. (1995). The constructivist view: a fashionable and fruitful paradigm for science education research and practice. In L. Steffe & J. Gale (Eds.), *Constructivism in education.* Hillsdale, NJ: Lawrence Erlbaum.

Dunn, C. S. (1983). The influence of instructional methods on concept learning. *Science Education, 67*(5), 647-656.

Dykstra Jr., D. I. (1992). Studying conceptual change: constructing new understandings. In R. Duit & F. Goldberg & H. Niederrer (Eds.), *Research in physics learning: theoretical issues and empirical studies. Proceedings of an international workshop held at the University of Bremen, March 4-8, 1991.* Kiel: IPN.

Eger, M. (1992). Hermeneutics in science education. *Science and Education, 1*(1), 337-348.

Eger, M. (1993a). Hermeneutics as an approach to science: part I. *Science and Education, 2*(2), Jan-29.

Eger, M. (1993b). Hermeneutics as an approach to science: part II. *Science and Education, 2*(2), 303-328.

Eger, M. (1995). Alternative interpretations, history and experiment: reply to Cushing, Crease, Bevilacqua and Gianetto. *Science and Education, 4*(4), 173-188.

Emmott, C. (1997). *Narrative comprehension : a discourse perspective.* Oxford: Oxford University Press.

Entwistle, T., Gentile, D., Graf, E., Hulme, S., Schoch, J., Strassenberg, A., Swartz, C., Chiaverina, C., Clark, R. B., Durkin, T., Gavenda, D., Peterson, F., Robertson, C., & Sears, R. (1999). Survey of high-school physics texts. *The Physics Teacher, 37*, 296-308.

Everitt, B. S. (1988). Cluster analysis. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: an international handbook* (pp. 601-605). Oxford: Pergamon Press.

Evrard, N., Huynen, A.-M., & Borght, C. d. B.-V. (1998). Communication of scientific knowledge in class - from verbalization to the concept of chemical equilibrium. *International Journal of Sceince Education, 20*(8), 883-900.

Fawns, R., & Sadler, J. (1996). Managing student's learning in classrooms: reframing classroom research. *Research in Science Education, 26*(2).

Finegold, M., & Gorsky, P. (1991). Students concepts of force as applied to related physical systems: a search for consistency. *International Journal of Science Education, 13*(1), 97-113.

Fletcher, P., & Garman, M. (1988). Normal language development and language impairment: Syntax and beyond. *Clinical Linguistics and Phonetics, 2*(97 -114.).

Flores, A., & Turner, E. E. (2001). Inclined planes and motion detectors: a study of acceleration. *School Science and Mathematics, 101*(3), 154-160.

Forbus, K. (1988). Qualitative Physics: Past, Present and Future. In H. Shrobe (Ed.), *Exploring artificial intelligence: survey talks from the national conferences on artificial intelligence* (pp. 239-296). San Mateo, CA: Morgan Kauffmann Publishers Ltd.

Fosnot, C. T. (Ed.). (1996). *Constructivism: theory, perspectives and practice*. New York: Teachers College Press.

Fox, R. (2001). Constructivism examined. *Oxford Review of Education, 27*(1), 23-35.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: lexicon and grammar*. Boston, Massachusetts: Houghton Mifflin Company.

Franklin, A. (1976). *The principle of inertia in the middle ages*. Boulder, Colarado: Colorado Associated University Press.

Frick, R. W. (1996). The appropriate use of null hypothesis testing. *Psychological Methods, 1*(4), 379-390.

Fry, D. B. (1979). *The physics of speech*. Cambridge: Cambridge University Press.

Gadamer. (1989). *Truth and method* (J. Weinsheimer, Trans. 2nd revised ed.). New York: The Continuum Publishing Company.

Galili, I. (1994). Weight and gravity: teachers ambiguity and students confusion about the concepts. *International Journal of Science Education, 15*(2), 149-162.

Galili, I., & Bar, V. (1992). Motion implies force: where to expect vestiges of the misconception? *International Journal of Science Education, 14*(1), 63-81.

Galili, I., & Bar, V. (1997). Children's operational knowledge about weight. *International Journal of Science Education, 19*(3), 317-340.

Galili, I., & Kaplan, D. (1996). Students' operations with the weight concept. *Science Education, 80*(4), 457-487.

Gardner, H. (1987). *The mind's new science: a history of the cognitive revolution*. USA: Basic Books. A division of HarperCollinsPublishers.

Gardner, P. (1974). Language Difficulties of Science Students. In M. Aldaham (Ed.), *Science and the ESL student* (pp. 23-33). Canberra: AGPS.

Gardner, P. (1977). Logical connectives in science: A summary of findings. *Research in Science Education, 7*, 9-24.

Gardner, P. (1981). On centrifugal force. *The Australian Science Teachers Journal, 27*(3), 69-74.

Gardner, P. (1984). Circular motion: some post-instructional frameworks. *Research in Science Education, 14*, 136-145.

Gardner, P. L. (1975). *The structure of science education*. Hawthorn, Vic.: Longman Australia.

Gardner, P. L. (1982). Centrifugal force: a reply to Deakin and Troup. *The Australian Science Teachers Journal, 28*(3), 39-43.

Garrard, J. (1986). Creative learning in science education. *Research in Science Education, 16*, 1-10.

Gauld, C. (1987). Student beliefs and cognitive structure. *Research in Science Education, 17*, 87-93.

Gauld, C. (1993). The historical context of Newton's third law and the teaching of mechanics. *Research in Science Education, 23*, 95-103.

Gauld, C. (1998). Solutions to the problem of impact in the 17th and 18th centuries and teaching Newton's Third Law today. *Science and Education, 7*, 49-67.

Geeraerts, D. (1997). *Diachronic prototype semantics: a contribution to historical lexicography*. Oxford: Oxford University Press.

Geiger, H. B. (1968). *Effectiveness in Learning Newton's Second Law of Motion in Secondary School Physics Using Three Methods of Learning*. Philadelphia, PA: Temple University.

Gentner, D., & Stevens, A. (Eds.). (1983). *Mental models.* Hillsdale, NJ: Lawrence Earlbaum.

Gernsbacher, M. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Gernsbacher, M. A. (1994). *Handbook of psycholinguistics.* San Diego: Academic Press.

Gilbert, J., Watts, M., & Osborne, R. (1985). Eliciting students' views about an interview-about-instance technique. In L. West & A. Pines (Eds.), *Cognitive Structure and Conceptual Change* (pp. 11-29). London:: Academic.

Gilbert, J. K., & Watts, D. M. (1983). Concepts, misconceptions and alternate conceptions: changing perspectives in science education. *Studies in Science Education, 10*, 61-98.

Giles, H., Robinson, W. P., & Smith, P. M. (1980). *Language: social psychological perspectives.* Oxford, UK: Pergamon.

Ginsburg, H., & Opper, S. (1969). *Piaget's theory of intellectual development: an introduction.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Gleason, J., Perlmann, R. Y., & Greif, E. B. (1984). What 's the magic word? Learning language through routines. *Discourse Processes, 6*, 493 -502.

Gleason, J. B. (1980). The acquisition of social speech and politeness formulae. In H. Giles & W. P. Robinson & P. M. Smith (Eds.), *Language: social psychological perspectives.* Oxford, UK: Pergamon.

Gleason, J. B., & Greif, E. (1983). Men 's speech to young children. In B.Thorne & C.Kramerae & N.Henley (Eds.), *Language,Gender and Society.* Rowley, MA: Newbury.

Goddard, C. (1998). *Semantic analysis: a practical introduction.* Oxford: Oxford University Press.

Gove, P. B. (Ed.). (1976). *Webster's third new international dictionary of the English language.* Springfield: G & C Merriam Company, Publishers.

Graham, T., & Berry, J. (1996). A hierarchical model of the development of student understanding of momentum. *International Journal of Science Education, 18*(1), 75-89.

Grayson, D. J. (1994). *Concept substitution: an instruction strategy for promoting conceptual change.* Paper presented at the 25th Annual Conference of the Australian Science Education Research Association, Hobart Tasmania.

Greca, I. M., & Moreira, M. A. (1997). The types of mental representations - models, propositions and images - used by college physics students regarding the concept of field. *International Journal of Science Education, 19*(6), 711-724.

Green, S. K., & Gredler, M. E. (2002). A review and analysis of constructivism for school-based practice. *School Psychology Review, 31*(1), 53-70.

Greeno, J. G. (1983). Conceptual entities. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 227-252). Hillsdale, NJ: Lawrence Earlbaum.

Greif, E. B., & Gleason, J. B. (1980). Hi,thanks and goodbye: more routine information. *Language in Society, 9*(159-166).

Griffiths, A. K., & Preston, K. R. (1992). Grade-12 students' misconceptions relating to fundamental characteristics of atoms and molecules. *Journal of Research in Science Teaching, 29*(6), 611-628.

Gruneberg, M. M., & Morris, P. (Eds.). (1994). *Theoretical aspects of memory* (2nd ed.). London and New York: Routledge.

Gunstone, R. (1979). *Thesis.* Unpublished PhD, Monash University, Clayton.

Gunstone, R. (1980). Word association and the description of cognitive structure. *Research in Science Education, 10*(10), 45-54.

Gunstone, R. (1984). Circular motion: some pre-instruction aternative frameworks. *Research in Science Education, 14*(14), 125-136.
Gunstone, R. (1988). Learners in science education. In P. Fensham (Ed.), *Development and dilemmas in science education* (pp. 80-95). London: The Falmer Press.
Gunstone, R. (1989). A comment on "The problem of terminology in the study of student conceptions in science". *Science Education, 73*(6), 643-646.
Gunstone, R., Chia, T., Giam, K., Sung-Jae, P., Pattanasuwan, C., Rodriguez, E., Swamy, N., & Talisayson, V. (1989). *Conceptions in mechanics: a survey of students beliefs in seven asian countries. Asia-Pacific Teachers and Educators Research Report No. 1.* Clayton, Vic.: Monash University.
Gunstone, R., & Watts, M. (1985). Force and motion. In R. Driver & E. Guesne & A. Tiberghien (Eds.), *Children's ideas in science.* Milton Keynes: Open University Press.
Gunstone, R. F., Champagne, A. B., & Klopfer, L. E. (1981). Instruction for understanding : a case study. *The Australian Science Teachers Journal, 27*(3), 27-32.
Gunstone, R. F., Gray, C. M. R., & Searle, P. (1992). Some long-term effects of uninformed conceptual change. *Science Education, 76*(2), 175-197.
Gutierrez, R., & Ogborn, J. (1992). A causal framework for analysing alternative conceptions. *International Journal of Science Education, 14*(2), 201-220.
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, 37*(7), 701-719.
Haggerty, L. (1929). What a two-and-one-half-year-old child said in one day. *Journal of Genetic Psychology*(38), 75 -100.
Haiman, J. (1980). Dictionaries and encyclopedias. *Lingua, 50*, 329-357.
Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64-74.
Halliday, M. A. K. (1985). *An introduction to functional grammar.* London: Edward Arnold.
Halloun, I. (1996). Schematic modelling for meaningful learning of physics. *Journal of Research in Science Teaching, 33*(9), 1019-1041.
Halloun, I. (1998). Schematic concepts from schematic models of the real world: the Newtonian concept of force. *Science Education, 82*, 239-263.
Hansen, R. (1971). The influence of variables other than knowledge on probabalistic tests. *Journal of Educational Measurement, 8*(1), 9-14.
Hart, C. (2002). If the Sun burns you is that a force? Some definitional prerequisites for understanding Newton's laws. *Physics Education, 37*(3), 234-238.
Hayes, P. J. (1979). The naive physics manifesto. In M. A. Boden (Ed.), *The philosophy of artificial intelligence* (pp. 171-205). Oxford: Oxford University Press.
Heidegger, M. (1962). *Being and time.* Oxford: Blackwell.
Heinze-Fry, J. A., & Novak, J. D. (1990). Concept mapping brings long-term movement toward meaningful learning. *Science Education, 74*(4), 461-472.
Heller, P., & Huffman, D. (1995). Interpreting the Force Concept Inventory: a reply to Hestenes and Halloun. *The Physics Teacher, 33*(November), 503-511.
Hellingman, C. (1989). Do forces have twin brothers? *Physics Education, 24*, 36-40.
Hellingman, C. (1992). Newton's third law revisited. *Physics Education, 27*, 112-115.
Hendrickson, A. D., & Buehler, R. J. (1971). Proper scores for probability forecasters. *The Annals of Mathematical Statistics, 42*(6), 1916-1921.
Hestenes, D. (1992). Modeling games in the Newtonian world. *American Journal of Physics, 60*(8), 732-748.

Hestenes, D. (1996, August 1996). *Modeling methodology for physics teachers.* Paper presented at the International Conference on Undergraduate Physics, College Park.

Hestenes, D., & Halloun, I. (1995). Interpreting the Force Concept Inventory: a response to March 1995 critique by Huffman and Heller. *The Physics Teacher, 33*(November), 502-506.

Hestenes, D., Wells, M., & Swackhammer, G. (1992). Force concept inventory. *The Physics Teacher, 30*(3), 141-158.

Hewson, P. (1981a). Aristotle: alive and well in the classroom? *The Australian Science Teachers Journal, 27*(3), 9-13.

Hewson, P. (1981b). A conceptual change approach to learning science. *European Journal of Science Education, 3*(3), 383-396.

Hirsch, E. D. (1967). *Validity in interpretation.* New Haven: Yale University Press.

Holland, P., & Wainer, H. (1993). *Differential item functioning.* Hillsdale, NJ: Lawrence Erlbaum.

Hollingworth, R. (1913). Experimental studies in judgement. *Archives of Psychology, 29*, 1-119.

Hood, C. G. (1992). Teaching Newton's second law - another viewpoint. *The Physics Teacher, 30*(September), 358-359.

Howard, R. (1987). *Concepts and schemata: an introduction.* London: Cassell.

Huffman, D., & Heller, P. (1995). What does the force concept inventory actually measure? *The Physics Teacher, 33*(March), 138-143.

Hurford, J., & Heasley, B. (1983). *Semantics: a coursebook.* Cambridge, UK: Cambridge University Press.

Hutchinson, T. (1991). *Ability, partial information, guessing: Statistical modelling applied to multiple-choice tests.* Adelaide, SA: Rumsby Scientific Publishing.

Jackendoff, R. (1983). *Semantics and cognition.* Cambridge, Massachusetts: The MIT Press.

Jacobs, S. S. (1971). Correlates of unwarranted confidence in responses to objective test items. *Journal of Educational Measurement, 8*(1), 15-20.

Johansson, S., & Hofland, K. (1989). *Frequency analysis of English vocabulary and grammar: based on the LOB Corpus: Tag frequencies and word frequencies* (Vol. 1). Oxford: Oxford University Press.

Johnson, M. (1986). *A computer-based approach to the analysis of child language data.* Unpublished Unpublished PhD Thesis, University of Reading.

Johnson, M. (1987). *The body in the mind: the bodily basis of meaning, imagination and reason.* Chicago and London: The University of Chicago Press.

Johnson-Laird, P. (1983). *Mental models.* Cambridge: Cambridge University Press.

Jones, A. T. (1983). Investigation of students' understanding of speed, velocity and acceleration. *Research in Science Education, 13*, 95-104.

Jones, M. H., & Carterette, E. C. (1963). Redundancy in children 's free-reading choices. *Journal of Verbal Learning and Verbal Behavior, 2*, 489 -493.

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language, 54*(3), 610-646.

Keil, F. C. (1991). *The emergence of theoretical beliefs as constraints on concepts.*

Kennedy, G. (1998). *An introduction to corpus linguistics.* London: Longman.

Kikoin, I. K. (1979). The Treatment of the Laws of Dynamics in Higher Level Schools. *European Journal of Science Education, 41*, 57-64.

Kline, P. (1994). *An easy guide to factor analysis.* London: Routledge.

Koballa, T. R., Jr. (1989). Using salient beliefs in designing a persuasive message about teaching energy conservation practices to children. *Science Education, 73*(5), 547-567.

Koehler, R. A. (1974). Overconfidence on probablistic tests. *Journal of Educational Measurement, 11*(2), 101-108.

Kornblith, H. (1994). *Naturalizing epistemology: second edition.* Cambridge, Massachusetts: The MIT Press.

Kraft, D. W., & Motz, L. (1995). Geometric derivation of radial acceleration magnitude. *The Physics Teacher, 33*(December), 565.

Kuczaj, S. (1976). *-ing,-s and -ed:A study of the acquisition of certain verb inflections.* Unpublished doctoral dissertation, University of Minnesota.

Kuhn, T. (1974). *The essential tension: selected studies in scientific tradition and change.* Chicago and London: The University of Chicago Press.

Kurtze, D. A. (1994). Teaching Newton's Second Law--A Better Way. *The Physics Teacher, 29*, 350-351.

Labov, W. (1973). The boundaries of words and their meanings. In C. J. Bailey & R. Shuy (Eds.), *New ways of analysing variation in English* (pp. 340-373). Washington: Georgetown University Press.

Lakoff, G. (1987). *Women, fire and dangerous things: what categories reveal about the mind.* Chicago and London: Chicago University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago and London: The Chicago University Press.

Lan, B. L. (2002). Overcoming students' misconception of centripetal force. *Physics Education, 37*(4), 361.

Langford, P. (1987). *Concept development in the secondary school.* London, New York, Sydney: Croom Helm Ltd.

Larousse (Ed.). (1954). *Larousse de Poche.* Paris: Librairie Larousse.

Lawrence, E. (1970). *The origins and growth of modern education.* Harmondsworth: Penguin.

Leech, G. (1981). *Semantics: the study of meaning* (2nd ed.). Harmondsworth: Penguin Books.

Leonard, W. J., Gerace, W. J., Dufresne, R. J., & Mestre, J. P. (1994, 20 April 1994). *Concept-based problem solving in physics* [website]. UMPERG, University of Massachusetts, Amherst. Retrieved 18 February 1998, 1998, from the World Wide Web: http://www-perg.phast.umass.edu/default.htm

Lewis, C. S. (1961). *Studies in words.* Cambridge: C.U.P.

Lorayne, H. (1986). *How to develop a super power memory.* New York: Thursons Publishing Group.

Lyons, J. (1981). *Language, meaning and context.* London: Fontana.

Lythcott, J. (1985). "Aristotelian" was given as the answer, but what was the question? *American Journal of Physics, 53*(5), 428-432.

MacWhinney, B. (2000a). *The CHILDES Project: Tools for analyzing talk.* (3rd ed. Vol. 1: Transcription format and programs). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

MacWhinney, B. (2000b). *The CHILDES Project: Tools for analyzing talk.* (3rd ed. Vol. 2: The database). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Maloney, D. P. (1984). Rule-governed approaches to physics - Newton's third law. *Physics Education, 19*, 37-42.

Mansion, J. E. (1977). *Harrap's new shorter French and English dictionary.* London: Harrap.

Marsh, H. W., Roche, L. A., & Australia. Dept. of Employment Education and Training. Evaluations and Investigations Program. (1994). *The use of students'* *evaluations of university teaching to improve teaching effectiveness : final project report.* Canberra: Australian Govt. Pub. Service.

Marton, F., & Booth, S. (1997). *Learning and awareness.* Mahwah New Jersey: Lawrence Erlbaum Associates, Publishers.

Masur, E., & Gleason, J. B. (1980). Parent -child interaction and the acquisition of lexical information during play. *Developmental Psychology,, 16*, 404 -409.

Matthews, M. R. (1995). *Challenging New Zealand Science Education.* Palmerston North: Dunmore Press.

McDermott, L. C., Rosenquist, M. L., & van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: examples from kinematics. *American Journal of Physics, 55*(6), 503-513.

Mehrens, W., & Lehmann, I. (1984). *Measurement and evaluation in education and psychology.* London: Holt Rinehart and Winston.

Menn, L., & Gleason, J. B. Babytalk as a stereotype and register: adult reports of children's speech patterns. In J.A.Fishman (Ed.), *The Fergusonian Impact* (Vol. I). Berlin: Mouton de Gruyter.

Michael, J. (1968). The reliability of a multiple choice examination under various test-taking instructions. *Journal of Educational Measurement, 5*, 307-314.

Minstrell, J. (1982). Explaining the 'at rest' condition of an object. *The Physics Teacher, 20*, 10-14.

Morris, P., & Gruneberg, M. (Eds.). (1992). *Aspects of memory, second edition. Volume 1: the practical aspects.* London and New York: Routledge.

Morse, J. M., & Richards, L. (2002). *Readme first for a user's guide to qualitative methods.* London: Sage Publications.

Mueller-Volmer, K. (Ed.). (1985). *The hermeneutics reader. Texts of the German tradition from the enlightenment to the present.* Oxford: Basil Blackwell Ltd.

Murray, J., Bradley, H., Craigie, W., & Onions, C. (Eds.). (1971). *The Compact Edition of the Oxford English Dictionary* (Vol. 1). Oxford: Oxford University Press.

Newburgh, R. (1998). What do centripetal acceleration, simple harmonic motion and the Larmor precession have in common? *Physics Education, 33*(2).

Norman, D. (1976). *Memory and attention: an introduction to human information processing.* New York: John Wiley and Sons.

Nussbaum, J., & Novick, S. (1992). Alternative frameworks, conceptual conflict and accommodation: toward a principled teaching strategy. *Instructional Science, 11*, 183-200.

Osborne, J. F. (1996). Beyond constructivism. *Science Education, 80*(1), 53-82.

Osborne, R., & Freyberg, P. (Eds.). (1985). *Learning in science: the implications of childrens science.* Auckland: Heinemann.

Palmer, D. (2001). Students' alternative conceptions and scientifically acceptable conceptions about gravity. *International Journal of Science Education, 23*(7), 691-706.

Palmer, D. H. (1997). Students' application of the concept of interdependence to the issue of preservation of species: Observations on the ability to generalize. *Journal of Research in Science Teaching, 34*(8), 837-850.

Palmer, R. (1969). *Hermeneutics: interpretation theory in Schleiermacher, Dilthey, Heidegger and Gadamer.* Evanston: Northwestern University Press.

Pankratius, W. J. (1990). Building an organized knowledge base: concept mapping and achievement in secondary school physics *Journal of Research in Science Teaching, 27*(4), 315-333.

Parkin, A. (1993). *Memory: phenomena, experiment and theory.* Oxford: Blackwell.

Peters, P. C. (1982). Even honors students have conceptual difficulties with physics. *American Journal of Physics, 50*(3), 501-508.

Pfundt, H., & Duit, R. (1994). *Bibliography: students' alternative frameworks and science education.* Kiel: IPN.

Piaget, J. (1970). *Genetic Epistemology.* New York: Columbia University Press.

Pines, A. L., & Leith, S. (1981). What is concept learning in science? Theory, recent research and some teaching suggestions. *The Australian Science Teachers Journal, 27*(3), 15-20.

Pollock, J. (1986). *Contemporary theories of knowledge.* London: Hutchinson.

Popper, K. R. (1969). *Conjectures and refutations: the growth of scientific knowledge* (3rd ed.). London: Routledge & Kegan Paul.

Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accomodation of a scientific conception.Toward a theory of conceptual change. *Science Education, 66*(66), 211-227.

Rippey, R. (1968). Probabilistic Testing. *Journal of Educational Measurement, 5*, 211-215.

Rippey, R. M. (1970). A comparison of five different scoring functions for confidence tests. *Journal of Educational Measurement, 7*(3), 165-170.

Robinson, M., & Davidson, G. (Eds.). (1996). *Chambers 21st century dictionary.* Edinburgh: Chambers.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic Press.

Rosch, E. (1978). Principles of Categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum Assocs.

Rosch, E., & Mervis, C. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology, 7*(7), 573-605.

Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 7*(7), 573-605.

Roth, W.-M. (1993). The concept map as a tool for the collaborative construction of knowledge: a microanalysis of high school phyiscs students. *Journal of Research in Science Teaching, 30*(5), 503-534.

Roth, W.-M. (1994). Science discourse through collaborative concept mapping: new perspectives for the teacher. *International Journal of Science Education, 16*(4), 437-455.

Rowell, J., Dawson, C., & Madsen, P. (1993). Probing students' non-scientific conceptions: a new tool for conventional and action-research in science teaching. *The Australian Science Teachers Journal, 39*(1), 62-68.

Roychoudry, A., & Roth, W.-M. (1996). Interactions in an open-inquiry physics laboratory. *International Journal of Science Education, 18*(4), 423-445.

Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching, 33*(6), 5[illegible]-600.

Sakharov, L. (1994/1930). Methods for investigating concepts. In R. van der Veer & J. Valsiner (Eds.), *The Vygotsky reader.* Oxford: Blackwell Publishers.

Saul, J. M. (1998). *Beyond problem solving: evaluating introductory physics courses through the hidden curriculum.* Unpublished Ph D, University of Maryland, Maryland.

Savage, L. J. (1971). Elicitation of personal probablilities and expectations. *Journal of the American Statistical Association, 66*(336), 783-801.

Schau, C., Mattern, N., Zeilik, M., Teague, K. W., & Weber, R. J. (2001). Select-and-fill-in concept map scores as a measure of students' connected understanding of science. *Educational and Psychological Measurement, 61*(1), 136-158.

Scott, P. H., Asoko, H. M., & Driver, R. H. (1992). Teaching for conceptual change A review of strategies. In R. Duit & F. Goldberg & H. Niederrer (Eds.), *Research in physics learning: theoretical issues and empirical studies. Proceedings of an international workshop held at the University of Bremen, March 4-8, 1991*. Kiel: IPN.
Searle, P. (1985). Circular motion concepts of first year engineering students. *Research in Science Education, 15*, 140-150.
Sell, P. (1985). *Expert systems - a practical introduction*. London: MacMillan.
Shuford, E., Albert, A. & Massengill, H. (1966). Admissible probability measurement procedures. *Psychometrika, 31*, 125-145.
SIL. (2001). Speech Analyzer (Version 1.5). Waxhaw: Summer Institute of Linguistics.
Sinclair, J. (1985). Lexicographic Evidence. In R. Ilson (Ed.), *Dictionaries, lexicography and language learning* (pp. 81-94). Oxford: Pergamon Press.
Sinclair, J. (Ed.). (1987). *Collins COBUILD English language dictionary*. London and Glasgow: Collins.
Sinclair, J., & Coulthard, R. (1974). *Towards an analysis of discourse: the English used by teachers and pupils*. London: Oxford University Press.
Smith, M. (1991). Put to the test: the effects of external testing on teachers. *Educational Research, 20*(5), 8-11.
Smith, P. A. (1992). Let's get rid of "centripetal force". *The Physics Teacher, 30*(316-317).
Snow. (1964). *The two cultures and a second look: an expanded version of the two cultures and the scientific revolution*. Cambridge: Cambridge University Press.
Solomon, J. (1994). The rise and fall of constructivism. *Studies in Science Education, 23*, 1-19.
Sowa, J. (1984). *Conceptual structures information processing in mind and machine*. Reading, MA: Addison Wesley.
Spearritt, D. (1988). Factor analysis. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: an international handbook* (pp. 644-657). Oxford: Pergamon Press.
Stanford. (1998). *An introduction to the philosophy of history*. Oxford: Blackwell Publishers Ltd.
StatSoft, I. (1999). *Electronic Statistics Textbook*. [WEB: http://www.statsoft.com/textbook/stathome.html]. Tulsa, OK: StatSoft. Retrieved, from the World Wide Web:
Stavy, R., & Berkovitz, B. (1980). Cognitive conflict as a basis for teaching quantitative aspects of the concept of temperature. *Science Education, 64*(5), 679-672.
Stavy, R., & Tirosh, D. (1996). Intuitive rules in science and mathematics: the case of 'more of A - more of B'. *International Journal of Science Education, 18*(6), 653-667.
Stavy, R., & Tirosh, D. (2000). *How students (mis-)understand science and mathematics: intuitive rules*. New York and London: Teachers College Press.
Stead, K., & Osborne, R. (1981). What is friction? - some children's ideas. *The Australian Science Teachers Journal, 27*(3), 51-57.
Steffe, L., & Gale, J. (Eds.). (1995). *Constructivism in education*. Hillsdale, NJ: Lawrence Erlbaum.
Stine, E. L., & Bohannon III, J. N. (1983). Imitations,interactions,and language acquisition. *Journal of Child Language, 10*, 589 -603.
Stone, S. J., & Glascott, K. (1997). The affective side of science instruction. *Childhood Education, 74*(2), 102-104.
Styer, D. F. (2001). Letter to the Editor: The word 'force'. *American Journal of Physics, 69*(6), 631-632.

Taber, K. S. (2001). Shifting sands: a case study of conceptual development as competition between alternative conceptions. *International Journal of Science Education, 23*(71), 731-753.

Talmy, L. (1976). Semantic causative types. In M. Shibatani (Ed.), *Syntax and Semantics, vol. 6*. New York: Academic Press.

Talmy, L. (1985). Force dynamics as a generalisation over 'causative'. 67-85.

Tao, P. K., & Gunstone, R. F. (1999). The process of conceptual change in force and motion during computer-supported physics instruction. *Journal of Research in Science Teaching, 36*(7), 859-882.

Tasker, R., & Freyburg, P. (1985). Facing the Mismatches in the Classroom. In R. Osborne & P. Freyberg (Eds.), *Learning in science: the implications of childrens science*. Auckland: Heinemann.

Taylor, C. W., & Gardner, P. (1999). An alternative method of answering and scoring multiple choice tests. *Research in Science Education, 29*(3), 353-363.

Tennyson, R. D., Chao, J. N., & Youngers, J. (1981). Concept learning effectiveness using prototype and skill development presentation forms. *Journal of Educational Psychology, 73*(3), 326-334.

Terry, C., & Jones, G. (1986). Alternative frameworks: Newton's third law and conceptual change. *European Journal of Science Education, 8*(3), 291-298.

Thornton, R. K., & Sokoloff, D. R. (1998). Assessing student learning of Newton's laws: the Force and Motion Conceptual Evaluation and the evaluation of active learning laboratory and lecture curricula. *American Journal of Physics, 66*(4), 338-352.

Tornkvist, S., Petterson, K.-A., & Transtromer, G. (1993). Confusion by representation: on student's comprehension of the electric field concept. *American Journal of Physics, 61*(4), 335-338.

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.

Twigger, D., Byard, M., Driver, R., Draper, S., Hartley, R., Hennessy, S., Mohamed, R., O'Malley, C., O'Shea, T., & Scanlon, E. (1994). The conception of force and motion of students aged between 10 and 15 years: an interview study designed to guide instruction. *International Journal of Science Education, 16*(2), 215-229.

Tytler, R. (1998). Children's conceptions of air pressure: exploring the nature of conceptual change. *International Journal of Science Education, 20*(8), 929-958.

Unsworth, L. (2001). Evaluating the language of different types of explanations in junior high school science texts. *International Journal of Science Education, 23*(6), 585-609.

Viennot, L. (1979). Spontaneous reasoning in elementary dynamics. *European Journal of Science Education, 1*(1), 205-221.

Viennot, L. (1985). Analyzing students reasoning.Tendencies in interpretation. *American Journal of Physics, 53*(5), 432-436.

von Glasersfeld, E. (1989). Constructivism in education. In T. Husén & T. Posthelwaite (Eds.), *The international encyclopaedia of education: research and supplementary volume 2*. (pp. 162-163). Oxford: Pergamon Press.

von Glasersfeld, E. (1992). A constructivist's view of learning and teaching. In R. Duit & F. Goldberg & H. Niederrer (Eds.), *Research in physics learning: theoretical issues*

and empirical studies. Proceedings of an international workshop held at the University of Bremen, March 4-8, (pp. 29-39). Kiel: IPN.
von Glasersfeld, E. (1995). A constructivist approach to teaching. In L. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 3-15). Hillsdale, NJ: Lawrence Erlbaum.
von Heusinger, K. (1999). *Intonation and Information Structure.* Unpublished Habilitationsschrift, University of Konstanz, Konstanz.
von Humboldt, W. (1985/1836). The nature and conformation of language. In K. Mueller-Volmer (Ed.), *The hermeneutics reader. Texts of the German tradition from the enlightenment to the present.* Oxford: Basil Blackwell Ltd.
Vygotsky, L. (1986). *Thought and language.* Cambridge MA: The MIT Press.
Vygotsky, L. (1994/1931). Thinking and concept formation in adolescence. In R. van der Veer & J. Valsiner (Eds.), *The Vygotsky reader.* Oxford: Blackwood Publishers.
Warren, J. (1979). *Understanding force.* London: John Murray.
Watson, J. R., Prieto, T., & Dillon, J. S. (1997). Consistency of students' explanations about combustion. *Science Education, 81*(4), 425-443.
Watts, D. M., & Zylbersztajn, A. (1981). A survey of some children's ideas about force. *Physics Education, 16*, 360-365.
Way, E. (1991). *Knowledge representation and metaphor.* Dordrecht: Kluwer Academic Publishers.
White, K. G., Alsop, B., & Williams, L. (1993). Prototype identification and categorization of incomplete figures by pigeons. *Behavioural Processes, 30*(3), 253-258.
White, R. T. (1988a). *Learning science.* Oxford: Blackwell.
White, R. T. (1988b). Metacognition. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: an international handbook* (pp. 70-75). Oxford: Pergamon Press.
White, R. T. (1991). An overview of the Australasian perspective. In J. Northfield & D. Symington (Eds.), *Learning in science viewed as personal construction: an Australian perspective.* Perth: Curtin University of Technology.
White, R. T. (1997). Trends in Research in Science Education. *Research in Science Education, 27*(2), 215-221.
White, R. T., & Gunstone, R. F. (1992). *Probing understanding.* London: The Falmer Press.
Wierzbicka, A. (1985). *Lexicography and conceptual analysis.* Ann Arbor: Karoma Publishers.
Wierzbicka, A. (1996). *Semantics: primes and universals.* London: Oxford University Press.
Wierzbicka, A. (1998). The semantics of English causative constructions in a universal-typological perspective. In M. Tomasello (Ed.), *The new psychology of language: cognitive and functional approaches to language structure.* London: Routledge.
Williams, H. T. (1999). Semantics in teaching introductory physics. *American Journal of Physics, 67*(8), 670-680.
Williams, K. (2000). Inexpensive Demonstrator of Newton's First Law. *Physics Teacher, 38*(2), 80.
Winchester, S. (1998). *The surgeon of Crowthorne: a tale of murder, madness and the love of words.* London: Viking.
Wittgenstein, L. (1953). *Philosophical investigations.* Cambridge, UK: Cambridge University Press.
Woolgar, S. (1988). *Science: the very idea.* Chichester: Ellis Horwood.

Wrobel, S. (1994). *Concept formation and knowledge revision.* Dordrecht ; Boston: Kluwer Academic Publishers.
Yates, F. (1966). *The art of memory.* London: Routledge and Kegan Paul.

INDEX