ELICITING AND UNDERSTANDING COMMONSENSE REASONING ABOUT MOTION

by

Nancy Wai Ying LAW

A Thesis submitted in fulfilment of
the requirements for the degree of
Doctor of Philosophy

in

University of London Institute of Education

September 1990

ABSTRACT

The focus of the present research is on children's commonsense reasoning in mechanics.

The important effect of pre-instructional ideas on children's learning is now widely recognised and much effort has gone into investigating what these ideas are like in various domain areas in science in the past few years. Early researches in this area have provided us with a comprehensive catalog of phenomenological descriptions of various aspects of children's reasoning about forces and motion. A related line of research has grown over recent years, which attempts to probe into whether there are deeper explanations underlying these misconceptions. If we take scientific theories and commonsense reasoning as two ends of a dichotomy, then early researches in this field have predominantly started from the scientific end, looking towards the intuitive end, trying to find out where the intuitive ideas go astray. To look for deeper levels of analysis, some have since turned to looking from the opposite end, trying to take children's ideas seriously, in their own right and not as a distortion of the scientific view. This latter perspective is the one taken by the present research and is believed to be appropriate if an understanding of the phenomenological descriptions of children's intuitive ideas is to be attained.

The present research sets out to investigate the possible cognitive models used in the spontaneous interpretation of and reasoning about motion by students with varying amounts of Physics instruction. It is hoped that the resulting models will not only provide a context for interpreting children's misconceptions, but also provide insight into the evolution of naive cognitive models to more scientific ones.

The research consists of two tasks. The first is a classification task asking students to categorize comic strip pictures about motion and to explain their underlying reasoning. The second is a programming task, asking students to write expert systems about motion in the language PROLOG. The second task is in fact one of self elicitation of knowledge by the students themselves under the assistance of the researcher. The advantage of such an exercise is that the representation is not only open for inspection by the students but is also explorable. The results from both tasks will be analysed and synthesized in the thesis.

ACKNOWLEDGEMENT

It is with feelings of excitement and relief that I embark on writing this last addition to my thesis. The five years I have spent on this research have been a great learning experience for me and I am still amazed that I can keep my interest and spirits up till this day. I am deeply indebted to all those who have helped me in various ways to make this possible.

I am especially grateful to Prof. Jon Ogborn, my supervisor, who has taken on the extra pains of supervising a student who works on the thesis part time and at a distance for most of the time. During my short and infrequent visits to London, Jon has never failed to oblige me with intensive meetings and discussions. I am deeply appreciative of his encouragement, criticisms and suggestions throughout.

I would like to thank Mr. Jonathan Briggs for discussions on technical aspects of the research during the initial stages and for providing me with the source codes of MITSI. I am also indebted to Mr. C.S. Lee for customizing MITSI for use in the research.

My thanks also goes to Dr. Joan Bliss for her comments on the research design and to Mr. W.W. Ki for comments and discussions during various stages of the work.

I owe a special debt to all the students who have participated so cheerfully and cooperatively in the various tasks of this research.

I am grateful to Miss Choice Tse, my research assistant for helping with the translation and transcription of all the audiotape recordings made during the programming sessions. I am also greatly indebted to Mr. William Pang for drawing all the figures and to Mr. Stanton Kok for helping with the final revisions on the word processor.

Special thanks goes to Ms. L.K. Tse and Mrs. Cecilia Shek who read through the final drafts and helped to eliminate many of the typographic errors and generally improved the readability of the script.

I am indebted to the University of Hong Kong for providing a research grant to support this work.

When I first talked about the idea of doing a PhD with my husband, Hing Chung, I mentioned it only as a fantasy - something I might be able to do if there were no family and no children. I was totally surprised when he asked me what might the obstacles be for such a wish to come true. Ever since, he has been striving with me to pursue this dream. This work would not have been possible without his unfailing support, concern and encouragement.

CONTENTS

CHAPTER 3 LEARNING, KNOWLEDGE AND REPRESENTATION

CHAPTER 4 METHODS OF ENQUIRY

CHAPTER 5 THE CLASSIFICATION TASK

CHAPTER 6 THE PROGRAMMING TASK

CHAPTER 7 ANALYSIS OF THE PROGRAMMING TASK RESULTS

CHAPTER 8 TESTING AND MODIFICATION OF PROGRAM

LIST OF TABLES

LIST OF FIGURES

Chapter 1 INTRODUCTION

Now is it clear merely in the light of common sense that a body in the absence of any force acting on it moves perpetually in a straight line with constant speed? Or that a body subject to a constant weight constantly accelerates the velocity of its fall? On the contrary, such opinions are remarkably far from common sense knowledge; in order to give birth to them, it has taken the accumulated efforts of all the geniuses who for two thousand years have dealt with dynamics.

Duhem (1954), 263

1.1 The Choice of Problem

Science as the culmination of the knowledge man has gained in his endeavours to explore and understand the world is undoubtedly, even according to common wisdom, superior to commonsense. Deciding on the desirable qualities of science graduates, the sort of skills and abilities that have to be cultivated, has always been a major problem in science education. The fact that students hold pre-instructional ideas about the physical world around us has not been a cause of concern in science education until very recently. Commonsense ideas were assumed to be either addressing different issues or to wither away the moment they meet school science. This intuitive assumption is possibly the root of the so called "tabula rasa" approach in teaching.

Research findings pointing to the fact that a lot of students, including science undergraduates, still hold intuitive conceptions that are contrary to the scientific view brought about much surprise and interest. The first investigations started as pencil-and-paper tests of "misconceptions" or spontaneous conceptions, initially mainly in the area of mechanics. Growing interest in the phenomenon led to explorations of students' ideas in all the major areas of science and to the development of different techniques in the exploration of such ideas.

Early explorations in the area of children's ideas focussed on getting phenomenological descriptions of these ideas, especially from the perspective of where they departed from the scientific view. Results revealed largely common intuitive ideas being held by students across a wide age range and across different countries. Such results coupled with a general background of increasing reference to the constructive role of the learner in the learning process led many to believe that the intuitive ideas have deep roots in people's everyday interactions with physical phenomena. A related line of research has grown over recent years which attempts to probe into whether there are deeper explanations underlying children's "misconceptions" in science. Here, researchers try to explore students' ideas in their own right and not as a distortion of the scientific view, and they hope to find deeper structures of cognition/ understanding that can give a more ubiquitous explanation for the various misconceptions recorded. The present research is an attempt to explore students' reasoning about motion in this same general direction.

There are philosophical explorations of the differences between science and common sense (e.g. Nagel, 1961) as well as explorations of the relationship between science and common sense from a sociological perspective (e.g Berger & Luckmann, 1966). However, there is little in the literature about the relationship between these two kinds of knowledge at the level of the individual. Do a person's commonsense ideas about a specific domain form an integrated system, that is, a theory? How consistently would the same commonsense ideas be used under different circumstances? What roles do these ideas play when a person comes into contact with scientific ideas in the context of school science learning? Would there be attempts to integrate them or would the two just remain in co-existence, each coming into play under different circumstances? Under what circumstances would a person give up his/her intuitive conceptions? These are interesting questions that have important educational implications.

In attempting to answer such questions, the present research tries to elicit and understand students' commonsense reasoning in a particular domain area - the physical world of motion. In other words, this is an exploration into students' intuitive mechanics. The world of motion is chosen to be the domain of exploration for two reasons: Everyone has to cope with motion phenomena from birth and it is inconceivable for anyone not to have developed intuitive ideas about it. At the same time, this is also the area with the largest amount of empirical data accumulated about it.

1.2 Methodological Issues

In defining the objectives of the elicitation exercise, the present work attempts to expose not just the individual ideas in the commonsense conception of motion, but also the structural relationships between them. This is derived from the belief that the actual organization of knowledge is at least as important as the elements found in it. Furthermore, as the focus is on the actual conceptualizations used in commonsense reasoning, it is felt that an exploration of the knowledge structures in declarative format is more appropriate. Elicitations in problem solving contexts bring in problems related to meta-cognitive skills and mastery of knowledge not strictly within the subject domain, as well as the disadvantage of a tendency for such task contexts to cue school science knowledge. However, the methods of elicitation commonly employed in this area of research that do yield information about declarative structuring of knowledge, for example concept map or line labelling tasks, tend to use school science concepts/terms as starting points. What is looked for here, instead, is an open task that can elicit in a systematic manner students' declarative ideas without pre-defining a specific context or a specific set of ideas to work from.

Another desirable feature of the elicitation task here is that it should probe into the more stable and rational aspects of students' commonsense reasoning. Elicitation strategies developed in this area of research, for example written tests, interview about instances, interview about events, repertory grids, etc., have generally been geared towards probing the spontaneous responses of students. The disadvantages of such methods are that it is not clear how strong a subject's commitments are towards the various ideas elicited, and it is not easy to find out how the different spontaneous ideas orchestrate together for different problem contexts.

In the present research it was decided to employ a rather novel elicitation technique of expert system development by students, supplemented by a more conventional classification task. The programming task requires students to write expert systems about motion in the computer language PROLOG. This is an extended task where each student is required to work for six to eight weekly sessions in the production of an expert system capable of answering general everyday questions about common motion events. During the process of program development, the students have the chance to explore the reasoning ability of the expert systems built by running them and also to modify the programs as they see fit. This task is essentially one of self-elicitation of knowledge by the students themselves under the assistance of the researcher. It has the advantage of producing a declarative representation of the students' reasoning about motion that is open for inspection by the researcher and students alike, as well as being explorable. The testing and modification behaviour of the students when programs fail to yield expected answers provide important insights that are normally otherwise inaccessible: insights into the students' degree of commitment to various ideas and possible learning mechanisms when existing knowledge structures fail.

In as much as a novel technique may bring on new developments, there is a danger that the results of the exploration are just artifacts of the research technique itself. It is thus desirable that the results of the research can be checked against those elicited by more conventional methods. It was decided that a classification task asking students to put comic strip pictures of motion into groups according to some criteria and then to explain how the classification was done should be performed before the start of the programming task. The criteria used by students for putting motion events into groups should yield important perspectives for reasoning about motion events in everyday contexts. As mentioned, major ideas elicited from this task can be used as starting points for the students to elaborate on in their programming task. More importantly, conceptual frameworks used in the intuitive processing of motion events as elicited from analysis of the classification protocols would provide useful reference points for comparison with results from the programming task.

Two groups of subjects participated in this research: a group of F.3 students (about 15 years old) who had not received any instructions in mechanics, and a group of F.6 science students (about 17 years old) who had gained credits or distinctions in Physics in the HKCEE examination (equivalent to the GCE O-Level examinations) and were doing A-Level Physics. It is hoped that comparison of the performance of these two groups of students would shed light on the effect of school physics instruction on the conceptions of motion elicited. Students from different ability ranges were also represented in each age group so as to find out if there is any indication of different performance for the different ability groups.

1.3 What This Research Hopes to Achieve

It is hoped that this research can yield fruit on two fronts: a better understanding of commonsense reasoning about motion and an assessment of the viability of employing expert system development as a means of knowledge elicitation and cognitive modelling.

In terms of our understanding of commonsense reasoning about motion, it is anticipated that this research should yield information on the parameters and perspectives that figure importantly in the everyday processing of motion events as well as revealing the existence, if any, of prominent patterns of reasoning about motion in the commonsense context. It is also hoped that the research findings can provide insights into the structural relationship of the different ideas elicited and thereby address the question of whether the commonsense ideas elicited can be termed "theories", and if so, in what sense. The modification behaviour of students in the programming task may, in addition, provide information on possible differences in the strength of commitment to the different ideas elicited. Comparison of results from the two age groups should also provide clues to the effects of school physics instruction in relation to commonsense ideas.

The explorability and modifiability of the elicited ideas in the form of PROLOG programs allows for a close observation of the possible cognitive responses and outcomes in the face of conflicts, in this case a failure of the developed programs to yield expected results. It has been claimed by many (e.g Hewson, 1981, 1984) that cognitive conflict leads to conceptual restructuring and thus learning. The programming task provides a chance, though limited in time duration and scope, to look into the cognitive strategies employed in the face of conflict and the possible modes of learning under such circumstances. Here, results from the two age groups, each with students from different ability ranges, can be compared to look at the possible effects of physics instruction and general ability on the different issues explored.

The present research is novel in trying to develop a new method of knowledge elicitation. It may be breaking new ground in bringing artificial intelligence tools to the service of cognitive exploration. At the same time the results of the exploration must be given careful scrutiny and their validity assessed. What does expert system development actually offer as a means of elicitation? Are the results credible in comparison with other more conventional methods of exploration? PROLOG forces the elicited ideas to be represented as a logical system while commonsense ideas are generally believed to be fuzzy, fragmented and not necessarily consistent. Would it be possible that the results of the exploration are just products of such an "artificial" environment? Are there difficulties related to the administration of this method and what are its limitations? Hopefully these questions can be addressed in the light of the results.

1.4 Organization of the Thesis

Chapter two is a review of research in the area of children's pre-instructional ideas in science. The emphasis is on the philosophical and sociological background to the strong interest in such ideas and on the theoretical frameworks which underlie the different lines of approach to the problem.

Chapter three is a review of theoretical issues in two areas: learning and knowledge representation. Views on learning contribute largely to the formulation of the research questions and to the associated theoretical framework. Knowledge representation follows both as a theoretical issue closely related to learning, as well as a methodological one of choosing appropriate representation formalisms for the results.

Chapter four gives an overview of the different methodologies commonly used in the exploration of learners' conceptions. Again the emphasis here is on the theoretical underpinnings of the various methods. Their relative merits are compared with reference to their aims and perspectives. The actual tasks used in the research are described separately in the next two chapters.

Chapter five is a self contained unit by itself, describing the rationale for choosing a classification task as the method of exploration, the actual administrative details as well as presenting the results and their analysis.

Chapter six describes the programming task in detail. First, expert systems and the features of the actual expert system shells used are described. The results of a pilot study on the viability of such a task are then reported, followed by a description of the actual administrative details of the task in the main study, including procedures and formats for the data collection and data reduction stages.

Chapter seven presents an analysis of the knowledge structures for commonsense reasoning about motion as revealed by the written expert systems. The programs are analyzed both at the individual rule level as well as the program level, to look into the actual conceptions of motion they represent as well as the structural features of those ideas.

Chapter eight analyzes the testing and modification behaviour of the students and tries to look at the cognitive strategies employed in response to conflicting situations. The issue of what possible modes of learning may be represented by these behaviour is also addressed.

Chapter nine attempts to synthesize the results from the two tasks and tries to answer the questions set out in this chapter.

Chapter 2 REVIEW OF RESEARCH ON CHILDREN'S IDEAS IN SCIENCE

The worst evil resulting from the precocious use of speech by young children is that we not only fail to understand the first words they use, we misunderstand them without knowing it; so that while they seem to answer us correctly, they fail to understand us and we them... This lack of attention on our part to the real meaning which words have for children seems to me the cause of their earliest misconceptions; and these misconceptions, even when corrected, colour their whole course of thought for the rest of their life.

Emile. J.J. Rousseau, 1760.

2.1 The Problem: Phenomenon of Children's Ideas in Science

2.1.1 Extent of the problem.

The concern for learners' pre-instructional ideas has inspired a whole area of research in science education. This trend of work started in the early '70s, gathered momentum in the '80s and is still a major area of research, broadening into subject areas other than science. It began when some researchers found to their surprise that students had not learnt what they were thought to have mastered (Doran, 1972; Duncan & Johnstone, 1973) with sometimes very low percentages of correct responses being recorded for quite simple problems. One early piece of research that made an important impact was that done by Viennot (1977, 1979). She gave to students varying from last year secondary to third year University a paper-and-pencil test on some qualitative mechanics problems involving predictions of specific aspects of the motion of bodies. The test was given to students in France, Britain and Belgium. The result was remarkable in that despite the wide range of students tested, in general not more than half of the students could give the correct answers. Even more surprising was that there was no substantial difference in the results from the three different countries, nor was there any great improvement with the amount of formal education received. The results pointed towards the existence of widespread, common intuitive 'laws of Physics' which are different from the accepted scientific view and which are extremely robust and resistant to change.

2.1.2 Characteristics of children's ideas.

Alternatives to paper and pencil tests were developed to find out more about these intuitive ideas of natural phenomena. The following are generally agreed to be important phenomenological features of such pre-instructional ideas (Osborne & Freyberg, 1985; Confrey, 1987) :

1. The alternative conceptions held by students are very often related to their informal knowledge and the way words in science have been used in everyday language. From a young age, and prior to any teaching and learning of formal science, children develop meanings for many words used in science teaching and views of the world which relate to ideas taught in science.

2. These ideas are usually strongly held, and resistant to change. They may be significantly different from views of scientists, yet they are psychologically compelling, and may still be held intact by children and adults alike after having completed years of formal science instructions.

3. These ideas are sensible and rational from the children's point of view.

4. Although there is variety in the ideas children use to interpret phenomena, there are some general patterns in the types of ideas that children of different ages tend to use, and quite independent research studies have reported similar patterns of ideas held. This commonality found across large collections of individual constructs probably is one important driving incentive for research in this area - for if all we can say about individual constructs is that they are different and there is no predictable trend, then there is not much that we can do about them as science educators.

There is one important divergence of opinion concerning the self-consistency of children's ideas. While some (e.g. Driver et. al., 1985) saw them as incoherent, others (e.g. Osborne & Freyberg, 1985; Viennot, 1979) referred to them as sensible, coherent, and self-consistent. The lack of coherence is apparent if viewed from a scientific perspective, but how far such views are consistent within the children's own frame, or how far children are committed to logical consistency of their own conceptions is still a problematic issue.

Differences are even more striking when one tries to get an overview of the various pieces of work within this area of research. As Gilbert & Watts (1983) pointed out, there is no general agreement on the aims of enquiry, the methods to be used, criteria for appraising data and the use to be made of the outcomes. Furthermore, different epistemological and ontological status is ascribed to the same phenomenon by different researchers, resulting in a diversity of descriptors: 'misconceptions' (e.g. Helm, 1980), 'preconceptions' (e.g. Novak, 1977), 'alternative conceptions' (Driver and Easley, 1978), and 'children's science' (Gilbert, Osborne & Fensham, 1982). A lot of ground has been covered since then, but the lack of agreement on many of these critical issues is still characteristic of the present state of this area of work.

2.1.3 The review.

It is not the intention here to give a comprehensive review of research in this area. There are books devoted entirely to the discussion of extensive collections of research findings in the field (e.g. Driver, Guesne, Tiberghien, 1985; Osborne & Freyberg, 1985; West & Pines, 1985), excellent review papers (e.g Driver & Erickson, 1983; Gilbert & Watts, 1983; Duit, 1987; Confrey 1987) and bibliographies (e.g. Pfundt & Duit, 1987; Giordan, 1987). This review attempts to trace the main questions which seem to lie behind the conception of research in the field. The work on developing more effective methods of teaching that takes children's preconceptions into account has been deliberately left out of this review. Though the purpose of understanding children's ideas is ultimately to teach better, the focus of the present research is to learn more about the nature of these prior ideas.

Mechanics, possibly more than any other branch of science, is a subject which everyone has an awareness of from birth. It is inconceivable for someone not to have some kind of beliefs about motion. This subject has received more attention than any other branch of science from researchers all over the world. Mechanics has been chosen as the subject of the present research because a rich accumulation of phenomenological observations of children's ideas in this domain already exists.

The present review will start with a description of the main characteristics of children's ideas in mechanics. It will then try to locate this movement in science education research within the wider context of changing conceptions in philosophy of science as well as psychology. This will be followed by a brief overview of the variety of questions tackled in this area of research and the different theoretical frameworks embodied, which will show how the present research is to be seen in relation to other work.

2.2 Children's ideas in mechanics

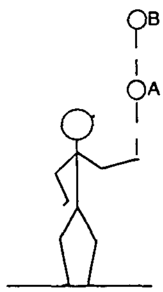
There are several very comprehensive reviews of work done in the area of children's ideas in mechanics: McDermott (1984), Gilbert and Watts (1983), Gunstone and Watts (1985) and Jagger (1985). A wide variety of research strategies were employed in the studies reported, varying from pencil and paper tests to in-depth interviews and laboratory tasks. Even so, a very similar pattern of understanding was found. The intuitive conceptions reported were found to fall into five main areas.

2.2.1 Force and motion

This is by far the most important area of difficulty in the learning of Newtonian mechanics in terms of both prevalence and persistence. There are several strongly held beliefs under this theme.

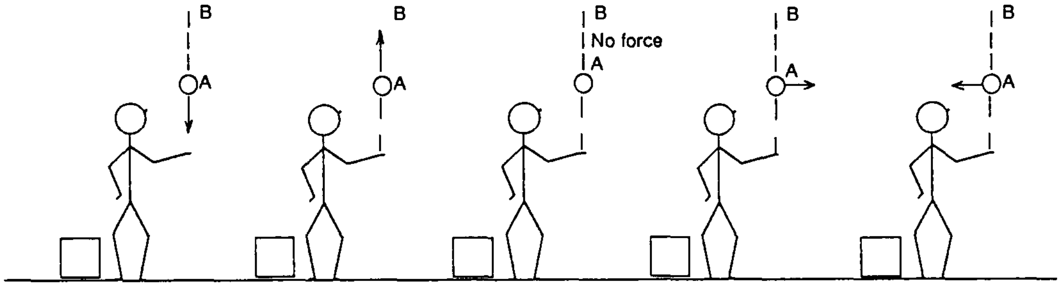
i) *Force is proportional to speed: the harder you push something the faster it moves.*
ii) *Force on an object acts in the direction of motion.*
iii) *There is no (net) force on an object at rest.*
iv) *Motion implies a force.*

The above four beliefs very often co-exist together and pervade nearly all the research into the learning of mechanics. These beliefs have emerged from people all over the world in all age groups and often in spite of intensive instruction, (Sjoberg and Lie (1981) in Norway, Watts and Zylbersztajn (1981) in Britain, Viennot (1979) in France). They can easily be elicited by variations of the tossed stone problem (fig. 2.1), or the oscillating weight on a spring problem (fig. 2.2).



A stone is thrown straight up in the air.
It leaves the person's hand, goes up through point A, gets as high as B and then comes back down through A again.

The arrows in the pictures are supposed to show the *direction of the force* on the stone.
Which picture do you think best shows the force on the stone on *its way up* through A ?



Explanation : ________________

Fig. 2.1 Question designed to elicit children's ideas about force and motion direction. (Watts & Zylbersztajn, 1981)

Are the forces acting on all the masses identical at the instants shown ?

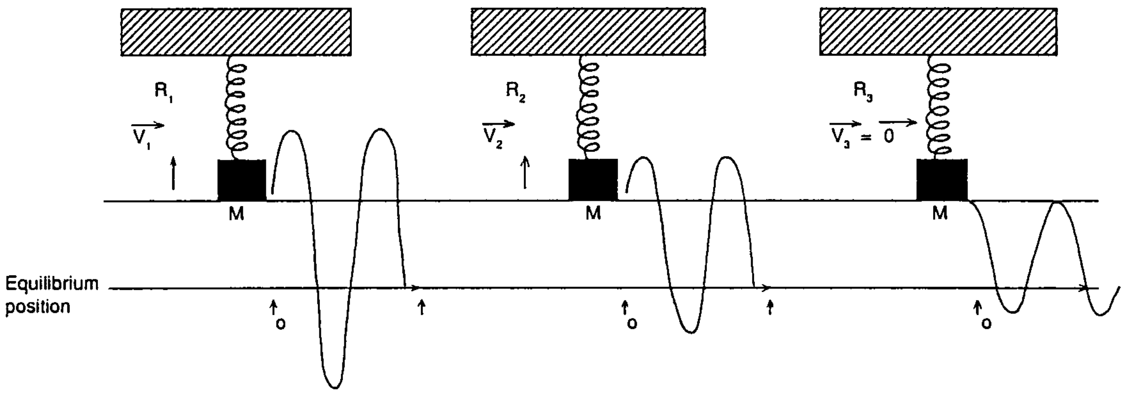


Fig. 2.2 Questions designed to elicit ideas about force, motion speed and motion direction. (Viennot, 1979)

In the first case, the Newtonian forces on the moving object at the various moments depicted are identical while in the second case the forces only depend on the position of the object. However, the provision of the trajectory information seemed to be a strong distractor and the students found it very difficult to explain why a particle could be moving upwards without having an upward force. Similarly, an object coming to a stop momentarily during the course of motion must have no net forces acting on it at the point when it was at rest. Changes in the speed of the object tend to be explained by a corresponding change in force. For example, a common explanation for a stone to slow down on going upwards after being tossed was that the upward force on it decreases as it goes up. Following a similar line of reasoning, an oscillating object on passing through its equilibrium position was very often thought to be experiencing the largest force. It was quite inconceivable to many that a moving object could have no force acting on it.

v)Constant motion requires a constant force.

This is another very prevalent intuitive belief found to be held by many, from children to in-service teachers (Lawson, 1984). This result has been consistently elicited by various methods: for example by interviews (Gilbert, Watts & Osborne, 1985), pencil-and-paper tests (Sjoberg & Lie, 1981), laboratory tasks and computer simulations (Langford and Zollman, 1982). The laboratory task method of elicitation was especially illuminating. Lawson (1984) asked students to make a relatively massive, almost friction-free puck move in a straight line with steady speed by blowing it with air from a hose. The students could try one or more of four procedures: constant blast, steadily decreasing blast, a series of short blasts or a single short blast. Only half of the students did it correctly at the first trial, though many corrected it on the second trial. However, there were some who persisted with blowing the puck with a constant force in the direction of the motion, and could not understand why they were not successful.

2.2.2 Force and Animacy

Many students tend to associate the ability to exert force with objects that are alive, and found it very difficult to conceptualize 'passive forces', for example, reaction forces of tables or chairs on other objects that exert forces on them. This difficulty has been reported by Driver (1973) and Minstrell (1982).

2.2.3 Force and Momentum

The words force and momentum have been found to be used interchangeably by children, and Sutton (1980) has pointed out that the current scientific meaning for the word 'force' was only adopted sometime in the nineteenth century. This difficulty with terms can also be related to a belief similar to the medieval 'impetus theory' of motion: An object will continue moving in its original manner until the initial 'force' that set it in motion is used up (Clement, 1982).

Another very interesting finding (McCloskey, Caramazza & Green, 1980) was that many students believed that an object following a curved path somehow 'remembers' the curvature of the track and continues to move in curvilinear motion after emerging from a curved tube or after a string holding it is being cut. (fig. 2.3) It was as if the students believed that the moving object possessed some kind of 'curvilinear momentum'.

These four line drawings were given on an examination. Three of the drawings represent horizontal curved tubes; the fourth is of a ball attached to a string and whirled in a circle overhead.

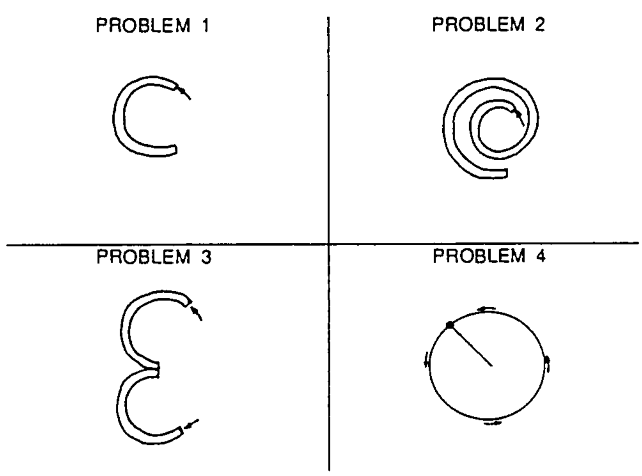


Fig.2.3 Question on curvilinear motion: Trace the trajectory the ball would follow after leaving tube (1-3), or after string is cut (4). (Caramazza, McCloskey & Green, 1981)

2.2.4 Kinematics

Very often, the difficulties with dynamical concepts are compounded by difficulties with kinematics. Piaget (1970 a, b) was probably the first to look into children's conceptions of time, distance and speed. He found that children often confuse the notions of speed and displacement, and the concept of speed as a coordination of space with time was not reached by most children until quite late in their development. Furthermore, speed is a more fundamental concept than time for young children. Trowbridge and McDermott (1980, 1981) carried out a detailed project to investigate university students' understanding of the concepts 'speed' and 'acceleration'. They carried out individual demonstration interviews to assess students' abilities to interpret simple motions of real objects. The principal conceptual difficulty with speed was the inability to discriminate between displacement and speed, and many thought that two objects with the same displacement had to have the same speed, at least at that instant. Problems with acceleration were much more numerous, including confusion between position and acceleration, between velocity and acceleration, between velocity and change of velocity. Many students failed to associate a particular instantaneous velocity with a particular instant, and they tended to neglect the time intervals during which changes in velocity took place.

2.2.5 Gravity

There is confusion over the exact meaning of the word 'gravity'. Jagger (1985) pointed out that inconsistencies can be found even in a dictionary as respectable as the Concise Oxford Dictionary. There gravity was defined first as 'the acceleration due to the gravitational field' which should then be the same for all objects irrespective of mass, and secondly as 'the force resulting from this acceleration' which should then depend on the mass of the body concerned. Similar confusions can be found in students (Gunstone and White, 1981). A number of other investigations (Watts, 1982; Champagne, Klopfer & Anderson, 1980; Stead & Osborne, 1981) into children's reasoning about falling objects disclosed the following prominent intuitive beliefs about gravitation and free fall:

i) *Gravity needs a medium and there is no gravity in a vacuum,*
ii) *Things are heavier (or lighter) the higher up they are,*
(iii) *Heavier objects fall faster.*

The above list of common intuitive conceptions/misconceptions in the area of mechanics, though by no means exhaustive, represent the most dominant phenomenological patterns found. Any attempt to produce an adequate theory of how such reasonings may come about needs to take them into consideration.

2.3 Wider contexts - psychological and philosophical roots of the concern

The increase in attention paid to pupils' wrong ideas by researchers in science education was accompanied by a parallel change in attitude towards such ideas: there was an apparent 'respect' or 'value' attached to children's ideas per se. At least two factors contributed to the provision of an intellectual climate for this change: changes in philosophy of science and in psychology.

In psychology, behaviourism, which emphasized the passivity of the mind, lost its hold on cognitive theorists after the war. Piaget (e.g.1954, 1970c), in his work on genetic epistemology, provided ample evidence that children were actively trying to make sense of the world around them, actively constructing their own understanding of the surroundings, right from birth. The processes of assimilation and accommodation are the processes leading to cognitive development in the individual due to varying degrees of dissonance between physical reality and their own constructions. Piaget's life-long ambition was "to explain knowledge, and in particular scientific knowledge, on the basis of its history, its sociogenesis, and especially the psychological origins of the notions and operations upon which it is based"(1970c). As such, he was concerned with the search for fundamental logico-mathematical structures that were supposed to provide the cognitive basis for scientific development.

Kelly, in his Personal Construct Theory (1955), used the metaphor of person-as-scientist and suggested that we are all scientists of a sort from a very young age in that we create or invent our own constructs to understand the world.

This shift of attention to the active processing of sensory input and the person's active construction of his/her own understanding, as well as a search for deep cognitive structures that may form the basis of our cognition, is found not only in psychology. In the field of linguistics, two different approaches have been adopted in searching for deep structures in the understanding and production of language. Chomsky(1965) looked at the problem as essentially one of syntax and developed the theory of transformational generative grammar. Fillmore(1968) held that the basis of natural language understanding in humans is essentially one of semantics and developed the approach called case grammar. In cognitive science, one of the major foci is to look for possible basic forms of memory structure that can cope with the processing of information and learning. Rumelhart and Norman's work (1985) is an interesting example. In the field of artificial intelligence, especially in the areas of machine learning and natural language processing, the search for basic structures of cognition is also a very important theme.

In philosophy of science, there has been a general weakening of the empirical-inductivist tradition which holds that true discoveries can be made by inductive generalization from carefully made unprejudiced observations. Current philosophical perspectives (e.g. Polanyi, 1958; Popper, 1972; Kuhn, 1970; Lakatos, 1970; Feyerabend, 1978) no longer believe that existing scientific models are 'final'. There are doubts about the nature of 'truth', 'objectivity' and 'reality' and a strong tendency to view scientific theories as constructions of the human mind. The following quotation exemplifies such a position:

> *"Science is not just a collection of laws, a catalogue of facts, it is a creation of the human mind with its freely invented ideas and concepts. Physical theories try to form a picture of reality and to establish its connections with the wide world of sense impressions."*
>
> *(Einstein & Infeld, 1938)*

Furthermore, the development of scientific knowledge is less often seen as an individual enterprise. Paradigmatic changes take place as an outcome of the cumulative efforts of the scientific community. The social institutions of science are seen as having important roles to play in the development of the sciences. They are the mediators of scientific knowledge and tend to influence what is considered a problem in a given tradition. They focus on opinions, appraisals and criticisms. They structure the socialization of young scientists and provide organized outlets for the announcement and evaluation of innovations by way of conferences and publications.

Thus on the one hand there is a recognition of the tentative nature of the status of our present scientific knowledge that scientific theories are also constructions of the human mind. On the other hand there is the growth of mistrust in the value of what science can offer, seeing that its misuse can lead to so much human misery.

Another major development in philosophy which has a more indirect influence on research in the area under review, but is nonetheless extremely important, is the work on phenomenology which started with Husserl (1913) and Heidegger (1926). As Dreyfus pointed out (in Magee, 1987), Husserl's claim that the foundation of all human understanding is in the directedness of mental content marked a culmination of the Cartesian tradition which sees the fundamental human situation as that of a subject in a world of objects. Heidegger reacted against this strong Cartesian viewpoint and proposed an entirely different view of this problem which Magee (1987, p.258) describes as follows:

> *"How do we as subjects gain knowledge of the objects that constitute the world? ... Can such knowledge ever be certain? ... Now Heidegger is saying that these questions are fundamentally misconceived ... it [is] a profound misconception to regard them as being the most important*

questions. Primarily, in our most characteristic modes of being, we humans are not subjects, spectators, observers, separated by an invisible plate-glass window from the world of objects in which we find ourselves. We are not detached from some external reality which is 'out-there', trying to relate to it. On the contrary, we are part and parcel of it all, and from the very beginning we are in amongst it all, being in it, coping with it. ... We are beings in amongst and inseparable from a world of being, existences in an existing world, and it is from there that we start."

In Heidegger's conception, there are different modes of operation in man's daily activities, the most fundamental of which is that of the coping being, as described above. Then there is the level which man, the problem solver and rational animal, the mode of operation characteristic of the Husserlian subject, directing his/her mental attention to specific aspects of the external world. There is also a level where man deals with the predicates and laws of science, but this would be at third remove from the level of everyday coping. Thus scientific theory, which can explain context-free causal relations very well, could not be expected to explain the everyday meaningful world of significance that Heidegger describes.

This phenomenological perspective has had a strong influence on many branches of contemporary thought. An important example is Alfred Schutz's work in developing a phenomenology of social reality, attempting to give an account of the foundations of the social sciences. Schutz's thought is well summarized in his posthumously published work "The Structures of the Life World" (Schutz & Luckmann, 1973). He started his thesis with the following statements:

"The sciences that would interpret and explain human action and thought must begin with a description of the foundational structures of what is prescientific, the reality which seems self-evident to men remaining within the natural attitude. This reality is the everyday life-world. It is the province of reality in which man continuously participates in ways which are at once inevitable and patterned."

This reality of the everyday life-world, according to Schutz, is characterized not only by the taken-for-grantedness of the natural attitude but also its intersubjective foundations. The everyday life-world is not a private world, but rather, the world of our common experience, much of our knowledge of which is social and contextual. As Zaner (translator's note, Schutz & Luckmann, 1973) pointed out, "Knowledge and society are recognized by Schutz as deeply interwoven. .... the reality of the everyday life-world is a social reality, it possesses social structures of relevance into which everyone is born, and in which he lives, and 'grows older' with his fellow-men. The child in his first interaction with others is included in a reciprocal motivational context with structures of relevance (goals, means, attitudes) that have been socially delineated and are 'taken-for-granted'." The phenomenological view of knowledge has made important impacts on educational research both in terms of the formulation of research perspectives as well as methodology, for example, the work of Marton and the Goteborg group (Gibbs, 1982; Marton, 1981).

Such developments present a serious challenge to the realist tradition in the philosophy of science and to the traditional view of learning as a quantitative process, consequently producing important impacts on those involved in science and science education. Against the psychological background of increasing concern over the learner's own constructions, researchers on children's ideas differ along the dimension of 'respect' for the wrong ideas found in the learner.

At one end of the spectrum there are those who draw a strong parallel between children's constructions and those of a scientist (Driver, 1983), allotting the status of 'children's science' to such ideas (Gilbert, Osborne & Fensham, 1982; Osborne & Freyberg, 1985). Because many of the intuitive ideas about natural phenomena held by children have been found to coincide with some of the ideas that appeared at various stages of scientific development, especially in the domain of mechanics, some researchers have drawn strong parallels between the two (e.g. DiSessa (1982) referred to children's ideas as Aristotelian, and McCloskey (1983) found strong similarities between students' intuitive ideas and the 'impetus theory'). Such a position was usually justified with reference to the non-final character of scientific ideas.

At the other end there are those who maintain that these ideas are misconceptions which arise mainly due to inadequacies in the learner's history of experience, and in the instruction which would be needed for the formation of coherent scientific concepts (McClelland, 1985).

Those currently working in the area of research into children's ideas in science are predominantly within the 'constructivist' tradition, giving due respect to children's ideas on grounds of personal validity. However, there are still important differences within this group as to the status to be given to children's constructions. Though both science and our stock of commonsense knowledge of the world are seen as socially constructed bodies of knowledge, some hold that there are fundamental differences between the cognitive activities of the scientist and that of the child (e.g. Ogborn, 1987). To call children's ideas 'children's science' is to obliterate this difference which may not be helpful towards our understanding of the phenomena to be studied. The epistemological difference between science and commonsense knowledge (which is deemed to have the same origins as children's ideas) is discussed in the following section.

2.4 Epistemological differences between Science and Commonsense

Given that both the natural environment and social forces are at work in shaping the development of scientific knowledge as well as commonsense ones, what distinguishes one type of knowledge from the other? In what sense can we say that scientific knowledge is superior to commonsense? The way any research in children's ideas is organized and perceived is very much affected by the positions taken by the researcher with respect to these questions. Thus an attempt is made in this section to present the epistemological issues as they will be understood in this thesis.

2.4.1 A difference in purpose.

A common sense construction of the world consists of a set of dependable expectations, self-evidences concerning what everyone ought to know about the common and basic activities of everyday life. These are to ensure dependable anticipation in matters of human action and to insure against unanticipated surprise. It thus has a basically pragmatic concern. In this respect, it is very similar to what Schutz and Luckmann (1973) described as the world of everyday life:

> *"The world of everyday life is ... man's fundamental and paramount reality. ..... the everyday life-world is to be understood as that province of reality which the wide-awake and normal adult simply takes for granted in the attitude of common sense. By this taken-for-grantedness, we designate everything which we experience as unquestionable; every state of affairs is for us unproblematic until further notice."*

Scientific activity on the other hand looks for a complete systematic theory of the world which can stand the most stringent test against reality. This aim largely derives from the metaphysical commitments of scientists, and from the purpose of scientific activity, which is the advancement of scientific knowledge as a whole.

The very different outlook taken by scientists is exemplified in the following quotation from Einstein and Infeld (1938):

> *"With the help of physical theories we try to find our way through the maze of observed facts, to order and understand the world of our sense impressions. We want the observed facts to follow logically from our concept of reality. Without the belief that it is possible to grasp reality with our theoretical constructions, without the belief in the inner*

harmony of our world, there can be no science. This belief is and always will remain the fundamental motive for all scientific creation."

Although the concept of reality may not be the same for different scientists, yet in all of them lie strong metaphysical commitments to this reality. Moreover, there is a strong commitment to a belief in 'the inner harmony', beauty and simplicity of our world and a strong desire to discover more about this physical reality. Polanyi (1958) called this *scientific passion* which "serves as a guide in the assessment of what is of higher and what of lesser interest; what is great in science, and what relatively slight", and argued that any process of enquiry unguided by such passions would inevitably fall into trivialities. The scientists' vision of reality, to which their sense of scientific beauty responds, suggest to them the kind of questions that may be reasonable and interesting to explore. It is such intellectual commitments that give rise to the distinctively uncommon ways in which scientists go about enquiring about the world around us. While we may be surprised to find in children's ideas myriad understandings of physical phenomena very different from the accepted scientific view, it is the latter and the process by which it has come to be regarded as such that is really uncommon, amazing and difficult to understand (Ogborn, 1987).

2.4.2 A sociological perspective

Because of the distinctly different purposes served by these two types of knowledge, the kinds of social influences in the shaping of such knowledge are quite different.

In our non-problematic functioning in our everyday life-world, we require our own conception of this reality to be not a private conception but an intersubjective reality shared by our fellow men. The fundamental structure of this reality must be shared. We also have to take for granted that the significance of this everyday life-world is fundamentally the same for everyone in order that a common frame of interpretation is possible. This is the only way in which a common, communicative world can be constituted. And it is in this sense that the province of things belonging to the outer world is also social for us.

Our stock of commonsense knowledge has to be non-critical by definition. In some sense, it develops as a result of a collusion, and indeed this is necessary to allow us to operate effectively in our daily business. The origin of such knowledge is diverse: besides personal experience, it may be derived from folklore, myths, literature, religion, morals, etc. This allows us to pass on the stock of life-world knowledge not only from person to person, but also from generation to generation through society's cultural heritage. As cultures can differ, the actual contents of commonsense knowledge can also differ from culture to culture. Another characteristic of commonsense knowledge is that a lot of this knowledge is tacit, and is unexposed for conscious examination. Even in those areas where this knowledge is expressible using our common language, indeterminacy is a prominent feature of the expression. According to Nagel (1961), there are typically two types of language indeterminacy in this area: (i) the terms in ordinary speech may be quite vague in the class of things designated and not sharply and clearly demarcated from the class of things not so designated (for example, water may be used to refer to a whole range of different compositions of liquids), (ii) the range of presumed validity for statements employing such terms usually do not have determinate limits.

Thus, because of the social role of commonsense knowledge and its pragmatic purpose, it is very difficult to alter commonsense radically. This, coupled with the indeterminacy in everyday language, makes it very difficult to challenge commonsense ideas.

Science, on the other hand, has developed entirely different characteristics in its pursuit of models of reality which give a unified systematic picture capable of accounting for a whole range of seemingly unrelated phenomena as well as predicting unknown phenomena and directing further inquiry into the world around us. Scientific activity is characterized by its continual conscious critical examination and testing of its own claims. This criticalness is not only found in the scientist as an individual, but largely reinforced by the scientific community. All the claims as well as the assumptions and procedures have to be made public and explicit. Nothing will be taken for granted, any new finding has to be repeatable, and new theories have to stand up to all conceivable challenges before they can be accepted. The establishment of a paradigm is a long and arduous process, resulting from the concerted critical effort of a whole community of scientists.

The criticalness required in the development of science is facilitated by its use of formal representations and precision in the use of language. The use of formalism serves as a device to reduce vagueness, to be used both as a tool to think with, as well as a language for communication within the scientific community. By being very precise, statements become capable of more thorough and critical testing. By making claims explicit and specific, scientific theories are more exposed to challenges and refutations. This need for precision leads to very different meanings being attached to words which are common to both science and ordinary language. As Oppenheimer (1954) very clearly pointed out:

> *"Often the very fact that the words of science are the same as those of our common life and tongue can be more misleading than enlightening, more frustrating to understanding than recognizably technical jargon. For the words of science .... have been given a refinement, a precision, and in the end a wholly altered meaning."*
>
> *Oppenheimer (1954), p.3.*

This problem of language is thus also a major difficulty which science teachers face.

2.4.3 Structural Differences

Nagel (1961) suggested distinguishing scientific laws into two types: experimental laws and theories. The former are basically empirical findings proposed and asserted as inductive generalizations. Scientific theories, e.g. the kinetic theory of matter, are often concerned with entities which are not observable, eg. molecules. Theories are not arrived at by induction from experimental data, but by relating the abstract terms used to observable phenomena, explanations being logical deductions from the premises/axioms of the theory. Thus theories are fallible and experimental laws would persist. However, the former are much more powerful tools with which to describe and understand the world. In a scientist's mental conception, the most general comprehensive theory has the highest priority and the hierarchy of theories forms the foundational basis of his reasoning.

It would be unreasonable not to expect some sort of structure in the organization of commonsense knowledge, but it remains a problem how much structure to expect. Larkin (1983) has found some interesting differences in problem solving behaviours between experts and novices. The main differences seem to be related to the use of different problem representations. Novices use what she called a naive problem representation, composed of objects that exist in the real world with solutions attempted by using operators that correspond to developments that occur in real time. Experts were found to construct a second mental representation, which Larkin termed 'physical', containing fictitious, imagined entities such as forces and momenta. Such a finding would be an expected consequence of structural differences between scientists' knowledge and common sense knowledge.

2.5 Children's Ideas and the History of Science

Many of the intuitive ideas about natural phenomena held by children have been found to coincide with ideas that appeared at various stages of scientific development. The comparison between children's ideas and pre-classical mechanical ideas is one that has been most widely noted. DiSessa(1982) and Whitaker (1983) referred to children's ideas as Aristotelian, while McCloskey (1983) found strong similarities between student's intuitive ideas and the 'impetus theory'.

2.5.1 Comparison with the Aristotelian View

The ideas often labelled as Aristotelian are: (i) that motion needs a force and (ii) that motion takes place along the direction of the force. However, this is as far as the similarity goes.

The most important argument against such labelling lies in the fundamental difference in the nature of the two sets of ideas. Children's ideas are intuitive ideas normally held tacitly for pragmatic purposes. The Aristotelian conception, on the other hand, was a well thought out, explicit set of arguments that ties into the wider context of the Aristotelian world view, providing support for a whole philosophy. The distinguishing characteristics of the Aristotelian conception of motion according to Wartofsky (1968) are:

i) The provision of a consistent framework which orders everything in the world - hence the definition of a natural place for all things (the ground) leading to the division of motion into 'non-violent' ones which move towards this natural place and motions that go against this natural order which need 'violent' forces.
ii) The employment of mathematical tools - a reliance on Euclidean geometry as a formal language.
iii) The whole conception of motion is not only explicit but is a formal system in itself. For example the term velocity was carefully defined using the language of Euclidean geometry.

Actually, Aristotle's dynamical views could almost be derived from his cosmology - the view that the terrestrial realm consists of four concentric spheres each with its particular properties.

2.5.2 Features common to the Impetus Theory

McCloskey (1983) argued for strong similarities between children's ideas and Medieval 'impetus theory'. The main ideas in common are:

i) Motion needs a cause to be found *inside* the moving body. When a mover sets an object in motion, he implants something into it (labelled 'impetus' by impetus theorists like Buridan and very often labelled 'force' by children) enabling the object to move in the direction in which the mover starts it.
ii) Motion stops as this 'impetus' or 'force' gradually dissipates.
iii) This 'impetus' can be increased or supplied by external agents.

These modes of reasoning do seem to account for a large number of the departures of the children's ideas from the classical view.

2.5.3 Reservations about the use of historical labels

As discussed above, parallels can be drawn between children's ideas of motion and those held by philosophers and theorists in the past. Yet there are reservations against the use of historical labels for children's intuitive ideas of motion.

First of all, the historical ideas were serious, systematic attempts at developing a theory of motion, whereas children's ideas are just intuitive beliefs about how things are, held tacitly and not explicitly exposed. One example is the idea of different kinds of impetus for different types of motion, which do not exist in intuitive ideas of motion (Saltiel & Viennot, 1985).

Secondly, both the Aristotelian view and the impetus theory are very much guided by an overall philosophy. There is no evidence that children hold philosophical stance of the theorists with whom they are identified.

2.5.4 What can we learn from comparisons with past theories?

Comparison studies always end up with a set of similarities and a set of differences. If we wish to look at the differences, it may be interesting to find out why children's ideas seem to coincide more with the medieval impetus theory of motion than with the Aristotelian view. If we leave the differences aside, there are striking similarities in the conception of motion in the Aristotelian view, in impetus theory, and in children's conceptions of motion. An investigation into such similarities might offer clues to where the major conceptual difficulties lie.

One major recurrent theme is that rest and motion are completely distinct states. It is related to the idea that motion needs a cause. As Gilbert and Zylbersztajn (1985) have described, a major difficulty in children's learning of classical mechanics lies in the counter-intuitive Newtonian concept of uniform motion and rest as being on the same ontological level. 'Force' seems to be the single most frequently used term to describe something that is salient in affecting motion. 'Force' may be ascribed as external to the moving object. But there is also a tendency to ascribe the ability to move as coming from an internal characteristic - a 'force residing in the moving object'. The use of the same term 'force' to stand for these two different notions is reflected even in Newton's *Principia* (1687) - 'impressed force' used to stand for what we now call force acting on something and 'innate force' used to stand for the notion of linear momentum (Saltiel & Viennot, 1985). The term 'force' was also attached to the notions of kinetic energy and potential energy as these were slowly worked out by scientists after Newton, and continued to be used even within the scientific community long after these concepts had been painfully worked out and defined.

2.5.5 Fallacy of the children-as-scientist metaphor

It seems inappropriate to compare children's ideas or constructions too closely with scientists' activities. We find that there is no exact parallel between children's ideas and past scientific theories. More importantly, science is a sophisticated human activity demanding a host of formal abilities as well as the adoption of an entirely different epistemology. DiSessa (1984) gave an interesting account of how the learning process may differ as a result of the different epistemologies adopted by the students. Bliss, Morrison & Ogborn (1988) while reporting on a longitudinal study of development of dynamics concepts in secondary school students in which they found that some of the students' informal ideas changed but some persisted, and also suggested that the difficulties the children face in changing some of those concepts have deeper roots. The Newtonian view of force and motion involves a complete imaginative reconstruction of concepts of motion, not merely seeing motion as it is seen by another person. Evading or neglecting these issues would be a serious mistake if we intend to tackle the problem of why learning science is difficult.

2.6 Theoretical Frameworks of this area of Research

2.6.1 Status of children's conceptions

Two dimensions serve to locate the large numbers of studies of children's conceptions of various topics in science. One is the interpretation of the nature of the phenomena and the status given to such conceptions by the researcher. As indicated previously, there are those who see the unorthodox ideas as just wrong, as mistakes in the learning process (Helm, 1980, McClelland, 1985). However, this perspective is not shared by the bulk of researchers in this area of study who are all broadly classifiable as constructivists. The second dimension is the kinds of the research questions asked.

2.6.2 Kinds of research questions asked

Some researchers explore the problem within the context of science teaching, while others see it as related to human cognition in general. The former are engaged in two main kinds of investigation: identification of the various conceptions held by students in specific domains, and studies of models of conceptual change and of possible teaching strategies to bring about desired changes. By contrast, those involved in the latter kind of study generally believe that children's conceptions in science are in fact specific instances of much more general modes and structures of cognition, and see their main task as to identify such structures. Thus this kind of study is much more influenced by the school of cognition to which the researchers subscribe. Few researches in this category are empirical explorations into ways of promoting conceptual change, though various psychological schools of cognition hold implicit views on how this might be achieved.

2.6.2.1 Identification of domain specific conceptions and units of analysis employed

Studies aiming at identifying domain specific conceptions may differ from one another in the kind of entity they are ultimately interested in and this affects the approach taken in the analysis of data. Gilbert and Watts (1983) made a comprehensive review of the different approaches used in the analysis of data obtained by various methods of elicitation of children's ideas done within this context. They propose to differentiate between three levels of analysis.

The first level was that of describing "personalized theorizing and hypothesizing of individuals" and the label 'conception' was given for this level of analysis. They further proposed that "conceptions are accessed by the actions (linguistic and non-linguistic, verbal and non-verbal) of the person, often in response to particular questions". This level of analysis is normally employed in the context of case studies (e.g. Hewson, 1980) which have the merit of being able to gain rich data about a specific case, but are limited in generalizability.

The second level of analysis is the construction of "groupings of responses which are construed as having similar intended meanings" and which represent an interpretation of statements at a functional level. This includes analyses like Osborne's (1980) categorization of one-line statements from interviews and Engel's (1982) categorization of students' answers to a particular question. This level of analysis attempts to generalize beyond the individual, but is complicated by the fact that a particular response may be arrived at via very different modes of thinking (Engel, 1982; Gilbert & Pope, 1982).

The third level of analysis is at the level of 'frameworks', or "short summary descriptions that attempt to capture both the explicit responses made and the construed intentions behind them". These are second order constructions, "thematic interpretations of data, stylized, mild caricatures of the responses made by students", and are thus generalizations across individuals as well as across different contexts. Examples of such frameworks include propositional statements capturing some aspects of children's reasoning (e.g. constant motion requires a constant force) and descriptions of attributes given to certain entities or scientific terms (e.g. a reflection is an image on a mirror or surface). Driver, Guesne and Tiberghien (1985) suggested relating the analysis with the elements of information stored in the long term memory and called such elements or groups of elements 'schemes'.

As yet, there is no consensus as to the appropriate level of analysis, though there is a trend towards identifying frameworks rather than analyzing data at the first two levels. However, it is not clear what suffices to establish a framework. A related problem is the absence of a general formalism for the representation or presentation of children's ideas. Results from empirical research are normally presented in the form of aspects of the children's ideas as inferred from the interviews and exemplified by quotations. There is generally no explicit mention of the level or kind of analysis adopted; indeed different levels of analysis are sometimes intermingled.

2.6.2.2 Instructional design and promoting conceptual change

Most of the work in this area adopts a notion of cognitive conflict in their intervention/ teaching strategies. It generally begins by encouraging students to make explicit their views and understanding of the domain. The teaching stage is aimed at heightening students' awareness of different possible ways of looking at the domain, raising doubts about their own conceptions and hopefully showing the superiority of the accepted scientific view, employing strategies like Socratic dialogues (which might include doing experiments) (e.g. Nussbaum & Novick, 1981; Champagne, Gunstone & Klopfer, 1985), or working in computer-based learning environments (e.g. diSessa, 1982; White, 1984).

Some of this work makes little explicit reference to a psychological model of learning (e.g. Nussbaum, 1985). Others have a clear commitment to a particular model of learning and intervention (e.g. Rowell & Dawson, 1983; Cosgrove & Osborne, 1985; Hewson, 1981).

Results of intervention have been very mixed. As Duit (1987) pointed out, a great "breakthrough" is not yet in sight.

2.6.2.3 Identification of deeper structures of cognitive processing

As mentioned above, some work identifies possible general cognitive structures underlying conceptions. Some is undertaken by those whose main interest is in developing general theories of cognition, seeing science as an area potentially providing rich ground for exploration. Others have moved in this direction from a starting point nearer to the first position. Many have pointed out the need for a deeper level of analysis, and for some theoretical constructs to provide a more comprehensive description of responses to various questions/situations, and predictive explanatory power.

2.6.2.3.1 A common sense theory of motion. Ogborn (1985) proposed to look for a 'theory of the content of alternative conceptions' in mechanics to 'provide a consistent integrated framework, more general than a set of rules about particular cases, which could hope to explain the relationships between situations and responses, and to account in a uniform way for the various conceptions of motion that researchers have characterized.' Such a theory was seen as a formalized version of commonsense, following the spirit of a theory of naive physics put forward by Hayes (1979). The theory was not to be taken as describing deviant or mistaken ideas, but rather as reflecting the basic nature of intuitive conceptions of motion.

Tests of such a theory might be an assessment of the descriptive adequacy of the theory as an account of a range of known conceptions, or its ability to generate acceptable (by commonsense standards) explanations of motions.

Two further levels of explanation were suggested to be necessary. One level was a 'physical' one, showing how organisms living in a Newtonian world can have experiences that would lead naturally to their developing a non-Newtonian theory of motion. The second was a 'psychological' one, explaining the existence of the theory in people's minds in terms of a deeper level of mental functioning.

2.6.2.3.2 Duplication schemes or Prototypes. Guidoni (1985) also proposed investigating 'natural thinking'. The proposed sketch is different from Ogborn's in that it made explicit assumptions about what kind of mental representation this type of 'natural thinking' is supposed to take. Another point of departure was that Guidoni held the view that 'a person's natural-thinking system is *not* a "coherent theory about the world", and if we try to interact with it as if it were, we shall deceive ourselves'.

'Natural thinking' refers to the 'background' complex mental functioning that is always going on, inseparable from any specialized (e.g. scientific) thinking. Its function is in '*driving behaviour within a context according to a purpose*' and it is achieved through a '*schematized duplication of the context, the purpose, and of their features which allows for off- and on-line efficient control of that fit, between subjective purpose and objective context*'. Such a '*radical schematization process*' is supposed to be at the root of our knowledge structures. The ideal structure of a resulting 'duplication scheme' or 'prototype' should be such that it is able to incorporate the maximum number of repetitive features which can be handled with the minimum attention, and at the same time be so structured as to make meaningful 'variables' salient and to hide those which are not. It should also contain basic blocks which allow for easy combination and structuring so as to permit the handling of increasing complexity.

Guidoni further proposed that in 'natural thinking', analogical reasoning is the basic reasoning strategy. A theory of 'natural thinking' should 'seek out keys to mechanisms which might yield patterns or correlations' and provide insight into processes of conceptual change.

2.6.2.3.3 Mental models. Gentner and Gentner (1983) are also concerned about deeper cognitive structures, exploring the conceptual role of analogy. Analogical comparisons, sometimes explicit and sometimes implicit, abound in people's descriptions of complex systems. The problem was: are they thinking in terms of analogies, or merely borrowing language from one domain as a convenient way of talking? They describe their '*generative analogy hypothesis*' in which analogies play the role of structure mapping, with the source domain providing a structural relationship between different elements which can be used for the target domain. An experiment designed to test this view is discussed in the next chapter.

2.6.2.3.4 Experiential gestalt of causation. Andersson (1986) again reiterated the need to find a deeper cognitive basis for understanding pupils' reasoning. He proposed the existence of a common core to pupils' explanations and predictions in such widely differing areas as temperature and heat, electricity, optics and mechanics. He called this core the experiential gestalt of causation, an analysis inspired by Lakoff and Johnson's (1980) work on language and metaphor.

According to Lakoff and Johnson, causation as a gestalt starts to be constructed at a very early stage. A feature common to many actions is that there is an agent which directly, with its own body, or indirectly, with the help of an instrument, affects an object. This causation gestalt also embraces related experiential generalizations such as: that the greater the effort by the agent, the bigger the effect on the object; that different objects resist to different degrees; that several agents have a greater effect than just one; and that the nearer the agent, the greater the effect. Andersson tries to demonstrate, using examples from a variety of areas in science, that many of the alternative frameworks can be explained as attempts to employ this experiential gestalt of causation in understanding phenomena in diverse areas.

2.6.2.3.5 Phenomenological primitives - a matter of cognitive priority. DiSessa (1983, 1985) also tries to look for general cognitive structures that are employed by novices in their intuitive interpretation of phenomena across different subject domains. He sought for small units of structure as the fundamental building bricks of cognition, which he called phenomenological primitives, or p-prims for short. Examples of p-prims given include springiness, dying away, and 'Ohm's p-prim' which postulated the relation of effect of effort exerted by agent to resistance encountered.

Like Guidoni, he argues against the 'Theory theory' interpretation of children's conceptions, contending that children's intuitive conceptions are fundamentally fragmented systems, and that it is wrong to attribute theory-like status to them. He further suggests that the most important difference between novices and experts are not at the level of content but at the level of the extent and kind of systematicity in the knowledge system.

One important feature of diSessa's work is his commitment to the belief that children's intuitive conceptions are necessary starting ground for *continual* growth and development toward scientific conceptions, thus differing from the dominant view that conceptual conflict is the key to development. Learning was perceived as an "evolution along the dimension of selecting some naively recognizable phenomena for more systematic and general application as knowledge structures".

2.6.2.4 Arguments against searching for general cognitive structures

Despite these various attempts, there is no concensus that a search for more general cognitive structures is necessarily a positive direction to follow.

Viennot (1985) instead proposes a much more pragmatic line of approach. She suggests looking for regularities in different kinds of students' productions (verbal or non-verbal responses to questions/situations) and inferring from them elements of mental organization. However, instead of looking for a single integrated framework to account for all sorts of conceptions in a content domain, the units of structure looked for need only be some sort of correlational description between situations and responses. According to Viennot, a set of regularities can be "explained" by different sets of "underlying reasons" or "mediators" and there is no clear set of criteria that can be used to decide between such alternatives. For example the incorrect response in fig. 2.4 may be explained by : i)students reasoning as if motion implies a force, ii) students employing a "spring" metaphor (or mediator) in reasoning - pushing on the spring (firing engine) changes its state, and releasing the spring (stopping engine) restores it to its original state, or, iii) students employing a "swimming" metaphor in reasoning - the initial motion being like a swimmer drifting in a river and the firing of engine evoking the idea that the "swimmer" now swims.

a. The problem (Clement 1982) :
A rocket is moving along sideways in deep space, with it engine off, from point A to point B. It is not near any planets or other outside forces. Its engine is fired at point B and is left on for 2 seconds while the rocket travels te some point C.
(a) Draw in the shape of the path from B to C (show your best guess for this problem even if you are unsure of the answer).
(b) Show the path from point C after the engine is turned off, on the same drawing.

b. Correct and incorrect answers to rocket problem.

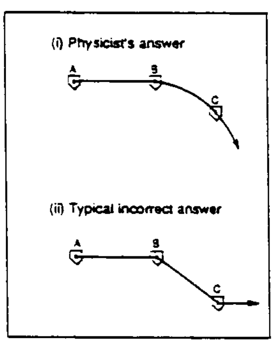


Fig. 2.4 Rocket problem together with the correct answer and a typical incorrect answer.

Viennot thus argues that whether the metaphors chosen should be purely phenomenological accounts of a regularity or whether they should have explanatory power was deemed to be difficult to answer and anyway not easy to distinguish. A pragmatic attitude was taken by Viennot in this: "let us give up trying to sort out what cannot be distinguished, and let us speak simply of more or less useful 'mediators'".

Arguments along a similar vein are proposed by Solomon et. al. (1985). They suggest that the personal constructions of students are better seen as an extension of a familiar life-world way of thinking requiring neither abstract concepts nor consistency rather than as a scientific kind of construction. They propose that pupils do not realize the demands for consistent explanation made by the scientific way of thinking, and that they just pick whatever they want from the available stock of socially acquired knowledge in different situations. However, the effect of the child's own cognitive development on the kind of selection made from available knowledge, and how far children will attempt to weld together pieces of the social stock of knowledge into a more coherent generalized world picture, require further empirical investigation.

2.6.3 Deep structures - what are they like?

From the foregoing survey, there are several pertinent features common to nearly all of these proposals:

i) A call for taking children's ideas seriously *in their own right* and not as a curious deviation from orthodox scientific ideas. Thus a level of description of children's ideas in their own terms and not in comparison with scientific terms or concepts is aimed at.

ii) There is a general reference to some kind of relationship between children's ideas and commonsense or life-world knowledge.

iii) A few of the accounts explore the nature of possible 'deep structures' with a view to the kind of mental/memory representation that may be at the root of our cognitive functioning.

iv) Some of the accounts point to the desirability of the deep structures to provide for a mechanism for conceptual change. A search for deep structures should thus preferably be carried out with some possible learning mechanisms in mind.

2.8 Where does the present research stand?

The author shares the same belief, as many others in this area do, that children's intuitive conceptions in science are not simply mistakes but represent outcomes of positive endeavors of children trying to make sense of the physical world around them. Further, this phenomenon is to be seen as part of a wider, more general issue in learning, and it is certainly possible that more general cognitive structures underlie people's intuitive reasonings in widely different subject domains. However, what is not clear is the interpretation of evidence. Most of the empirical work to date has been gathered as responses to specific tasks or problems and the analyses tend to highlight the prominent deviations from the accepted scientific views. Thus whether the intuitive conceptions do form a 'theory' or not has so far been a matter of conjecture without conclusive evidence either way. Further, because the data were gathered as on-task or spontaneous responses of the subjects, the higher level general conceptions are necessarily second or third level constructions by the researchers concerned.

The research questions here are:

How would students describe their intuitive knowledge about a certain subject domain if they are not confined by specific problem or task contexts?

Would they come up with an intuitive 'theory' or just fragmented pieces of ideas?

If one succeeds in identifying such conceptions, what cognitive role do these play in the daily processing of information?

Further, would the students' descriptions of their intuitive knowledge be affected by the amount of science instruction they receive in class?

Can any 'evolutionary' difference amongst different students' productions be detected?

The research questions formulated above evidently present problems of finding a suitable methodology for the investigation. This will be discussed in the chapter on research methods. Mechanics was chosen as the subject area for investigation because this was the best researched area, and one in which there is evidence that people have developed a rich stock of intuitive knowledge.

Ch. 3. LEARNING, KNOWLEDGE AND REPRESENTATION

3.1 Learning

Anyone in science education wanting to deal seriously with the problem of children's ideas will very soon come up against the questions: How do we learn? How is it possible for anyone to learn new things or to change strongly held ideas? One's views on these questions determine to a large extent the kind of questions asked, the actual research design and analysis for any research in children's conceptions in science. This synopsis does not purport to be a complete review of the works of major theorists in this area, but rather tries to document the author's personal understanding of the issues concerned, as this represents a major dimension of the theoretical framework underpinning this piece of research.

The word 'learning' here carries a broad reference to all learning that may take place and not just learning in a school context. Thus this includes both the case of informal learning as an individual in the everyday context, as well as learning as a social enterprise, with scientific development as one of its branches.

This review begins by introducing briefly some major schools of thought on learning as an epistemological problem. However, as the main purpose of the review is to clarify the theoretical foundations relevant to the research, discussion is restricted mainly to perspectives that might be broadly labelled as 'constructivist'. The idea of classifying learning into various types will then be discussed. Finally some views on how conceptual changes may be effected through metaphors (including similes, analogies and models) will be presented.

3.1.1 The Meno Paradox - How can we learn anything new?

> *You argue that a man cannot enquire either about that which he knows or about that which he does not know; for if he knows, he has no need to enquire; and if not, he cannot; for he does not know the very subject about which he is to enquire.*
>
> *Plato, Meno 80E; Jowett translation*
> *(quoted from Petrie, 1979)*

This age old problem of how people could ever learn anything that they did not know before has troubled humanity ever since Plato wrote his famous paradox, and is in no way settled as yet.

The main issue at stake is how knowledge is related to experience - whether all our knowledge comes *from* experience or just that knowledge comes *with* experience. One school of thought which can be broadly labelled as rationalist (or structuralist) believes, in some sense, that all our knowledge is essentially a priori or innate. Plato's solution to this problem is a classic example of the rationalist view. He believed that the soul has gone through many lives and in so doing has seen all things, though it has forgotten them. Thus learning is just a matter of the person recollecting what he has already known before, and teaching then is appropriately a process of reminding. Though probably no one would now accept Plato's view in toto, the belief in innate knowledge is still shared by many. Experience is essentially seen as the occasion for our coming to know something but not as where we directly derive our knowledge.

The empiricist camp, on the other hand, believe that all our knowledge comes from repeated exposure to experiences of a similar kind - learning by some kind of induction. The process itself is relatively mechanical as the repetition at one stage is thought of as sufficient for the development of the next. To borrow the analogy of *tabula rasa* as Hamlyn (1978) used it to describe the empiricist view, 'the growth of experience is a function of indentations of a given kind on the wax, so that the mind is structured according to the predominant role given to the deepest and therefore most prominent indentations'.

Though very few of us may be classified as entirely rationalist or empiricist, the sort of position we take up along the dimension of this dichotomy does very much affect the way we think about the very nature of the research involved in explorations into learning. Indeed, the way cognitive theorists go about their investigations into the processes of learning are implicitly governed by their views on this epistemological problem. To use Hamlyn's (1978) characterization, the empiricist stand represents a view of learning as genesis without structure, and the rationalist stand represents a view of learning as structure without genesis. Skinner and Chomsky respectively represent these two different camps while Piaget might be described as taking up a middle position.

Skinner believes that all behaviour, including verbal behaviour, is learnt through building up linkages between items of experience, by a process which he called operant conditioning. This brand of conditioning works like other brands of conditioning but differs in that instead of the stimulus-response link being reinforced, it is the reward-operant link.

Chomsky, a linguist who sees himself in relation to rationalists like Descartes and Leibniz, was interested in constructing models of human linguistic competence and thereby to specify the 'universals' of language. He held that 'the environment per se has no structure, or at least none that is directly assimilable by the organism. All laws of order, whether they are biological, cognitive or linguistic, come from inside, and order is *imposed* upon the perceptual world, not derived from it.' (Piattelli-Palmarini, 1979). Thus he was convinced that language is represented in the mind in an extremely abstract fashion, in so called linguistic deep structures, and claimed that human inheritance specifies, or at least highly constrains, the rules of syntax that a child could invent.

If we use structure and genesis as criteria for the characterisation of any epistemological stance, then Piaget saw the process of learning as genesis with structure. For anyone who would not want to accept that we have any knowledge that is innate, and at the same time is not satisfied with the haphazardness of association as the basis of our cognition, then he/she is faced with the problem of trying to explain how progressively more complicated structures can be built up through experience. Piaget, in his work on genetic epistemology, tried to build up an account of human learning as a building up of the necessary structures at each stage of development, with each stage providing the necessary structures for a subsequent stage, and the stages happening in a necessary order. Thus Piaget (1970c) described his own stance as follows:

> *'... for the genetic epistemologist, knowledge results from continuous construction, since in each act of understanding, some degree of invention is involved; in development, the passage from one stage to the next is always characterized by the formation of new structures which did not exist before, either in the external world or in the subject's mind.'*
>
> *'Genetic Epistemology', p.77*

Though different cognitive psychologists have different theories of what the learning process is like, those that have come to be identified as constructivists share the same views as Piaget on the following: (i) learning is an active attempt of the individual to make sense of the world around him, (ii) every individual possesses the ability to build up more complicated knowledge structures from his interactions with the environment and his existing knowledge structures.

3.1.2 Types of Learning

We are learning new things every day, and most of the learning we do is quite unproblematic as well as uninteresting. I learn that the electricity charges will be raised next month, or that the telephone district code for my district has just been changed. We take in new information all the time, and much of it we do not find difficulty in learning. On the other hand, there are a lot of things I find very difficult to learn, and some I may not be able to master at all within my whole lifetime: I don't know how to ski, I took a few French lessons and then dropped out, and high energy physics is (so far) beyond me. It would be reasonable to say that learning tasks do differ in their levels of difficulty and it would be helpful to examine what characterizes such differences.

According to Piaget's (1975) latest model of learning, equilibration plays a central role in a person's cognitive development (or 'self-regulation' of human knowing). According to this theory (see Furth, 1981, 253-283), equilibration regulates the cognitive 'cycles' and keeps them in more or less permanent balance(equilibrium). It operates through the dual processes of assimilation and accommodation: during an individual's encounters with the environment, there is always a tendency for the individual to assimilate where and when it is possible by means of available cognitive schemes, and at the same time, for the schemes to accommodate, to make modifications in response to the specific situation. Thus schemes assimilate content, at the same time and considered from the other direction, schemes accommodate to content. Knowledge then has bi-directional tendencies: the outgoing direction is accommodation which assures contact with the world and thus defines the object of knowledge; the inward direction is assimilation which indicates the regulatory function through which this contact is organized and coordinated and defines the subject of knowledge. This prescribes a radical view of knowledge as a relational concept, relating the subject to the object and encompasses a truly constructive interaction between the two.

In Piaget's model of equilibration, assimilation and accommodation, though opposed in direction, do not function against each other, but rather necessarily go together. Accommodation does not always mean the permanent change of a scheme in response to external pressure and may very well be an episodic adjustment of the same scheme to the constantly changing contents to which it is applied. Undoubtedly, assimilation and accommodation do contribute to developmental changes of schemes. His latest version of the model hypothesized that such development is made via three types of regulated balance, or 'knowing cycles': within-scheme, between-scheme and totality cycles. The within-scheme cycle considers the knowing act in its present interaction and is essentially concerned with positive-negative regulation, what the object is and what it is not, in the subject-object relation. The between-scheme cycle concerns scheme differentiation and has to do partly with object knowledge and partly with subject knowledge. The totality cycle focuses on integration of differentiated parts into structured wholes. It derives from subject knowledge but is abstracted. At the formal stage, this regulation is context-free. The whole theory is very complicated and it is not appropriate here to go into it in any detail.

However, a relevant point to our discussion here is that Piaget sees cognitive development as the building up of more comprehensive and abstract knowledge structures.

A relatively new approach to the study of cognition is the information processing approach. Again, it would be inappropriate to go into great detail about it here. It suffices to say that the development of this branch of cognitive psychology is stimulated by developments in information technology, and benefited from the opportunity to test theories by machine implementation. Though the success of any machine implementation of a psychological model in no way guarantees the psychological validity of the model proposed, it does provide a method of testing its plausibility. Even more important in the author's opinion is that, for the exploitation of machine implementation, a microscopic level of analysis is required. The psychologist has to deal with detailed contents of the memory as well as a detailed account of how the particular cognitive mechanism under investigation proceeds and how it affects the memory. Thus it is not surprising that a variety of formalisms for the representation of knowledge in long term memory have been developed in recent years by various information processing psychologists. Such formalisms also provide us with a more expressive language to describe knowledge structures.

Two information processing psychologists, Rumelhart and Norman (1978), proposed to differentiate between three modes of learning: accretion, tuning and restructuring. They used the phrase 'knowledge modules' (KM) (Norman 1978) as a neutral name for the units of memory structure so that their argument for the three modes of learning does not rely upon the particular knowledge representation formalism chosen. Norman (1978) defined the three modes as follows:

> *"... accretion is simply the addition of of new information either within or guided by existing structures. There are two ways accretion might occur. First an existing KM may have incomplete structures (amounting to unfilled arguments or unspecified arguments or unfilled slots). Accretion in this case is simply acquiring the appropriate information to fill out the existing KM structure. Second, new knowledge may be guided by an existing KM. Thus an existing KM acts as a prototype module for the construction of a new KM.*
>
> *........[Restructuring] is often characterized by new insight into the structure of the topic. .... if accretion is knowledge acquisition, restructuring is knowledge understanding. .... there need be no formal addition of knowledge by the student during restructuring.*
>
> *......[For] tuning, the proper KMs and concepts are assumed to be present, but the structures need to be refined. Unnecessary computation needs to be eliminated, unnecessary variables need to be replaced with particular values. Some information that was previously*

computed or inferred is stored directly. Thus, during tuning, many (thousands) of special cases are acquired as specialized KMs"

It can be seen that Rumelhart & Norman do not focus on a progression in generality as the main criterion for learning, as Piaget does, but look at the kinds of structural change that take place. Accretion is the normal kind of fact learning, including our daily accumulation of information. The acquisition of memories of the day's events, the learning of lists, dates, names, telephone numbers are examples of learning through accretion. There are no structural changes in the information-processing system itself.

Tuning is a more significant kind of learning through which we gain skill and proficiency in the performance of certain tasks. It requires more than just an addition to the database. It involves continual tuning (or minor modification) of the knowledge structures to bring them more into congruence with the functional demands placed on these structures. Improvement in typing speed and increase in proficiency in solving a particular class of problems are examples of tuning.

Restructuring is by far the most significant and difficult process and is necessary for the learning of complex material, usually requiring a time scale of years to achieve. Through this process new structures are created. Once the appropriate structures exist, the learner may be said to "understand" the material.

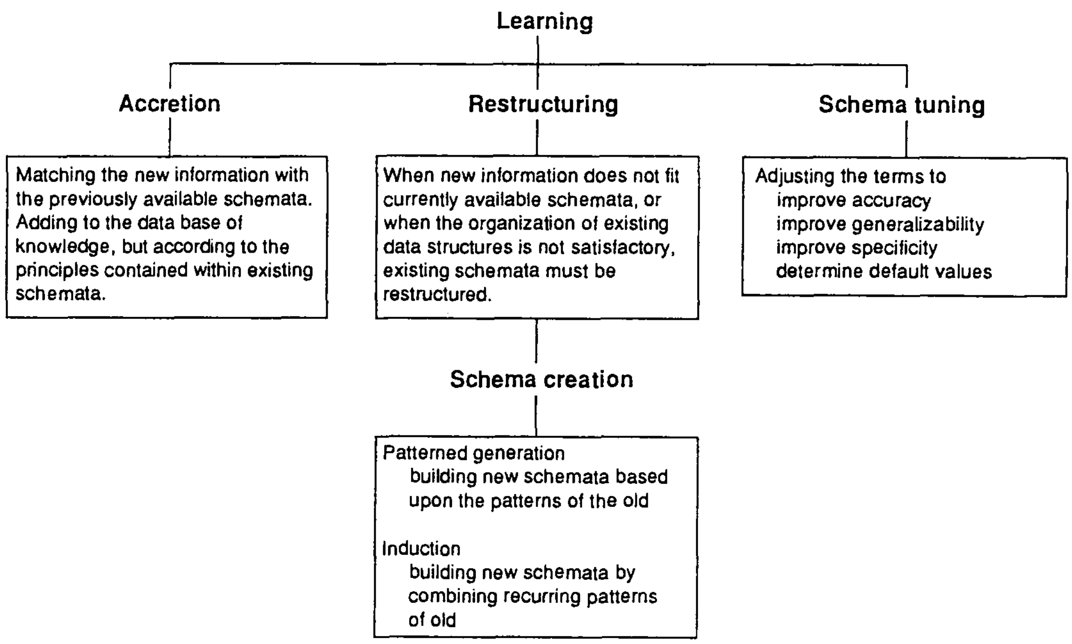


Fig.3.1 A classification of the mechanisms by which learning might occur (Rumelhart & Norman, 1978)

From the point of view of research into children's ideas in science, restructuring seems to be the most important aspect. Indeed, many researchers in this field acknowledge that there needs to be a fundamental change in the cognitive structure of the learner in order to dispose of misconceptions. It is important to point out here that Rumelhart and Norman do not see these three processes as entirely separate events, and very often all three processes may coexist, although one process may be dominant at any one time.

3.1.3 Mechanisms of learning at a contextual level

The foregoing section discusses learning from a general perspective. The kind of learning that is of most interest as well as relevance to the issue of children's intuitive conceptions in science would probably be at the level of Piaget's totality cycle or Rumelhart and Norman's restructuring. However, the issue about learning here is not so much a matter of general development but how a particular knowledge structure changes into a radically different one. There is a need to describe and analyse cognitive structures and their development at a much more microscopic and concrete level. Thus the problem of learning is not investigated from a Piagetian perspective and the nature of cognitive structures assumed is quite different. Piaget was interested in looking for the basic cognitive structures that are necessary for a person to be capable of the logico-mathematical activities that are required to carry out scientific investigations, and the temporal order in which they appear. Research in the field of genetic epistemology has made important contributions to our understanding of human cognition, probably having the most pervasive impact too. On the other hand, as Bliss (1986) pointed out, the limitation of Piaget's description of cognitive structures at the formal stage to operational schemes alone, (without domain specific references), coupled with the fact that the formalisms employed (propositional logic, lattice theory, etc.) are only capable of expressing static states rather than dynamic aspects of cognition, makes it difficult to be of direct use to the task at hand.

A major strand of development that warrants our attention in looking at the possible mechanisms for the genesis of specific knowledge structures is work in the area of metaphors (or analogies) as instruments for cognitive change. Those who claim the cognitive importance of metaphors use it as a broad label to encompass similes, analogies and models as well. They vary in their views on whether metaphors are the only means for creating new cognitive structures. Rumelhart and Norman (1978) suggested that there are logically only two ways for forming new schemas: (i) a new schema can be patterned on an old one, consisting of a copy with modifications, (ii) it can be formed by induction from co-occurring configurations of schemata. The former was labelled patterned generation and the latter schema induction. Though they see schema induction as a possible mechanism for schema creation, they believed that most learning takes place through creation of new schemas by patterned generation. This is in fact very similar to Schank's (1986) idea of creating new knowledge structures through 'pattern-based explanation' when one fails to find rules or beliefs that explain a given event, and this is done through attempting to find rules from some other domains that might fit the case at hand. Petrie (1979) also made a strong claim for the case of metaphors - "the very possibility of learning something radically new can only be understood by presupposing the operation of something very much like metaphors".

To understand the cognitive role of metaphors better, it might be helpful for us to first examine a characterisation of metaphors from the linguistic perspective. Black (1962) made a distinction between the principal subject and the subsidiary subject in a metaphor. The principal subject is the subject (or context) of the discussion and the subsidiary subject is the subject (or context) from which reference is drawn. Black also identified two elements in a metaphor: a focus and a frame. For example, for the sentence "the chairman ploughed through the discussion", the principal subject is the discussion and the subsidiary subject is farming, the word "plough" would be the focus and the discussion is the frame. Thus the 'frame' is the principal subject and the focus is the particular aspect(s) of the subsidiary subject to be focussed on. Metaphors may be used purely for stylistic reasons, and as condensed similes, so that the metaphor replaces an equivalent literal comparison. Black (1962) argued however that the richest and most interesting form of metaphor operates at the interaction level. This *'interaction view'* of a metaphor holds that "the metaphor selects, emphasizes, suppresses and organizes features of the principal subject by implying statements about it that normally apply to the subsidiary subject".

Petrie (1979) was interested in metaphors as learning tools and identified their most interesting feature to be "the use of a familiar rule-governed device for dealing with the material to be learned in ways which require the bending or even breaking of the familiar rules.... It provides a mechanism for changing our modes of representing the world in thought and language."

Gentner and Gentner (1983) adopted an information processing approach to the investigation of the cognitive role of analogies. They gave a more detailed description of analogies in terms of knowledge structures. Analogies were defined as 'structure-mappings between complex systems'. An analogy would convey that the relations that hold among the objects in the *base domain* (the known domain) can be applied in the *target domain* (the domain of inquiry). A structure mapping asserts that identical operations and relationships hold among non-identical things. The relational structure is preserved, but not the objects nor their properties.

Though the perspectives taken by each of the views mentioned above are not exactly the same, they all see metaphors as using the structural relationships in one context to structure an understanding of the second context.

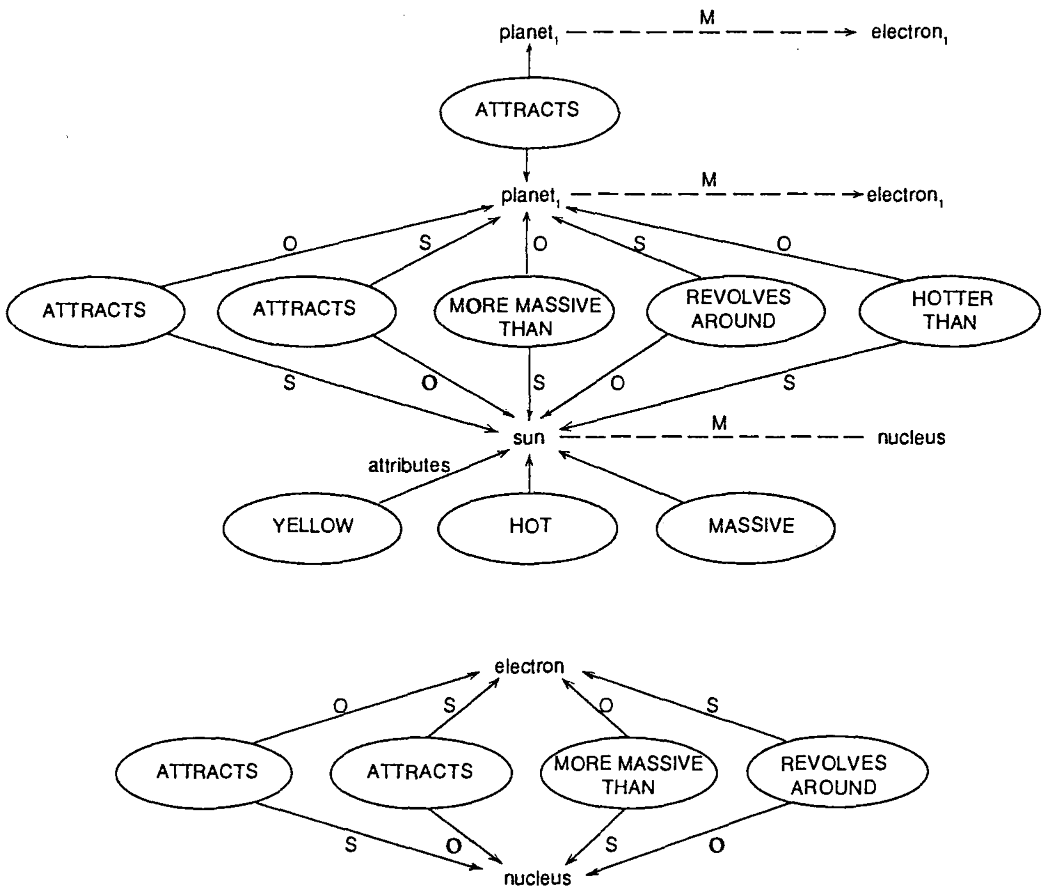
3.1.4 Metaphors and education

If metaphors or patterned generations play a central role in the generation of new knowledge structures, then they should also figure prominently in teaching situations. According to Petrie (1979), this is the case and he elaborated on the mechanisms involved. A new schema may be produced by the use of metaphors through the process of *'accommodation of anomaly'*. This process depends very much on one important characteristic of metaphors: the understanding of metaphors, be it comparative or interactive, *depends on the presupposed cognitive scheme*. The same metaphor can be either comparative or interactive, depending on the understanding of the person trying to make sense of it. "The atom is a miniature solar system" is probably a comparative metaphor to the science teacher because s/he can recognise the similarities between the two systems. However, to a student just starting to learn about the atomic structure, in order for the metaphor to be successful, it has to be interactive. In fact, the metaphor in this case ***creates*** the similarity for the student. This is the same as Black's (1962) assertion that 'it is more illuminating ..... to say that the metaphor creates the similarity than to say that it formulates some similarity antecedently existing'.

Thus, in Petrie's view, learning something radically new starts when a teacher makes a statement (understood by the student as an assertion) in the form of a metaphor, and the student lacks a cognitive structure sufficient to recognise it as a comparative one. In this case the assertion would be seen to be false if only the literal meanings are taken, and this creates an anomaly so great that it cannot be assimilated into the student's existing cognitive structure. The student may then modify his or her cognitive structure by focusing on the literal aspects of the subsidiary domain and slowly working out which are the aspects which do apply and which do not. Thus, through this process, new structures may be built by modelling on existing ones, breaking and modifying some of the features that do not apply.

Petrie also attributed the creation of new knowledge structures in the growth of science to scientists' search for new metaphors or models in order to resolve the anomalies created when problems cannot be solved by the current paradigms. The main difference he saw between the scientist on the frontiers of knowledge and the student is that in the student's case the metaphor is provided by the teacher and so is more likely to be immediately helpful than the variants tried out by the scientist.

Gentner and Gentner (1983), adopting an information processing formalism for the representation of metaphors, made the concept of new schema generation even more specific. In their structure-mapping definition of metaphors, both the base and target domains are viewed as systems of objects and predicates. Predicates are either relationships between objects or attributes of the objects. To create a new schema, the learner starts with the known base domain and maps the object nodes (objects represented at the nodes of a structural representation of the knowledge structure) of the base domain to the object nodes of the target domain. After setting up this correspondence of nodes, the analogy conveys that the relationships that hold between the nodes in the base domain also hold between the nodes in the target domain. This can be illustrated by the following structural representation (fig.3.2) of the analogy between the solar system and an atom.



key:
M mapped to
S subject of relation
O object of relation

Fig. 3.2 Representations of knowledge about the solar system and the hydrogen atom, showing partial identity in the relational structure between the two domains. (Gentner & Gentner, 1983, p.103)

There are two basic principles to be observed in Gentner and Gentner's model of schema creation :

i) *Preservation of relationships* - it is only relationships existing in the base that have a high probability of being transferred to the target. Attributes tend to have very low probability of transfer.

ii) *Systematicity* - Since interconstraining relations are particularly important in explanatory analogy, a relation that is dominated by a potentially higher-order relation is more likely to be transferred to the target system than an isolated relation.

In the same paper they also reported on an experimental investigation to look for observational evidence that analogies do have genuine effects on a person's conceptions of a domain. Basically, they wanted to look for evidence that a person's reasoning in the target domain is genuinely affected by the analogical model they adopts, and is not just used as a linguistic shorthand. Thus they need to find observational evidence that cannot be derived from the surface features of the analogies. The domain they chose was simple electric circuits and the analogies they used were two commonly used analogies in teaching: the hydraulic model and the moving objects model. They predicted that people using the hydraulic model should find it easier to deal with circuit problems involving parallel batteries whereas those using the moving crowds model should find it easier to deal with problems involving parallel resistors. Their results showed systematic differences in the patterns of inferences in the target domain as predicted. Since such differences cannot be attributed to shallow verbal associations, their results do give support to the view that the analogies we use do affect our understanding of the domain of study.

3.2 Learning and Knowledge Representation

The nature of the present research is essentially one of knowledge elicitation, trying to explore how intuitive conceptions about motion are actually structured. The issue of knowledge representation arises in at least two contexts: the medium for representation of knowledge during the elicitation process as well as the representation used in the analysis of data. In fact a particular choice of representation formalism would largely determine the kinds of research methodologies employed. The main part of the knowledge elicitation in this research is done through the building of expert systems in Prolog by students, and it is essentially that the theoretical assumptions as well as implications in using first order logic (on which the programming language Prolog is built) for representation, and the effects of translating intuitive knowledge into a declarative program are examined. An awareness of the broad kinds of formalisms available for knowledge representation, what they were designed for and their underpinning psychological assumptions, would also be helpful in the development of a formalism for representation in the final analysis. The review is not to justify the formalisms employed in this research in the sense that they are the best (and indeed the criteria for the best is arguable and dependent on the context of use), but to find out which are the resources available that could reasonably be employed.

3.2.1 Knowledge Representations used in science education research

In discussing the problem of knowledge representation, one important issue that needs clarification is the locus of the knowledge that is to be presented: is it the domain knowledge as a body of public knowledge for an established discipline, or is it the private understanding of certain individuals in a certain domain area as a psychological entity? Some of the problems faced by these two kinds of representations would be quite different, and it is the latter type that is more difficult and complex. The main problem here is that public knowledge tends to be much better defined and accessible whereas private understanding tends to be idiosyncratic and illusive. Another difficulty is the criteria for choice of appropriate representations. For representation of public knowledge in a certain discipline, it is easier to judge whether a certain representation is suitable by finding out whether it serves a particular purpose, e.g. solving certain kinds of problems. For representation of private knowledge, how such representations could be utilized is debatable, and the issue of assessment is even more difficult.

3.2.1.1 broad categorizations - a third order perspective

Attempts to represent children's intuitive conceptions in science broadly fall into two categories. Many of the first investigations into children's conceptions tend to broadly categorize the elicited conceptions according to representative viewpoints, prominent characteristics, or certain dimensions of the knowledge structures. These are third order perspectives in the sense that they are the researchers' categorizations on their interpretations of what the students' conceptual structures are like. In all such cases, an important justification for the categorization is the compelling need for reduction of data, to produce a generalised, manageable and meaningful reduction of the data by trying to encapsulate the essential features of the conceptual structures. Necessarily, how such reduction should best be done depends largely on the assumed characteristics of cognitive structures, which in turn are determined largely by the underpinning psychological model adopted.

White (1985) proposed reducing interview protocols into nine `dimensions' of cognitive structure: extent, precision, internal consistency, accord with reality or generally accepted truth, variety of types of element, variety of topics, shape or form of organization of cognitive structure, ratio of internal to external associations and availability of knowledge. This proposal was based on a static model of cognitive structure developed by Gagne and White (1978).

Whilst White's proposal focused mainly on the contents and properties of the knowledge elements, Pines (1985) proposed analysing cognitive structures in terms of a taxonomy of conceptual relations, which gave sole consideration to the kinds of relation existing between knowledge elements. His thesis for the value of such a taxonomy is that it emphasizes the primacy of relations in all meaning and this is justified because symbolic relations play a central role in human cognition. Pines illustrated his idea by giving examples of how two powerful conceptual relations: set-element relationships and whole-part relationships can be used to explain a lot of the domain knowledge in biology as well as in some areas of human knowledge. This is an admirable attempt and is similar to efforts of some computer scientists in their attempt to develop knowledge representation languages for the construction of computational models of intelligent behaviour. However, this task is dauntingly formidable: whilst the task of developing a language for representing well-defined domain knowledge for computational purposes is still a major research issue in the area of artificial intelligence, the apparent idiosyncracy of knowledge in everday cognition makes one doubt whether it is possible, at least theoretically, to reduce all of human cognition to a few basic conceptual relations. The ultimate success of such an attempt would depend very much on whether commonsense conceptions essentially form hierarchically structured and logical systems.

Many of the data reductions made by researchers in the analysis of data about children's conceptions in science tend to be less bound to particular psychological models, and are focussed at highlighting the actual viewpoints typically found in a group of respondents. For example, Gunstone and Watts (1985) identified five distinctive types of understanding concerning the use of the word `force'. The categories developed from the interview data were intended to encapsulate the main ideas found in the form of `conceptual frameworks'. This kind of categorization is helpful in providing an overview of the main conceptual viewpoints for a particular population group.

Data reduction is important and useful in any kind of research, and research in the area of children's conceptions is no exception. However, it is very often insufficient to just represent or report on third order perspectives without giving due attention to the second order perspectives - the individual subject's actual conceptual structures as perceived by the researcher. These are important for several reasons. First of all, there are very often rich and subtle differences between individual's conceptions belonging to the same category. Such differences would be even more important if we want to investigate the evolution of ideas: how some students have managed to change their intuitive conceptions into more scientific ones whilst others have not, all starting from roughly similar conceptions. Secondly, research attempts differ from one another in terms of techniques employed, dimensions of categorization as well as in the actual categorization elements developed. It is sometimes very difficult to understand the rationale and value for a particular categorization without gaining access to the actual conceptual structures that the researcher generalized on.

3.2.1.2 Representations preserving knowledge elements and structure

Some of the work in the area of children's conceptions in science tried to explore ways of representing children's knowledge so that both the contents of the knowledge elements and the relations between them are preserved. Some tried to elicit the conceptual structures by asking students to draw a concept map of their own ideas (e.g. Novak et al., 1983), while others (e.g. West et al., 1985) tried to work out the structure used by students from interview protocols. One of the main difficulties in developing such knowledge structures is in providing for `conceptual anchors' to initiate the knowledge elicitation process, and most of the research in this category start off from textbook knowledge of the domain that the students have to learn and may thus not be able to represent how the individual actually think about problems in that domain in a spontaneous/intuitive context.

Bliss, Monk & Ogborn (1983) developed a relatively context free knowledge representation formalism for the presentation of qualitative data that may also be employed in representing children's knowledge.

If we ignore the problem of elicitation for the moment and focus on the problem of representation formalism, much of the work in this area is inspired by developments in artificial intelligence and language research. It is helpful to review briefly the main types of knowledge representation formalisms.

3.2.2 Some knowledge representation formalisms

The advent of digital computers presented the possibility of creating mechanical models of the mind. By treating all information processing, human and otherwise, as a formal system that can be studied without reference to the physics or biology of the system carrying them out, computation offers a powerful metaphor for thinking about human cognitive processes. This computational metaphor is also the main theoretical basis of cognitive science, a newly developed discipline trying to understand cognitive phenomena, which views the human mind as a complex system that receives, stores, retrieves, transforms, and transmits information (Stillings et al., 1987).

In the context of knowledge representation, though we can see two perspectives for the problem: "What kind of representation do people use?" and "What kind of representation is best for machine intelligence" (Winograd, 1975), there is as yet no clear distinction between these two questions.

Representational formalisms were first developed to deal with representation at the concept and sentence levels, and these can all be classified as semantic type representations. Subsequently, there was pressure for development of formalisms that can provide for better mechanisms for structuring of concepts, and for supra-sentential structuring of knowledge to cope with complex tasks. The ensuing two sections will discuss briefly these two broad classes of representation formalisms.

3.2.2.1 Semantic nets type representations

Knowledge representation formalisms started as networks of nodes with linking lines (or links for short). Though more and more sophisticated formalisms are continually being evolved, and some of them do not look like networks anymore, they can still be viewed as consisting of nodes and links, only that the nodes and links themselves have acquired more complicated structures. This is the only feature that is common between all the various formalisms that have evolved; whilst the actual *semantics of the formalism*, that is, what the nodes and links represent are different from formalism to formalism.

The idea of a semantic network representation for human knowledge is generally acknowledged to have originated from Quillian's work (1967,1968) to build a semantic memory model of the human long term memory. The focus of the representation was on word concepts. This network model has two important features: a superclass-subclass taxonomic hierarchy and the description of properties for each class deriving from the two link types, one indicating a "subclass" relationship and another indicating a "modifies" relation. Thus properties true of a class are assumed true of all its subclasses except for the modifications. Consequently, the superclass chain extending upwards from a concept embodied all properties true of that concept. These features lead to the notion of *inheritance of properties* in a semantic net (see fig. 3.3 for details). The plausibility of a hierarchical model for human memory seemed to be supported by results of reaction time studies (Collins and Quillian, 1970).

Fillmore's work on linguistic case structure (1968) helped network development to focus attention onto a particular class of concepts - verbs, thereby stimulating developments for

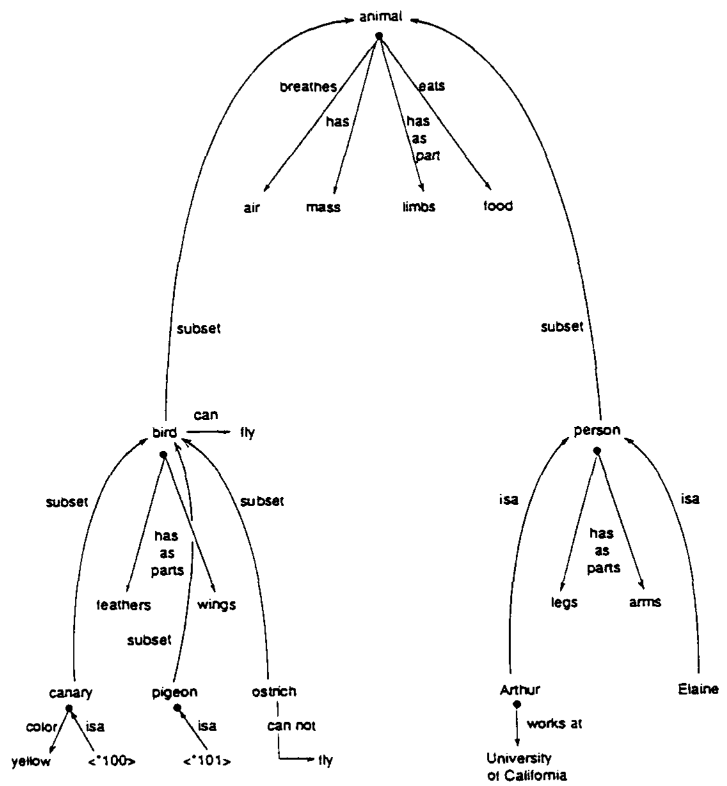


Fig. 3.3 A simple semantic network, chosen so as to illustrate the use of inheritance. (Rumelhart & Norman, 1985, pp.25)

the unit of knowledge representation to be raised to the level of a sentence. An example of network formalism that incorporated case structures was that of the LNR group (Rumelhart and Norman, 1973). It can be seen from fig. 3.4 that links coming out from the verbs were all case pointers. This was an attempt towards defining a limited set of unambiguous link types.

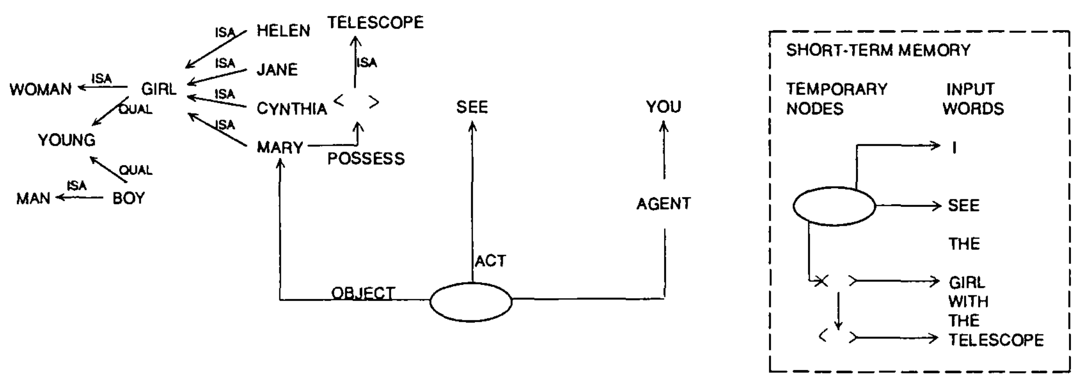


Fig. 3.4 A semantic network representation of the sentence "I see the girl with the telescope". The links coming out from the verbs are all "case pointers". (Rumelhart & Norman, 1973,pp.456)

Further developments in semantic net type representations include the search for knowledge primitives (e.g. Schank's conceptual dependency representation, 1973), and the search for a logically adequate formalism (e.g. Schubert & Cercone's propositional network, 1975).

3.2.2.2 Higher level units of knowledge structuring

So far, the representations discussed were at the lexical and sentential level and no provisions were made to express the relationships between such small units of knowledge as larger chunks, nor knowledge at the supra-sentential level (Rumelhart and Norman, 1985). Furthermore, these earlier notations all employ uniform formats of expression so that all link types were at the same level and processed in the same way. Such representations prove to be inadequate for the expression and processing of more complex knowledge required for more intelligent tasks. This inadequacy is with reference to a different level of demand than those discussed above. The nature of this inadequacy was clearly expressed by Minsky :

"It seems to me that the ingredients of most theories both in Artifical Intelligence and in Psychology have been on the whole too minute, local, and unstructured to account - either practically or phenomenologically - for the effectiveness of common-sense thought. The "chunks" of reasoning, language, memory, and perception ought to be larger and more structured; their factual and procedural contents must be more intimately connected in order to explain the apparent power and speed of mental activities."
(Minsky, 1981, p.95)

Several formalisms focussing on higher level units of knowledge were put forward at roughly the same time: Minsky's *Frames*, Rumelhart and Norman's *Schema*, Schank's *Script* and Abelson's *Plan*. Though these notations differ in their representational format, there were a number of pertinent features common among them.

Each unit of knowledge at the highest level (with their names as highlighted above) in some sense represent the stereotypes of some aspects of the outside world: events, actions and sequences of actions, etc. The processing of information was equivalent to finding the unit that best fits the incoming information. The configuration of such higher level units together form the interpretation of the input. They all share these same features:

i) These units have variables, which have default values, so as to capture a lot of the knowledge which we may term as "experience". Thus each unit may be thought of as a packet of information containing variables which may be assumed to take on the default values unless specified.

ii) Such units can embed, one within another to form bigger and more complex units.

iii And as a result of the embedding of units, these notations may be used to represent knowledge at all levels of abstraction.

iv These formalisms were developed with a common underlying view of the human memory - we have episodic as well as semantic memory. The latter is some sort of definitional knowledge of the concepts whereas the former represents the facts and relationships we know about those concepts and which may be personal. These higher level units were representations of episodic memory more than semantic memory. Thus these units in a sense really represent our knowledge as opposed to definitions.

These higher level units of representation provide a much richer syntax as well as a mechanism to deal with information at different levels of specificity. However, the specificity of these formalisms bring with them an associated difficulty: to employ such formalisms one has to be sure that the nature of the knowledge to be represented is of the same type as that which the formalism is designed for.

3.2.3 Problems of knowledge representation for the present research

As mentioned in an earlier section, it is hoped that in this research, children's commonsense ideas about motion can eventually be analysed in the form of sets of metaphors/mental models that they employ in their thinking. Thus it is possible that some sort of semantic net or schema/frame type repesentation will be used in the representation of the final analysis. However, it is not possible at the outset to pre-determine which formalism would be more appropriate, given that the essential features of the knowledge structure to be explored is largely unknown. The fact that Prolog programming is used as the main activity for eliciting students' ideas requires a formalism for the representation of the programs developed. The systemic network notation developed by Bliss, Monk & Ogborn (1983) for the representation of qualitative data differs from the formalisms reviewed above in that it does not impose any particular node or link type nor structure on the knowledge to be represented. It is thus considered to be an appropriate formalism for the first level representation of the PROLOG programs developed by the subjects. Representations for more structured conceptualizations would probably have to be developed on the basis of deeper levels of analysis done on the data collected.

CHAPTER 4. METHODS OF ENQUIRY

4.1 Review of techniques for investigating students' conceptions

With the growing interest in finding out what the learner knows, in the belief that it has a strong relevance to teaching, many techniques have been developed to investigate children's conceptions. The development of each technique has been guided by some theoretical considerations - mainly psychological. Very often, the purpose of the research determines to a large extent the nature of the data collected. What sort of data can be collected is as much dependent on the technique chosen as on the aim of the research. Thus, before we can sensibly ask the question "What are the strengths and weaknesses in each of the different techniques used?", we have to first look at the underlying theoretical commitments and the purpose of the investigations they were used for, rather than the techniques per se.

Research methods used for exploring into learners' conceptions can broadly be classified into two categories: those that have a clear focus on exploring structural relations in the learners' conceptions, and those that do not have such an intention. Different techniques have been developed for different purposes in the former category by researchers subscribing to three different theoretical traditions, the associationist, Ausubelian and information processing approach. Techniques in the second category are used by many in the area of research into children's learning in science and it is difficult to classify the theoretical frameworks they subscribe to beyond the fact that they are all broadly constructivists. The first two sections of this chapter give a review of the main techniques in these categories in terms of their respective research interest, and the nature and characteristics of the data collected. The final section discusses the main methodological considerations in the present research and gives the broad rationale for the actual research design.

4.1.1 Research Methods that Explore into Structural Relations in Learners' Conceptions

Implicit in any research methodology for exploring knowledge structures is the underlying assumption of what knowledge structure is, and there is certainly no widely accepted definition for knowledge structure. Three psychological inclinations can be identified amongst those who have made efforts to define and measure cognitive structure: associationist, Ausubelian and information processing approach.

Table 4.1 *Summary of research methods used for exploring structural relations in learners' conceptions.*

Psychological School	ASSOCIATIONIST - Alleged to be of the behaviorist school (Stewart, 1980) - Believe that the closer two concepts are in cognitive structure, the closer they will be on the mental map or tree. As a result, the time for access (or semantic distance) would be smaller.	INFORMATION PROCESSING APPROACH - Draws analogy between human intelligent behaviour and the way a computer works. - Need not hold a strong position on the analogy. May only use a computer as a way to test a psychological model. - Takes seriously the task of expressing and exploring the human cognitive system as a formal system.			AUSUBELIAN - Belief: "most important single factor influencing learning is what the learner already knows". - Need to compare what learner knows with what he has to learn in terms of conceptual structures. - Concern with starting point of learners mainly in relation to the structuring and logical prerequisites in learning of a particular curriculum.	
Research Interest (Purpose of research)	- Attempt to answer question: "To what extent does cognitive structure mirror instructional content for a piece of information?"	Procedural Representation - Used in research on problem solving behaviour. - Comparison between expert and novice behaviour in task performance. - Strong interest in metacognitive skills, e.g. possession of a search strategy other than a general search, and to propose models for the problem solving behaviour in the two groups. - Focus very much on the 'skills' of performance.		Declarative Representation - Develop instructional materials, strategies, study aids, as well as measuring the 'amount of learning'. - Focus is more on the 'knowledge' possessed.		
Nature of Data	- Graphic representation of concept relatedness available for tree construction task and for digraph analysis of text. - Ultimately to arrive at a measure of 'semantic proximity'.	- Transcripts of think aloud protocols, or interviews after performing a task. - To arrive at a plausible process model composed of i) current knowledge of problem, ii) actions for developing problem representation, iii) interpreter for selecting rules.		- Graphic display of students' knowledge structure in the form of concept maps, semantic networks, etc. in which the relations between the concepts are explicitly labelled.		
Technique	Free Word Association Task - give stimulus words and ask subject to list related words under a given context. Controlled Word Association Task - similar to above, list limited for example to 5. Tree construction task - given a list of concept labels, subjects are directed to pick 2 closest in meaning, connect with a line, label 1, continue with remaining labels, 2, 3, 4...etc. until all labels are used up	Think Aloud Protocol - list of verbal statements about what the subject is doing while solving a problem. - can probe further by asking questions about why s/he performs a certain action after the task is completed.	Stimulated Recall - Subject does a certain task and is videotaped. - once finished, view videotape together with subject to discuss what and/or why certain actions were done.	Clinical Interview - About concrete phenomenon. - Description, explanation or predictions about the phenomenon. - Intention is to probe subject's cognitive structure in a narrowly circumscribed area of science.	Concept Map/Line Labelling Task - Similar to tree construction task of associationists', but subject has to explainthe relationship between 2 linked concepts.	Concept Relations Task - give definition of terms and then student asked to describe how any pair of concepts can be related.
Characteristics of the data collected	- Starting point is from instructional content and try to find out the degree of match between learner's memory and the material to be learnt. - Does not provide information on 'natural thinking' of learners. - comes up with numerical measures on various scales like hierarchical clustering and multi-dimensional scaling.	- The knowledge is applied with a purpose, which give a good and valid context for exploring a knowledge structure. - An ultimate test of intelligence is problem solving ability, though not necessarily quantitative.	- Less disturbed by the talking process in the performance of the task. - The interview answers may not reflect actual processes, but a rationalization of what happened after the event.	- Less structured, can be useful for probing into intuitive thinking as well as for more directed probing.	- Provide graphic result, easier to analyse.	- Asking for definition, can tell what understanding the subject has about some scientific terms.
References	Preece (1976) Shavelson (1971, 1974) Gunstone (1980) Stewart (1979)	Larkin & Rainard (1984), Stewart (1980), Chi et al. (1980).		Stewart (1980), Pines et al. (1978), Novak, Gowin & Johansen (1983), Nussbaum & Novak (1976).		
Comments	- Does not seem to be of much use as relevance of semantic proximity as a measure of cognitive structure is dubious. - Neglects inter-relatedness above that of a link between two nodes.	- Rigorous and persuasive results. Provides illuminating evidence on why even hard-working students find problem solving difficult. - The claim that novices only have available general search strategies and that they may not be aware of the possibility of making use of knowledge other than physical principles (e.g. spatial knowledge) is a strong point. However, model does not offer explanation for transition from novice to expert. - Problem solving should be affected by misconceptions, mainly in terms of interpretation of physical principles and the choice of relevant ones. - Scope exists for modifying this approach to take intuitive ideas into account.		- Can have more qualitative data, but no feedback on how realistic the model elicited is. - Incomplete model in terms of how this structure will be used by the subject under any context. - Reseach on subject's understanding essentially with reference to school science curricular, and not understanding in the subject' own terms.		

4.1.1.1 Techniques to measure structure as strength of association

For associationists, allegedly derived from the behaviorist school (Stewart, 1980), relationships between concepts only differ from each other in terms of the strength of association between the concepts, and the nature of such relationships is not an issue for consideration. Here, to find out the learner's knowledge structure means to measure the 'semantic proximity' (Shavelson, 1971) for the concepts to be explored. Furthermore, within such a psychological framework, it is not evident that the learner's pre-instructional knowledge would have any influence on his/her subsequent learning beyond the fact that the semantic proximity of some concepts may not be the same as what would be desired. Thus the main motivation for exploring into learner's knowledge structures by this school is to answer the question "To what extent does cognitive structure mirror instructional content for a given topic area?"

Different techniques can be employed for measuring semantic proximity, depending on the number of pre-defined concepts that is to be explored (Shavelson, 1974). A free word association task gives stimulus words to the subject and ask him/her to list related words under a given context. A controlled word association task works in a similar way, only that the list of related words would be limited to a given number, say 5, thus presetting the maximum semantic distance to be explored. In a tree construction task, a subject is given a list of concept labels, and directed to pick 2 closest in meaning, connect with a line, label 1, continue with remaining labels 2,3,4, etc. until all labels are used up.

The utility of using semantic proximity as a measure of cognitive structure is very doubtful since two different cognitive structures could produce identical or nearly identical responses to such tasks (Stewart, 1979). The proximity of concepts does not reveal anything about the nature of the meaning in the relationships, nor does it handle inter-relatedness above that of a link between two nodes.

4.1.1.2 Techniques for arriving at a graphic display of cognitive structure

Techniques to construct graphic representations of learners' knowledge structures have been developed by those subscribing to the Ausubelian belief that the "most important single factor influencing learning is what the learner already knows" (Ausubel, 1968). However, the concern with the starting point of learners here is mainly in relation to the structuring and logical prerequisites in the learning of a particular curriculum. The main obstacle to learning is seen as the absence of essential conceptual building blocks on which the target learning materials are built. Thus investigation normally starts by analyzing the school curriculum to find out a logical sequence of concepts to be learnt and how they should be related together. Then tasks are set for the learner to compare what s/he knows with what s/he has to learn in terms of conceptual structures. These tasks result in declarative representations of the students' knowledge structure in the form of concept maps, semantic networks, etc. in which the relations between the concepts are explicitly labelled.

The most popular techniques within this category include the concept relations task and concept map (or line labelling) task. The concept relations task asks the subjects to give the definitions for a given set of concepts, and then to describe how any pair of concepts can be related. The concept map (or line labelling) task is similar to the tree construction task described above except that the subjects have to explain the relationship between 2 linked concepts. These two are highly structured tasks, confining the exploration to a given set of concepts from the school curriculum. A less structured task, clinical interview about concrete phenomenon, can also be conducted to ask subjects for descriptions, explanations and/or predictions about the phenomenon, and the graphic representation can then be constructed by the researcher from the interview protocol.

Unlike the case of measurement of semantic proximity, techniques in this category do provide qualitative data on the nature of the relationship between concepts. However, because the elicitation is confined to a set of narrowly focussed concepts, the resulting information reveals only the subject's understanding with reference to school curricular, and not understanding in the subject's own terms. Thus it is not clear how realistic the elicited model of the subject's knowledge structure is in terms of the subject's actual understanding of the particular subject domain investigated. Furthermore, the elicited model does not provide information about how this knowledge structure would be used by the subject under any particular context.

Amongst the set of techniques in this category, the extent to which the elicitation is confined and structured can be most easily altered for the clinical interviews. This technique can thus be modified for use in probing into subjects' intuitive understanding. The disadvantage of this technique, however, is that the form of the graphic representation of knowledge resulting from an analysis of the interview protocols is very much up to the interpretation of the researcher and is much less well-defined.

4.1.1.3 Techniques for probing into dynamic aspects of knowledge structure: organization of knowledge for problem solving

A very different line of research into knowledge structures is carried out by those subscribing to the information processing approach in cognitive psychology. Here, analogy is drawn between human intelligent behaviour and the way computers work. The analogy is normally not held as a strong psychological position, especially for those engaged in explorations in the area of education. Rather, such an analogy promotes the expression and exploration of the human cognitive system as a formal system, and the computer can also be used as a medium for testing psychological models of cognition. Possibly because computers have conventionally been used for problem solving and execution of tasks, most of the explorations into knowledge structures by those subscribing to such an analogy is in the area of investigating problem solving behaviour, looking at the skills, processes and strategies in the performance of such tasks.

Verbal data collected from the subjects about the actual cognitive processes going on during the problem solving process forms the main source of data for all investigations in this category. These are collected either as think aloud protocols while the subject is actually performing the problem solving task, or by stimulated recall immediately after the task is completed. The latter method of protocol collection has the advantage of eliminating the possible disturbance that may arise due to the need to handle the extra talking process while performing the task, but the protocol so collected may suffer from being a rationalized account of what is thought to have happened on retrospect. Transcripts of the protocols collected can then be analyzed to arrive at a plausible process model of the problem solving process, comprising i) current knowledge of problem, (ii) actions for developing problem representation, and iii) an interpreter for the selection of rules for processing.

In contrast to the methods of elicitation described in the previous sections, investigations into problem solving behaviour yield procedural representations of the subject's knowledge, revealing what concepts and how they have been used by the subject in arriving at a solution for a particular problem context. In many cases, models of the cognitive processes hypothesized to have taken place have been implemented on the computer, and the intermediate outputs from the computer compared with the protocols collected from subjects during the problem solving process. Such investigations have succeeded in showing that experts may not differ from novices in terms of the declarative knowledge they possess about the subject domain, but in other aspects. For example experts have access to other related knowledge required for arriving at the solution, use different representations for problems, and possess search strategies other than a general search (Larkin, 1983; Larkin & Rainard, 1984).

Results from investigations in this category have generally been rigorous and persuasive and have provided illuminating insight on some issues, for example why some very hard-working students still find problem solving difficult. One important limitation of this type of investigation is that the elicitation has to be confined narrowly within the context of the problem solving task, and cannot be used for probing intuitive ideas (in declarative format) about a broad domain area.

4.1.2 Methods of exploring into learners' conceptions without focussing on structural relations between concepts

In terms of psychological inclination, those who have investigated into learners' conceptions without displaying a clear focus on exploring structural relationships between knowledge elements can all be broadly classified as constructivists. They all share the common belief that learning is a constructive process involving the active participation of the learner, and that the pre-instructional, intuitive ideas held by learners have important effects on learning. However, two main research interests can be identified within this broad categorization: There are those who use the scientific concepts as reference points and try to find out specific instances where intuitive ideas differ from scientific conceptions. Others believe in the possession of integrated intuitive/commonsense frameworks of understanding by learners about phenomena in a particular domain area. They try to find means of probing into learners' understanding in their own terms and are thus concerned that the subjects would not be cued into adopting a particular mode of reasoning during the course of elicitation.

A variety of techniques has been used to find out how intuitive ideas depart from scientific ones: Carefully designed tests of misconceptions (for example, Viennot, 1979; Sjoberg & Lie, 1981) consisting of multiple choice or short questions can be used to show up a subject's misconceptions, and these are normally used for confirmatory rather than exploratory purposes. Open ended essay tests (for example, Bell, Brook & Driver, 1985) asking subjects for explanations of certain phenomena provide a means of exploring subjects' understanding, but suffer from the inability to gain interactive feedback from the subject and consequently present problems for data interpretation. Clinical interviews about instances (Gilbert , Watts & Osborne, 1981) or events (Osborne & Gilbert, 1980) can be used to explore subjects' understanding of a scientific term or principle, or their reasoning about given phenomena. Such interviews provide rich information on the intuitive reasoning employed if well structured and carefully conducted. The main disadvantage of these interview techniques is that they prescribe a set of concepts or phenomena for the elicitation and may not reflect the actual priorities in everyday reasoning contexts.

Two methods of probing into commonsense reasoning patterns can be identified as having made explicit efforts to avoid giving any cues to subjects in the elicitation process: A variant of the clinical interview about events, interview about comics (Bliss, Ogborn & Whitelock, 1990), may be more useful in eliciting intuitive reasoning by asking subjects about events

Table 4.2 *Summary of research methods used to explore into learner's conceptions, with focus not on structural relations.*

Psychological School	CONSTRUCTIVIST (Piaget, Kelly)					
Research Interest	- To find out specific instances where intuitive ideas differ from scientific concepts. - To develop teaching strategies in view of such ideas.				- to probe into commonsense reasoning, how students interpret phenomema in the domain under investigation in their own terms.	
Nature of Data	- List of alternative conceptions. - Though some researchers called them alternative frameworks, few attempts were made to put such ideas into a comprehensive tangible structure.				- An important concern here would be trying not to cue the subjects into adopting a particular mode of reasoning.	
Technique	Test of Misconceptions	Essay Test	Clinical Interviews (meticulous about being non-directive, non-suggestive, first developed by Piaget)			Repertory Grid
			About Instances	About Events	About Comics	
	- Use of some carefully chosen multiple choice or short questions (usually qualitative) designed to show up a subject's misconceptions.	- Open ended questions asking for explanations of certain phenomena.	- To find out in a more naturalistic manner about a subject's interpretation/understanding of a scientific term. - Give a set of drawings to find out whether the subject considered that as an instance of the concept.	- Focus on a single phenomenon and to explore the various aspects of a student's reasoning about the given phenomenon. - Can be used to explore socially agreed meanings and reasonings.	- Similar to the technique of interview about events, only that the event is an event taken from comic strips, not necessarily a possible one in everyday life.	- Identifying and using bipolar dimensions in a 'semantic space'. - Subject is given or asked to choose 3 objects and then to choose 2 from the 3 which resemble each other in some way and different from the 3rd and to say how they are different. - This creates a list of bipolar dimensions which reflects the criteria the child is consciously or unconsciously using to make sense of the world.
Characteristic of Data Collected	- It is a quick and easy way to find out how a student population stands in terms of displaying certain misconceptions in an area or in diverse areas. - Questions are usually designed for confirmatory purposes, not for exploring students' conceptual structure.	- Requires students to generate rather than select responses, therefore more likely to reflect students' existing knowledge. - Allows for no interactive feedback and thus presents problems for data interpretation.	- Provide a much richer and realistic information on how a person actually understands a scientific term/principle.	- Flexible nature of the interview means that the nature of the data collected depends on the focus of the questions. - Technique and planning very important here.	- It is different from interview about events in that the event may not be possible. These are often exaggerated events. - Comics provide instances which are just on the verge of impossibility. This may be a good medium for finding out children use to determine when something is possible.	- Good for identifying which aspects of some events (e.g. motion) are important for a subject (or a group of subjects).
References	Viennot (1979) Sjoberg & Lie (1981)	Bell, Brook & Driver (1985)	Gilbert & Watts (1980)	Osborne & Gilbert (1980)	Whitelock(1990)	Fransella & Bannister (1977) Whitelock (1988)

that may not be possible in real life. The repertory grid technique (Fransella & Bannister, 1977) can also be adopted for exploring intuitive conceptions. Here, a subject is given or asked to choose 3 objects and then to choose 2 from the 3 that resemble each other and different from the third and to say how they are different. This creates a list of bipolar dimensions which reflects the criteria the child is consciously or unconsciously using to make sense of the phenomena represented.

4.1.3 Comparison of the two broad categories of techniques

It is interesting to note that of the research trying to explore children's ideas in their own terms (rather than in contrast with textbook knowledge), relatively few attempts were made to explore these conceptions in terms of a coherent conceptual structure. On the other hand, most of the work in exploring children's learning of given subject domains, whether they be starting from an associationist, information processing or Ausubelian stand, attempt to explore the possible knowledge structures of the learner.

Another notable observation is that all explorations into knowledge structures have adopted more directive (suggestive) techniques, trying to find out the learners' knowledge structure in relation to given scientific terms and principles in the curriculum, which may not necessarily reflect the most crucial aspects of their 'true' understanding. 'True' here refers to the degree of spontaneity of that particular structure being employed in their normal everyday reasoning. In fact, it is fair to say that all these techniques tend to *prescribe a set of concepts* and then investigate how the learner may possibly relate them under the context set by the investigation (but as understood by the learner).

The above two observations are probably related. As mentioned in the previous chapter on knowledge representation, it is much easier to explore knowledge structure starting from some ground concepts and much more difficult to evolve into comparable structures starting from less structured starting points.

Within the group of explorations into knowledge structures, research in the area of problem solving is particularly interesting. Typical in such research, an *explicit, testable model* of the solver's knowledge, as well as the rules it uses to search for the solution is given. The advantage of such an approach offers a unique means of assessing the plausibility of the proposed psychological model. Moreover, it provides a relevant, *meaningful context* for the structuring of the concepts for the individual (it is believed that the cognitive structure of a person is not static but changing with time as well as with the context). Moreover, the interest in metacognitive skills by researchers in this area is well founded. It would be interesting to see how such metacognitive skills develop and how they may affect a student's learning. The disadvantage of this approach is that it tends to cue textbook knowledge and offers no mechanism for asking about intuitive ideas.

Techniques employed by researches in children's ideas, on the other hand, have mostly adopted more naturalistic approaches to data collection. The most commonly used technique is clinical interview. This form of data collection is costly in terms of time and thus puts severe limits on the possible sample size investigated. The data collected is purely qualitative, and is not easy to analyze. Yet this seems to be the most suitable way of exploring into the deeper levels of cognitive functioning, especially in terms of commonsense reasoning. Interview about events seems to offer good potential for exploring into how children reason spontaneously about different physical situations, given that it tends to go with a more relaxed setting and children tend to feel less on the spot. Another technique, the use of a repertory grid, provides another form of probing into children's conceptions. One major difference of this method from interview about events is that instead of focussing the interviewee's attention on specific aspects of an event, the subject is free to choose which aspects s/he wants to talk about. The bipolar dimensions chosen reflect the most pertinent aspects of the situations that show up in the person's interpretation of the situations. Moreover, the order in which the dimensions are chosen reflects the priority unconsciously assigned to these various aspects in the person's reasoning and interpretation of the situations. Another very important advantage of this technique is that the task is free of any judgement about how to respond to the elements from the side of the researcher. There can be no right or wrong answer to what sort of dimensions may be picked up, and so release the subject from possible anxiety of being judged in his performance of the task.

An important issue in research of this kind is the validity of the claims. How far are the conceptual structures found in the investigations affected by the research design is a question that has to be answered. In some investigations, multiple methodologies have been employed to provide a cross reference on the results.

4.2 Considerations in the design of the present research

As was clarified in the earlier chapters, the main interest in this research is to probe into deeper structures in children's commonsense reasoning about motion and, to find out if these intuitive views do in any sense form an integrated and consistent whole. Technically, the task is one of knowledge elicitation and representation. Issues related to the latter aspect was discussed in the previous chapter. The remainder of this chapter will be devoted to issues related to elicitation.

There are four main methodological considerations in the design of the knowledge elicitation task(s):

i) the task should be aimed at eliciting the kind of reasoning used by subjects in everyday commonsense contexts,
ii) it should not prescribe any particular types of motion or aspect of motion for elicitation,
iii) it should not in any way suggest any particular line of reasoning or focus for the subject to follow,
iv) as an in depth study of children's reasoning, it should provide opportunities for the researcher to follow up on particular points coming up in the process to seek further clarification in a non-threatening, non-judgemental setting.

It is considered to be more appropriate for the purpose of the present investigation to focus on elicitation of declarative knowledge as problem solving contexts tend to cue textbook knowledge, and it is not evident that the issues considered in such contexts do coincide with people's major considerations when reasoning in a commonsense context.

As can be seen from table 4.1, existing methods for exploring declarative conceptual structures need '*conceptual anchors*', that is, the research has to specify the focus and a starting point for the elicitation to start. This introduction of a pre-determined set of parameters for elicitation is not desirable for the present work. Furthermore, these methods tend to elicit *static* aspects of children's views. The dynamic aspects of the elicited conceptions, that is, how these conceptions are actually applied for solving everyday problems, their relative priority and how they work together are not accessible to the researcher.

Some of the methods listed in table 4.2, for example, interviews and repertory grids, have the merit of being able to probe into aspects of children's understanding as seen from their own perspectives, in their own terms. However, the direct data normally do not yield any comprehensive picture of a particular child's conceptions and the reported findings are normally third order perspectives (c.f. Gilbert & Watts, 1983 & discussion in chapter 2). One important consequence of this is that only the most prominent aspects of the subjects' reasoning used in specific situations can be generalized in these types of investigations. One cannot find out easily using such methods how the various conceptual frameworks coordinate or orchestrate together in their actual application, nor whether they do orchestrate.

A good context for eliciting in-depth reasoning under non-threatening contexts would be one in which the subject plays the role of a teacher. This has proved to be a successful mode of elicitation in exploring children's knowledge in arithmetic (Johnson, 1983). In Johnson's study, the subject acted as the teacher, trying to explain to the researcher how s/he solved certain arithmetic problems. The researcher played the role of the student, trying to understand the knowledge that has been passed on, and she tried to ensure that her understanding matches the intended meaning by re-explaining what she thought was explained and the process repeats until both are satisfied that they have reached the same understanding. This technique is attractive as it provides a very good context for feedback and clarification on elicited ideas. Further, the focus is on how the subjects actually think and work and they are under no pressure of being judged on the validity of their knowledge. However, this method was employed for elicitation of procedural knowledge and there is great difficulty in employing this for elicitation of declarative knowledge. It is not easy to find out what someone actually means by a declarative statement without seeing how that particular statement is actually used in specific problem situations, that is, finding out the dynamic aspects of the declarative knowledge.

The advent of expert system shells creates another possibility for the realization of a teaching context. Conceptually, a computer with an expert system shell running on it is like an intelligent being without any worldly knowledge, but ready to learn. Thus instead of teaching human beings, subjects can teach such computer systems. In so doing, they can also find out how well this machine 'learner' has mastered the concepts taught by asking it to answer queries in the domain of interest. Such a computer learner has several advantages over human learners because of its mechanistic nature. First of all, the knowledge taught to the learner is always open to inspection both by the teacher and the learner. Secondly, these learners are entirely free from any domain specific knowledge of their own. Consequently, if the subjects externalize their ideas in the form of generalizable rules, they can then explore the logical implications of these ideas. Furthermore, if they are not happy with any of the externalized ideas, whether because the ideas are not well expressed, or because they have changed their minds, they can always change them and be sure that the learner can always 'unlearn' when required to. The researcher can also have further access to the thinking processes during the teaching process through discussing with the subject. (For details see Law & Ki, 1987 (appendix 1) and Law, Ogborn & Whitelock, 1988 (appendix 2)).

One possible objection to employing such a technique for the present investigation is that the way such a computer system handles knowledge is very different from what is commonly believed about how human beings operate in commonsense everyday contexts. The computer here only accepts knowledge in the form of rational logical statements, and it treats the complete set of knowledge statements entered as one formal system. On the other hand, commonsense ideas are believed to be fuzzy, possibly fragmented, pragmatic and possibly not logical.

The above objection essentially boils down to the fact that this mode of elicitation imposes a particular format onto the elicited knowledge. This objection, though valid, is true also for other techniques employed in the exploration of children's ideas in science. For example, concept map construction forces subjects to express their ideas in the form of networks of nodes and links. The particular ideas elicited via repertory grid technique may be influenced by the actual elements chosen, and the elicitations tend to highlight contrasts amongst the elements. On possible advantage of the expert system development technique (with the metaphor of teaching the computer) over many of the other techniques in this respect is that it provides an environment for the explorations of the dynamic aspects of the externalized knowledge. Thus while subjects may not be aware of the consequences (if any) of a concept map in terms of its utility, or the semantics of the network formalism itself, subjects developing the expert systems can find out whether they think the computer actually reasons the same way as they themselves do (or as they expect it to), and the researcher can have access to their reactions concerning this.

Because of the advantages of expert system development as a knowledge elicitation technique listed above, it was employed as the main method of investigation in the present research. However, because of its novelty, it would be helpful if another more conventional technique is employed to provide some form of cross-referencing. An interview based on a classification task was used to provide some general perspectives on how students reason generally about motions. Results from the interviews will be analyzed independently in their own right, as well as providing starting ideas for the expert system development task. Details about these two tasks are described in the following two chapters.

Chapter 5 THE CLASSIFICATION TASK

5.1 Objectives of the task

As was discussed in the last chapter, none of the commonly used techniques in the elicitation of children's ideas in science is entirely satisfactory for the purpose of exploring the commonsense reasoning of children in the area of mechanics, though some of the techniques have certain attractive features. In a review on the different methods used in the documentation of conceptual frameworks, Driver and Erickson (1983) placed the different methods employed by researchers in this area along a conceptual-phenomenological continuum. Conceptually framed tasks like word associations and concept mapping were obviously unsuitable as the present study was not interested in finding out what students understand by certain conceptual labels. Contextual tasks like rule assessment and observational studies of students' theories-in-action are designed to probe the commonsense spontaneous knowledge of students, but are very much tied to the operational aspects of the specific situation under investigation. The aim here is to explore the in depth reasoning used by children without biasing them by pre-determining the concepts/terms with which they have to use to express their reasoning. Thus, to use the conceptual/ phenomenological distinction in the classification of techniques, the elicitation strategy needs to start from open discussions about phenomena to elicit the concepts/ terms subjects naturally use in their everyday commonsense reasoning about motion, which can be used as starting points to explore into the detail reasoning they use.

It was decided that deep case studies would be the appropriate level to work at for the purpose of the present research and a programming task (to be described in the next chapter) was chosen as the main elicitation task for exploring into students' commonsense reasoning about motion. However, there is a need to do another task prior to the programming task to provide some initial exploration into the terms and features about motion that figure prominently in commonsense reasoning. This would form a worthwhile investigation in its own right as well as feeding useful information for the execution of the programming task.

This chapter reports on a classification task administered for the purposes just described. Its objectives are to find out the perspectives taken and parameters looked for by children in their daily perception of motions. The design for this task is a modification of the repertory grid technique. Instead of asking subjects to dichotomize a set of objects, they are free to put them into two or more groups according to a single criterion. The specific research questions for this study are:

What are the main issues of interest for children in their everyday perception of motion?

Can we identify a common pattern or 'framework' of perception employed by the students?

Does age or amount of physics instruction received affect a child's everyday perception of motion?

5.2 Administration of the task

The task. The method of investigation used in this part of the study is a classification task. Each subject was shown nine to ten pictures depicting different kinds of motion and told to put the pictures into groups according to one consistent criterion. After the classification was completed, the subject was asked to explain how the classification was done until the interviewer was satisfied that she understood the detailed workings of the criterion used. The subject was then asked to re-group the same pictures again using a different criterion. This process was repeated until no further different classification could be done. The whole process was audio-taped and the recording was translated and transcribed for analysis (the interviews were conducted in Chinese).

Materials. The pictures used in the classification task were all taken from children's comic books. This ensures that the situations depicted were all easily comprehensible everyday situations of interest to children, and that they would not resemble textbook examples and unduly direct them onto using learnt terms and concepts. Further, the performance of a classification task does not require in depth reasoning or analysis about any particular situation but rather demands a certain degree of generalization across situations.

In order to find out whether the outcome of the investigation depends on the particular situations depicted in the pictures, three different sets of pictures (Appendix 4) were used in this classification task. Set A were all pictures of 'bouncing' motions where the moving objects follow a more or less parabolic trajectory, set B were pictures of motions in the air, and set C were pictures of either falling/tripping motions or motions in water. Each student was only required to categorize one set of pictures and any systematic difference in the resulting categorization schemes due to the choice of pictures could then be detected.

Subjects. As one important point of interest for the present investigation is to find out whether or not the commonsense frameworks of perception employed by the subjects depend on their age and on the amount of physics instruction they have received, two groups of students were used as subjects. The first group of eleven Form six students were about 17 years old, all of whom had completed an O-Level course on Physics and were receiving instructions in A-Level Physics in their school. These eleven students were chosen so as to represent different levels of Physics attainment in their O-Level examination results. The second group of twelve Form three students were about 15 years old and had not yet received any formal instructions in mechanics. They were selected to represent a range of different overall academic attainment according to their order of merit score obtained in the last school examination.

5.3 Classification as a cognitive task

The nature of this classification task is essentially one of categorizing motion events. The main reason for choosing a categorization task is that as a cognitive activity, it is somewhere between perception and thinking (Neisser, 1987). Gibson (1979) described perception as 'resonating' to the invariants in the objectively existing information, giving rise to the direct recognition of the real state of affairs. Categorization, whilst being often a spontaneous act, goes beyond the direct perception of objective information and rests on our beliefs about the world. Keil and Batterman's work on developmental change in categorization behaviour (1984) showed that there is a characteristic-to-defining shift in the transition from naivete to expertise at any age. He also demonstrated that such a shift is related to a change in the subject's beliefs about the nature of the entities being categorized (Keil, 1986). Furthermore, though such underlying theories and beliefs can easily be inferred unambiguously, they are very often held tacitly by the subjects themselves. This is considered to be an appropriate level to start the exploration.

Rosch (1978) proposed that there are two general and basic principles underlying the formation of categories. The first is the principle of *cognitive economy* and is concerned with the function of category systems. It asserts that the task of category systems is to provide maximum information with the least cognitive effort. This is in line with the commonsense notion that, as an organism, what one wishes to gain from one's categories is a great deal of information about the environment while conserving finite resources as much as possible. To categorize a stimulus thus means to consider it, for the purposes of that categorization, not only equivalent to other stimuli in the same category but also different from stimuli not in that category. The second is the principle of *perceived world structure* and is concerned with the structure of the information so provided. It asserts that unlike the sets of stimuli used in traditional laboratory tasks, the perceived world is not an unstructured total set of equi-probable co-occurring attributes but rather comes as structured information. Combinations of what we perceive as the attributes of real objects do not occur uniformly: some pairs, triples, etc. are quite probable; others are rare; others logically cannot or empirically do not occur. However, the perceived world structure is not just the metaphysical world without a knower: the kinds of attributes that can be perceived are not only species-specific, but also culturally dependent.

According to these two principles, it is reasonable to expect that the schemes used in the categorization of motion events would reflect the kinds of general questions children ask when encountering motion events in everyday life, the features they judge as important and the parameters they would wish to be able to predict in such situations.

There is little in the literature about studies in the actual categorization of events, but a number of studies pointed out the relationship between the categorization of objects and the perception of events (Nelson, 1983; Fivush, 1987). Some studies indicate that young children have a preference for grouping objects together on a functional or thematic basis (Denney & Moulton 1976; Mandler 1979). Thus candles, cakes and party hats go together because they belong to the same event: the birthday party. This is a possible indication for the primacy of event representations in the conceptual development of children, and that thematic and functional categories have their basis in children's representation of events. Furthermore, experiments on children's memory performance (Lucariello & Nelson, 1985) show that hierarchical categorical organization is not readily available to young children, and that categorical structures possibly emerge and develop from a conceptual knowledge base initially organized schematically around items that can be used interchangeably for the same function in similar events. It is hoped that the present study could also contribute to an understanding of events as a unit of cognitive/knowledge structure.

5.4 Scheme of analysis

Both groups of students found it easy to categorize the pictures. All of them could find more than one possible categorization scheme for classifying the set of pictures given to them, though some failed to always adhere to one single criterion in a particular attempt. An analysis of the detailed contents of the categorization schemes reveals that they can be labelled as either descriptive or causal. Descriptive schemes are those that classify pictures according to physical features of the motion, e.g. descriptions about the moving object, the shape of the trajectory, speed of the motion, how does the motion develop, etc. Causal schemes are those that classify motions according to analyses of the causes, e.g. the agents involved, the forces causing the motion, etc. Descriptive schemes are thus schemes that attempt to answer the '*what*' and '*how*' questions about the motions while causal schemes are those that attempt to answer the '*why*' questions. The former are thus concerned with issues roughly within the domain of kinematics while the latter deals with dynamical issues of motion.

It is possible to further subdivide each of these two broad groups of schemes into three different levels of categorization. The descriptive schemes can be classified as: (i) D1 - non-physical, pragmatic concerns, (ii) D2 - physical and holistic, or (iii) D3 - single physical parameters. The causal schemes can be classified as: (i) A1 -causal agency (concrete or abstract), (ii) A2 -characteristics of causal agency, or (iii) A3 - mechanism for producing the motion. A detailed list of the categorization schemes belonging to each of the above subdivisions is given in table 5.1.

It can be argued that in doing this classification task, the categorization schemes the students use actually derive from the kinds of questions they would naturally ask themselves about motions as they encounter them in everyday life. Close examinations of the schemes should provide crucial insight into what are the most important questions to be considered in commonsense perceptions of motion.

5.4.1 Descriptive Schemes

Of the 37 different categorization schemes identified, about two thirds are descriptive ones. These are roughly evenly distributed amongst the groups D1, D2 and D3. This shows that non-physical, pragmatic concerns like purpose of the motion, intention of the moving objects and the social situation under which the motion occurred are as important as considerations of the physical aspects of the motion even for the group of Advanced Level Physics students. (However, such pragmatic concerns rarely come into classroom discussions during mechanics lessons.)

Another interesting observation is that of the schemes concerned with physical descriptions of the motions, about half were concerned with global descriptions, mostly in the form of assigning global labels/identifiers to the motions. Some of the global identifiers used were labels for 'types of motion', i.e. words we use in everyday language for naming different kinds of motion, like: bouncing, falling, flying, slipping, tripping, swinging and rising. These are similar to the stereotypical motions identified by Ogborn, Bliss & Whitelock (Ogborn, 1989). From the students' elaborations in the classification protocols on how the

Table 5.1 List of all categorization schemes grouped under subdivision titles

CATEGORIZATION SCHEME	INSTANCES F.3	F.6
I. Descriptive schemes [D1]		
(i) pragmatic concerns, non-physical aspects		
* intention of moving object	2	3
* ability of object to control the motion	1	2
* social situation in which the motion occurred	2	1
* motivation for the object to perform the action	1	0
* feelings of the objects involved in the motion	1	0
* moving object's reactions to the situation	1	0
* suitability of object's orientation for reacting to the situation	1	0
* purpose of the motion	0	1
* whether the pictures together form a tangible story	0	1
(ii) physical and holistic [D2]		
* type of motion	6	7
* shape of trajectory	0	4
* motion direction	1	2
* type of position change	0	2
* likelihood of event happening	1	1
* expected development of event	1	0
* change in type of motion?	0	1
(iii) Single physical parameter [D3]		
* nature of the moving obj.,e.g. animate/inanimate	3	1
* speed	1	2
* motion involves a turning of the object?	1	1
* change of speed during motion?	0	1
* shape of object	0	1
* height reached by object during motion	1	0
* duration of motion	1	0
* location of motion	0	1
II. Causal Schemes		
(i) Causal agency [A1]		
* concrete causal agent	2	1
* abstract causal agent, e.g. forces involved	1	2
* agency for change in motion	0	1
(ii) Characteristics of the causal agency [A2]		
* type of agent, e.g. self, ext. obj., ext. forces	5	4
* causal agent also moving?	0	1
* object carried along by agent?	0	1
* agency required for the motion?	0	1
* number of forces involved	0	1
(iii) Mechanism for producing the motion [A3]		
* analogical/theoretical model of the motion	3	2
* how motion occurred in relation to gravity	2	2
* motion resulting from diff. rel. dir. of forces	0	1
* motion resulting from interaction of forces	0	1
* effect of the phy. configuration of system	0	1
* effect of own action on motion	0	1

categorization schemes worked, it was evident that these words carry very rich meaning for them. They seem to stand for mental representations of prototypical motion schemas stored in their long term memory. The basic contents of such schemas would include the following items: likely instantiations of the kind of object involved, location of the motion, shape of the trajectory, agents for the motion and likely course of development.

The extract from student S's classification task protocol below illustrates the rich meaning behind one such label used - 'bouncing':

S. (about the group of pictures 7,10) ... I find that things seem to be bouncing up and down .....

(about the group of pictures 2,3,4,8) ...something is being thrown out. There is a person ... some object which give it a force by which the object is ejected or thrown out.
.....

I. What about the group of pictures 6 and 9 (the group she classified as jumping earlier?

S ......These two seem to be done by themselves. Seems to be more natural.

I. More natural? I do not understand.

S. I mean, they do it themselves and not that somebody makes them do it. Things happen and things will naturally develop in that way accordingly. ...... for the horse, since there is a fence in front, it will jump over it.

.....

I. How about the group of pictures 1,5?

S ..... It seems that it meets something and then it bumps outwards. At first, it appears normal but then it suddenly meets something, crashes or hits on something, then it has another kind of motion ... that is, it should be in a sequence, but something makes it change.

This example also shows that during the process of categorization, the students were actually trying to map each of the motions in the pictures onto a typical motion prototype. In other words, they were trying to place them under some already existing representation schemas. Furthermore, such schemas are higher level, holistic units roughly at the level of complexity similar to Rumelhart and Ortony's (1977) schemas, Schank and Abelson's (1977) scripts and Minsky's (1975) frames. The protocols collected during the classification task seem to reveal that each representation schema comprises a set of parameters (or slots) that may be instantiated differently according to the kinds of motion involved: the kind(s) of object involved, location for the motion, shape of the trajectory, direction of motion, likely subsequent development of the trajectory, whether the motion involved sudden changes from an expected course of development, probable causes for the motion and agents for the motion. Categorization schemes using global identifiers as described above can be interpreted as categorizing according to labels/names for the motion schemas.

These representation schemas as inferred knowledge structures will be referred to as 'scripts' in the ensuing discussion to avoid confusion with the actual classification schemes used by the students. There is, however, no intention that these 'scripts' are to be interpreted strictly in the way Schank defines them.

Some global physical categorization schemes do not use 'types of motion' labels for the scripts. Instead, certain labels for single physical parameters like trajectory shapes, motion directions and types of positional change involved were named as categorization criteria. Thus at first sight, these schemes look like D3 schemes - i.e. they look as though the categorization was made according to the value of a single physical parameter. However, careful inspection reveals that in many cases this was not true and that these parameters were actually used as script labels for global categorization schemes. For example, student BS, when categorizing the picture set B (App.4), used 'direction of motion' as the categorization criterion in one of her schemes. Pictures B1, B2 and B5 were put into one group - the group of horizontal motions. However, B1 was a plane moving in the air and B2 was a bird flying and it was not evident from the pictures that they were moving horizontally. B5 was about a ball being thrown out by a boy and the expected path should be a parabolic one. BS's explanation for putting them into this group was:

> 'In B5, it is pushed out, a horizontal motion. In B1 and B2, they use their own force to rush forward. So they are similar.'

It thus seems that she put these three together because they were all 'pushing' motions and not because of their actual motion directions. Similarly, in the same scheme, pictures B3, B4 and B8 were put into the group of downwards motions but only B3's motion was strictly vertically downward motion. The motions in the other two pictures had parabolic paths. BS explained:

> '....(B4 & B8) have instruments to help them to spring up, and they sprang up before coming down. ... They go up and down, not horizontally. In 3, the motion is downwards, so I place it together with the one involving the planks.'

However, the boy in picture B6 was also moving up and down - and doing so repeatedly as illustrated. She placed it in the group of 'rising' motions. The actual criterion used for the downwards motions group was probably that they were all motions that eventually went down - they were 'falling' motions.

Some categorization schemes belonging to the holistic physical descriptive type (D2) were not concerned with assigning script labels. Rather, they were concerned with expectations about the motion, e.g. likelihood of the event happening, the expected course of development of the motion, and whether the picture depicted a sudden change in the motion type. This seems to reflect a very pragmatic concern in our everyday encounters with motion - our cognitive system must be able to formulate anticipations about the likely development of the event very quickly and efficiently so as to be able to react adequately to the situation. If the motion is classified as a 'normal' or 'natural' motion (many students use these as category labels), this means that the motion falls into one of those stereotypical motions for which there are ready, packaged scripts that hold default values for the various slots, including possibly one on the optimum reactions appropriate for dealing with the type of motion at hand. Thus the cognitive resources required for the handling of the event in such cases are minimal. This is in line with Minsky's (1975) justification for looking for larger and more complex units of representation than semantic nets - to account for the effectiveness of commonsense thought. On the other hand, if there is a sudden change in the expected development of the motion, the cognitive system would have to go into an alerted state and make greater efforts to make sense of the new situation and to formulate appropriate reactions. Here, we can actually see the inherent relationships between the D1 and the D2 schemes: the D2 schemes reflect our need to ask the 'what' questions (what is the motion like?) in order to formulate answers to the 'how' questions (how should one deal with the situation for the purposes in hand, i.e. the current context?). It is a reasonable expectation that our long term memory representation schema (or scripts) for the motion events should be able to handle the pragmatic concerns of the 'how' questions.

Another indication of the importance of being able to give a holistic description to the motion by way of cognitive understanding of the phenomena is the fact that only 5 out of the 23 subjects do not use at least one D2 type scheme in all their categorizations and 10 subjects used it as the first categorization scheme.

The third kind of description based categorization is those that categorize according to the instantiations (i.e. expected or default values) for one of the slots (parameters) in the motion scripts as described above. The most popular categorizations in this group are those that classify according to the nature of the moving object or according to the speed of the motion. Other less popular categorization parameters are: whether the object spins during the motion, whether there was a change in speed, the shape of the object, maximum height reached by the object, duration of the motion and the location of the motion. These physical parameters correspond to the slots that appear in the script representations for the motion events and thus represent important aspects to pay attention to in one's everyday encounters with motion.

It is interesting to note that the most popular slot for categorization concerns the animacy of the moving object (whether it is animate or not, etc.) rather than parameters relating to the motion itself. If the main concerns of everyday cognition are pragmatic, then the kind of moving object and the rough magnitude of the speed are possibly the most crucial aspects one has to take into account when considering how best to react to the situation. For example, one's reactions on seeing a baby fall from a table would be quite different from the reactions on seeing a parcel of similar size and weight falling from the same place. Such differences however, do not figure in the Newtonian view of motion where only the physical attributes (like weight, density distribution, size) of the moving object would have any effect on the motion and the animacy of the moving object is irrelevant - an object would still be falling with the same acceleration whether it is a stone or a person.

The reason for some students to give serious attention to the animacy of the moving object can be explained by the following extract from Student Y's protocol when he tried to explain the difference between the motion of dead (inanimate) objects and living (animate) objects:

> "Living objects can control their own movement but the dead objects are controlled by man. ..... [For the inanimate group] there is no specific rule to change their own motion. Like (pointing to picture C5) .. he pushed the ball, the ball has to follow the force to move, then the force decrease and it fell down following a parabolic path. ....
> The living object can generate force but the dead object can't."

Thus one main difference is in the ability to control the course of development. Inanimate objects cannot have control over their own motion and one would therefore be able to anticipate what the motion would be like unless some animate objects intervene.

Another reason for considering the nature of the moving object is to facilitate decisions about how best to handle the situation. For example, one of Student H's categorization schemes (used on picture set C) was prompted by such considerations. Picture set C comprised 10 pictures, 7 of which are about some kind of falling or tripping motion while the rest are objects moving on the water surface. Student H put the pictures into three groups: the first group are those where the moving objects are doing something to protect themselves against possible damage arising from the motion, the second group are those that could not protect themselves because of the circumstances of the situation and for the third group, the question of protection is irrelevant. Clearly, the question of trying to do something to protect oneself from danger would be irrelevant only for inanimate objects. Moreover, the same kind of motion, falling, would be classified into different groups under this pragmatic concern of safety. To decide on what one has (or does not have) to do when confronted by a moving object, one needs to know not only what type of motion is involved but in many cases also the kind of object involved. Thus the importance of this 'nature of the moving object' slot in the representation schema is probably not for its role in identifying the script type but for possible instantiations of the expectation slot and for identifying the reaction required to handle the situation.

5.4.2 Causal Schemes

As mentioned earlier, one third of the identified categorization schemes classified pictures according to analyses of the causes. These analytic schemes seem to be answering questions about why the object/system was moving the way it did and most of these schemes were about some aspects of causal agency.

In the A1 group of categorization schemes, students classify pictures into groups each involving a different causal agent. The agencies so named may be abstract, e.g. naming the forces responsible, or concrete, naming the actual object causing the motion. This seems to be the lowest possible level of explanation that can be offered - one tries to explain away something by simply pointing out that this is the kind of motion that would result in instances where the same agencies are involved. For example, student L categorized picture set C into 3 groups according to the causes involved: 'smoothness', 'obstructions' or 'wetness'. The following extract from her protocol illustrates her reasoning.

> 'This group (pictures 1,2,6) they fell down, they ... the floor surface is very smooth. The supporting force is not very good, ... it is smooth, so it is easy to slide over. ...
>
> This group (pictures 4,5,7) is just the opposite, there is something blocking you, obstructing you.
>
> The last group, my feeling is that they are related to water. Three of these pictures got water in them. ... I would think that water is even more smooth (making it easier for things to fall through).'

This explanation for the categorization scheme in terms of causal agency does not offer further analysis of the motions other than what a D2 type scheme in terms of types of motion offers. However, this act of labelling the cause of motion as a prominent feature of the situation reflects the cognitive desire to look for causal relations in motion events. That is, there is a cognitive desire to seek *explanations* for why a motion is developing the way it is when one meets it. Schank's (1986) work on the nature and kinds of explanations provides a helpful framework for understanding the causal schemes generated by the students. He proposes that there are three types of explanations. The first kind are those that are made for the sake of others, to tell them what one already knows. Such explanations are almost always present in the mind of the explainer before the explanations are given and he calls those *canned explanations.* The second kind are those intended to explain-away a phenomenon by showing it to be another example of something that is already well-understood, and he calls these *explaining-away explanations.* The third kind of explanations are those that are intended to add knowledge by the process of explanation and are thus critical to learning, and he calls them *additive explanations.* He believes that the first two kinds are much easier to generate and that an explanation is likely to be additive only if something important rests on the creation of an explanation. Thus when no real need is driving an explanation, it is likely that explanations created will have as their primary intention to explain away the phenomenon. In other words, we do not try hard to understand deeply unless there is a strong need to understand.

When seen within this framework, explanation in terms of naming the causal agencies concerned does satisfy the explanation need to the level of explaining-away the problem. Abstract constructs can also be used as labels to explain away the causes for the motions, and the favorite constructs in this category are adaptations from scientific terms - labelling (very often creating them at the same time) various kinds of forces as the responsible agencies. For example, student C, in classifying the same picture set C, put them into three groups. The following is an extract from the protocol:

> C. The group 1,2,3,5, they all involve the kind of force when things go from a high position to a low position, that type of force that is going downwards.
>
> ....
>
> Yes, that's it, why does it fall down? It is because the ground has force of gravity and the object has weight-force, when it is in mid-air with nothing attached, then it would fall down. This is the force going from high to low .. is this called 'falling force'? ....
>
> I asked her to explain the criterion for putting 4,6,7 together.
>
> C. This group are all about being slipped or tripped by something. All because while walking or moving, you meet an obstructing force from outside and tripped. ....
>
> The last group of pictures was 8,9 and 10 and she said they belong together because they were all about 'floating force, all about water'.

Again this categorization here is not too different from a D2 scheme that classifies the pictures into downward motions, slipping or tripping motions, and motions in water. The only difference is the application of a 'canned explanation' (Schank, 1986) to the motions at hand. The 'canned explanation' is in the form: motion is caused by force, motions are different from each other if the forces involved are different. The quality/nature of the explanation does not seem to be different whether concrete or abstract agency is used and a few of the A1 schemes actually employ both kinds of agency for the groups.

The categorization behaviour of the students seems to demonstrate a strong desire to give causal explanations for motion phenomena. This desire is evidenced by the fact that there are only 3 out of the 23 students participating in this task who did not offer any analytic categorization scheme. Thus, adhering to the earlier assumption that to make sense of a motion one encounters in everyday life, one needs to recognize the type of motion involved (to search for an appropriate event script), the above observation points to the primacy of causal agency as an essential slot in the memory representation of motion.

One of the A1 type categorization schemes deserves special mention. Instead of looking for the agents causing the motion, Student V categorized the pictures according to the agents responsible for the *change in course of motion*. Though she still grouped them according to the concrete agents involved, it was an extremely important step forward to look for (even when performing such a casual, non-academic task) agency for change rather than the agency for the maintenance of a motion. Her implicit classification of motion is not Newtonian: rather than using a uniform motion - acceleration motion distinction, she distinguished between motions that followed a natural course of development and those that had a sudden change in the course of motion. This classification highlights the kinematic observation that an object's trajectory would be completely defined if the initial conditions and acceleration are defined. According to her explanation, a moving object can only change its course of motion (which need not be a uniform motion) by an *external agency*, and this is valid even if the moving object is animate. The following extract from her protocol illustrates her reasoning:

> V. In B5, a child throws a ball out and I feel that the ball once thrown out cannot change its own motion. It cannot increase its speed or change its direction... Once it flies out, then it has to fall according to the law of nature.
>
> ....
>
> In B1,B,2,B3 & B9, what is peculiar is the use of air. Without air, parachutes cannot come down and kites, planes and birds cannot fly. .... They can control the direction of inclination, to change the course of motion. ....
>
> The parachute can be controlled by the person carried on it. .... As for the kite, though this child is pulled upwards suddenly, not of his own accord, he can still, like people gliding in the sky, by the movement of his body control the direction of flying. The common point here is that, they can only fly with the help of the air current.

I. asked her if the dog in B4 can control its own course of motion.

V. It can when it is not yet thrown up, but once it is thrown up it cannot. .... It can only have action but nothing can give it a reaction. Unless it knows how to unbuckle its neck belt and to throw it away, then because there is an action, it would have a reaction. ....'

It can be seen that she was trying to adopt Newton's first and third laws of motion and was applying them correctly and creatively to these common everyday motions. It is interesting to note that in this process, while still recognizing the difference between animate and inanimate objects in their motions, she confined that distinction to the ability of animate objects to 'control' parts of their own body, but whether this could lead to a change in the course of the motion would still be subjected to the same physical laws as inanimate objects. She has succeeded in mastering the scientific laws and principles to the extent that she can recognize the general applicability of these laws to the motion of both animate and inanimate objects, while still being able to explain for the differences that exist between them using the same principles. This is a rare attainment and she is the only one who has gone so far in clarifying such views in this classification task.

The second kind of analytic schemes (A2) categorizes motion according to certain perceived characteristics of the causal agency. The most popular A2 type scheme (accounting for 9 out of the 13 instances, and is the largest single categorization within the analytic schemes) was to categorize according to the nature of the causal agent, distinguishing mainly between self and external objects/forces as causal agency, or sometimes between animate and inanimate causal agents. For the latter categorization, the reasoning is actually the same as that used in the D3 scheme which categorizes according to the nature of the moving object. The main focus here was again on whether the motion can be changed at will. Student J (categorizing picture set B) explained this point very clearly in the following extract:

'For the smaller group (pictures B1, B7 and B9), the motions are caused by general factors like wind, the power of the balloon or the movement of the machine that caused the motions. ......

The motions (for the other pictures) are acting according to his own will. .... There is the idea of an animal. .... The action can change whenever he likes because he can do whatever he like. But it is difficult in the case of inanimate things, they can't change voluntarily.'

This is essentially the same as the reasoning given by student Y quoted above.

For those categorizing causal agents into self and external agents/forces, the motivation is two-fold. There is a pragmatic concern of whether the motion occurred according to the intention/wishes of the moving object. In this respect it is similar to the concerns of the most popular D1 type scheme, 'intention of the moving object', and as such is also related to the D3 scheme that classifies motion according to the animacy of the moving object. Another motivation behind this categorization was to predict possible changes in the subsequent motion. For example, extracts from Student T's protocol below show that the motions that are grouped as those assisted by external agencies have motion characteristics that would otherwise be unexpected. The other two kinds of motions in this categorization are actually self-initiated motions and motions due to 'natural circumstances'. The need to identify causes of motion was in line with the pragmatic need to identify the type of motion script (as exhibited also in the D2 categorizations) and to be able to anticipate subsequent developments of the motion encountered. She named 'potential energy' as the reason for motion in the third group, but actually 'potential energy' was only used as a label.

> T. (About 1,2,4,7,9,10) .... the actions are all assisted. Say, this man in the chair (picture C9), he would not move by himself, it is the boat, the string and the chair all put together that assisted him to have this motion.
>
> I.asked T. to explain further the role of the external 'assistance'.
>
> T. You can just stand there and move your arms about. Yes, you can swim, but you can't move in the same posture with the same movement as he is now.
> ....
> (About 6 and 8) These are actions coming out from themselves.
> ....
> (about 3 and 5) They themselves have the ability to move. They have the ability to move, they have the potential energy, ... anyway they can move.

Categorization schemes in the A3 group go into deeper levels of causal analysis. Though six categorization schemes were identified, there were basically only two forms of explanation: causal explanation in the form of analogical models or as the aggregate effect of forces.

Schemes in the first category attempt to offer analogical/operational models for the underlying mechanisms responsible for bringing about the motions. The schemes normally identify certain physical components as playing the crucial causal role which is analogous to certain common mechanical devices, e.g. springs, obstacles, levers. Some subjects went even further to build up theoretical models of the kinds of forces at work in such situations. For example, Student W theorized that some objects bounce up and down repeatedly on hitting obstacles because they possess 'elastic force':

> 'Objects (that can bounce after falling to the ground are those) that have the ability to expand and contract, the object possesses elastic force. .... Just like the ball, there is elastic force inside and it is not that heavy. ...... It (the object that possesses elastic force) can receive forces. If it is just placed

there, it can't move, but if you hit it, then because of your force, it touches the ground, so it bounces.'

Thus an object that possesses elastic force can receive a force and re-use it time and again even after hitting obstacles while another object that does not possess elastic force will stop after hitting an obstacle. Elastic force does not enable an object to start a motion, but would affect its subsequent motion once it has been started by receiving a force. Here a property of the material of the object is also classified as a kind of force. A possible reason is that the literal translation of elasticity in the Chinese language is 'elastic force'.

Another example of this kind of explanatory theory proposed was the 'transmission of rushing (impulsive) force'. Objects moving at great speed possess 'rushing force' which may be passed onto other objects during collisions. However, some motions may require an intermediate device to transmit this force from the causal agent to the other object (like in the case of lever systems).

The second form of explanation was built on the basic belief that motions occur in the direction of the aggregate effect of the forces acting. How exactly the aggregation takes place is not clear if the forces do not act along the same line, and rotation seems to demand additional turning forces. If however, the forces are acting in opposite directions, then a model of contention (the strongest force will rule) is frequently adopted so that the larger force '*wins*' and the object then moves as a result of the winning force and the weaker forces would fail to produce any effect unless the balance of forces changes. Furthermore, gravitational force is seen to be an all pervasive force and would cause objects to fall unless there is a large enough contending force to act against it. Student KC's protocol (while categorizing picture set B) illustrates this kind of reasoning:

"KC. For the group 4,5,8, they only stay in the air for a short period of time. They will fall down very soon. ..... Because of gravity...... (The other group can stay for a long period of time)because there is something that resist against gravity. ...
I. How about the bird? How can the bird resist against gravity?
KC. It can fly by itself."

One of the A3 categorization schemes stood out uniquely against the others both in terms of the question it poses and in the solution it gives. In one of her schemes, Student F classified picture set C according to whether the motion will be affected by actions on the part of the moving object.

'The first group is 8 and 10. For these two, in order to continue their motions, the people here have to move part of their body in a different way......
Whereas in the other group ..... But if he (the man in picture C9) wave his

arms about, it would still not affect his motion. The same situation for this man who is falling, and the same for the rest of the pictures.'

What is even more remarkable is his comment on the classification task as a whole:

F. I feel that quite a lot of my classifications are very vague (he did 7 different categorization schemes in all). They look different but as you think more clearly about it, they are more or less the same.

I. Can you give an example?

F. The nature of their motion does not really differ much. Say in my second way of classification, the one about force, ... actually, it should be that anything that moves does so because it is pushed by other forces. If you say that you want to push a force out, that is not very reasonable. When you push something, the other thing also has a reaction and so pushes it back. So it seems that a lot of the classifications are very superficial, and also very vague.....

I. Do you feel that you can cancel the second set of classifications?

F. Not exactly. It still has its value.

I. Where does this value lie? As you said, everything depends on some external force.

F. It depends on how you look at it. From a very fundamental perspective, most motions are of the same type. But you can go to a more complex level and classify it further into different motions. For example, turning is one type of motion, or linear motions is another. You can say about anything that if you push it, it would move. So if you say that anything that moves when pushed is in one group, then you would only end up with one group. But if you want to do it in greater detail you can subdivide it much further into many groups.

I. Do you find a more detailed classification meaningful? Just now you said that they belong very much together.

F. Yes. When things move in different ways, you would wish to know their past movement and how their motion would develop in future, then this is useful. But if you ask why they move, they are more or less the same.'

I. For the pictures 10 and 8 you said that he moved his arms in order to maintain this situation, while for the other group they do not rely on themselves. Is this distinction important?

F. Yes. It means that one can use one's own ways to influence one's own motion. Say, in everyday life, when you are running, you can use more efforts to run quicker.

One striking feature coming out from the above protocol is that Student F recognizes that all motion are caused by external forces, and that the dynamics for different motions are the same, the main difference is only at the kinematic level.

5.5 Results

5.5.1 Comparison between the two age groups

Despite the fact that the three sets of pictures depicted very different kinds of motions, there seems to be no significant difference between the picture sets in terms of the total number of categorization schemes that each student can come up with. However, the older group of students consistently produced larger numbers of different categorization schemes. On average the older group of students produced 4.6 categorization schemes each, nearly 50% more than the younger group which only produced an average of 3.1 schemes each. One possible explanation is that the older group of students employs different (and possibly complex and differentiated) cognitive schemas in their processing of information about motion events. Another possibility is that the older students have a broader knowledge base about motion and can thus come up with more ways of categorizing them.

Table 5.2 Total no. of categorization schemes resulting from the classification task for each group of students for the three sets of pictures

Picture Set	F.6		F.3		Total for picture set	
Set A*	14	(4.7)#	10	(2.5)	24	(3.4)
Set B	19	(4.8)	14	(3.5)	33	(4.1)
Set C	19	(4.5)	14	(3.5)	33	(4.1)
Total	52	(4.7)	38	(3.2)	90	3.9)

The figures inside brackets denotes the average number of categorization scheme per student

* There are only three students in this group while there are four in all the other groups

An examination of the distribution of categorization schemes over the different scheme types for the two age groups reveals where the main differences lie. As Table 5.3 indicates, not only is the ratio of descriptive schemes to analytic schemes essentially the same, but even the percentage distribution for the six finer scheme types are comparable for the two age groups. Thus there is reason to believe that the structure of the cognitive schemas for the processing of motion events are very similar for both groups. The largest difference in the number of instances of occurrence is in scheme type D2. If we try to examine the D2 schemes as listed in table 5.1, we can see that most of the subjects in the younger group could only give a holistic physical description in terms of giving a label to the 'type of motion'. However, for the older group, 'type of motion' schemes accounted for only 7 out of a total of 17 D2 schemes resulting from the categorization task. Parameters like trajectory shape, motion direction and type of positional change are familiar to the older group, as these are the kind of parameters that one would expect them to use in their study of mechanics in school, but these descriptions do not come to the younger group naturally. A frequency count from table 5.1 shows that of the 38 different schemes resulting from the categorization task, the F.6 students came up with 31 of these schemes whereas the F.3 could only come up with 21 different schemes, giving a similar ratio as the total number of categorizations. As discussed in the previous section, schemes within the same type normally reflect similar concerns and perspectives and the differences between them are usually nominal. Thus a reasonable conclusion from the figures is that the older students have more means to elaborate and express their ideas than the younger group, but there do not seem to be fundamental differences in their cognitive schemas.

Table 5.3 Distribution of classificatory categories over the scheme types for the two age groups

	Descriptive			Sub-total	Analytic			Sub-total	Total
	D1	D2	D3		A1	A2	A3		
F.3	9 24%	9 24%	7 18%	25 66%	3 8%	5 13%	5 13%	13 34%	38 100%
F.6	8 15%	17 33%	7 13%	32 62%	4 8%	8 15%	8 15%	20 38%	52 100%
Both groups	17 19%	26 29%	14 16%	57 63%	7 8%	13 14%	13 14%	33 37%	90 100%

5.5.2 Comparison of performance between the two sexes

There is no major difference in the performance of the categorization task between the two sexes, averaging 3.8 and 4.1 categorizations per student respectively for the female and male students. However, careful scrutiny shows that female students produce many more D1 type schemes than male students while male students produce a much higher percentage of A3 schemes. The more humanistic categorizations in the D1 type schemes are nearly all produced by the girls, e.g. the categorization criteria like feelings of the objects involved in the motion, the motivation for the motion, the social situation in which the motion occurred

Table 5.4 Distribution of classificatory categories by sex

	Descriptive			Sub-total	Analytic			Sub-total	Total
	D1	D2	D3		A1	A2	A3		
F	12 24%	14 29%	7 14%	33 67%	4 8%	7 14%	5 10%	16 33%	49 100%
M	5 12%	12 29%	7 17%	24 59%	3 7%	6 15%	8 20%	17 41%	41 100%
Both sexes	17 19%	26 29%	14 16%	57 64%	7 8%	13 14%	13 14%	33 36%	90 100%

and whether the pictures together form a tangible story. These categorizations probably could not be produced just on the factual information given in the pictures alone but were possibly done on imagined social situations, possibilities extending from the pictures. Male students on the other hand seem to exhibit more interest in working out mechanical/ analogical models for the motions depicted.

5.5.3 Effects of picture set used on categorization behaviour

In analyzing the results of this classification task, one obvious question to ask is whether the categorization behaviour of the students is dependent on the picture set used. It is not possible to give a definite answer to this question from the results of the present research as the number of cases involved are relatively small and no attempt has been made to make the three groups of subjects comparable (table 5.5).

Table 5.5 Age group and sex distribution of subjects by picture set

Picture set	F.6		F.3	
	M	F	M	F
Set A	2	1	2	2
Set B	0	4	3	1
Set C	3	1	0	4

However, results in table 5.6 do seem to indicate some differences in the categorization behaviour produced by the different picture sets. Set A pictures are all bouncing motions and they all possess similar trajectories while pictures in the other two sets are more varied in this respect. The smaller number of total categorization schemes arising from set A may be due to the strong similarities among the pictures.

Table 5.6 Distribution of classificatory categories by picture set

	Descriptive			Sub-total	Analytic			Sub-total	Total
	D1	D2	D3		A1	A2	A3		
Set A	4	7	4	15 (63%)	2	2	5	9 (37%)	24
Set B	2	11	6	19 (58%)	2	8	4	14 (42%)	33
Set C	11	8	4	23 (70%)	3	3	4	10 (30%)	33
Both sexes	17 19%	26 29%	14 16%	57 (63%)	7 8%	13 14%	13 14%	33 (36%)	90 100%

Another difference in categorization behaviour evident in table 5.6 is that set B is producing many more A2 schemes and set C is producing many more D1 schemes. Set B pictures are all 'flying' motions, that is, motion in the air while set C pictures are all 'falling' motions. Ogborn (1989) proposed a model for the origins of ideas of the causes of motion which identified 'effort' and 'support' as the two basic categories for thinking about motion. Falling and flying came out as two important stereotypical motions in his model. Falling is very special in that it is caused by a lack of something (in this case, support) and not by the presence of something (generally, effort). For flying to be possible, both effort for the flying motion as well as the source of support (possibly as a result of own effort) has to be accounted for. Such a model implies that there is a much greater inherent cognitive need to look for causes of motion for flying than for falling in the commonsense context. This may account for the much larger number of causal schemes arising from set B and the relatively few causal schemes from set C. Set C pictures tend to describe more dramatic situations many of which put the objects in motion into rather dangerous positions. This may explain the much larger number of D1 schemes arising from set C.

5.6 Summary

The categorization schemes produced by the students on motion event comic strip pictures reveal that students have two main kinds of concerns in the intuitive processing of motion events in everyday encounters: the need to give a broad and rough description as to what kind of physical motion script it is and the need to find out how the motion is effected/caused. In most of the cases, the explanations given by the students revealed that processing was not done by means of logical derivations according to rules relating the various parameters of motion, but rather involved a search for prototypical motion scripts like falling, bouncing, etc., that could be suitably applied to the motion at hand. This seems to support the idea that knowledge about motion events are stored in units that are larger and more structured than semantic nets. Such knowledge structures are likely to be at the level of complexity similar to Schank's (1975) scripts or Minsky's (1975) frames. They serve the purposes of dealing with the pragmatic concerns of trying to anticipate how the motion will develop and to search for strategies to handle the situation. The important distinction between animate and inanimate objects in commonsense considerations of motion can also be explained within such a context of looking for suitable strategies to handle the situation. The advantage of using such large and complex structures for memory storage is that the processing time can be much reduced since each script would hold default values for most of the motion parameters (slots).

Another interesting finding from the present study is that older students who have received more instruction in physics have more ways of describing and analyzing motions than their younger counterparts. However, there is no evidence that the two groups possess fundamentally different structures in their script representations, nor is there a difference in the pragmatic outlook for the two groups of students.

Though there is no fundamental difference in the performance of the subjects from the two sexes, girls seem to exhibit more humanistic concerns and to play more on their imaginations in their processing of motion events. They produce a much higher proportion of thematic classifications which focus on people's intentions, the social circumstances of the motion, etc. On the other hand, boys seem to be more inclined towards constructing analogical/ theoretical models to explain for the causal mechanism of the motions.

The A3 type schemes, that is, those schemes that try to propose causal mechanisms for the motions at hand, indicate attempts by the students to tackle the problem of causal agency beyond simply giving it a label, be it a concrete agent or an abstract one. There is an attempt to generalize across various specific agencies some general underlying mechanism whereby the causal effect takes place. The availability of such mechanisms in the knowledge structure about motion enables the cognitive processing about motion in everyday life to go beyond pattern matching and searching for an appropriate script to fit the event into. It represents some bootstrapping procedure where one can try to reason logically about the behaviour of a system even when no exact script for handling it is available. The level of explanation provided by this kind of schemes is thus deeper than a mere labelling of the causal agent and provides a link towards the scientific study of motion.

Two of the schemes produced by the older group of students do exhibit a creative application of the scientific principles learnt to the solution of everyday pragmatic concern of causal analysis and anticipation. The main conceptual jumps taken were that causal agency is only required to effect a change in motion, and that though animate object may be able to influence their own course of motion by moving parts of their own body, they still have to succumb to the same physical laws of motion.

Chapter 6 THE PROGRAMMING TASK

6.1 Introduction

The main concern of the present research is to build up an in-depth cognitive model of children's conceptions of motion, including how they reason about them in everyday circumstances. The review in chapter 4 points out that the method of elicitation affects the quality and nature of the knowledge elicited as the nature of the task determines the mode of access to the long term memory and the depth of processing of the information. The classification task, described in the foregoing chapter, was chosen because it is a cognitive task requiring spontaneous reactions, and appealing to the subjects' direct perception of motion events. This suffices to provide some insight into the pertinent features of children's long term memory structures and reasoning patterns about motion events. However, this research also intends to probe into features of deeper modes of processing about motion events: How do children reason about everyday motion events when they are required to operate at a formal, rational level? Commonsense knowledge is typically fuzzy, largely tacit and there is no general agreement as to the intra-personal consistency in the choice of rules in dealing with motions in different contexts. What if children's ways of reasoning about motion are explored as an explicit externalization from them instead of as implicit deductions from their responses to certain tasks? What if we look for more considered, rationalized responses instead of spontaneous ones? Would we see any difference in the reasoning so elicited?

To answer the above questions warrants the setting up of another task which can lead to a more explicit elicitation of ideas and presents a context for the researcher to probe deeper into the application of the elicited ideas to various situations. The task chosen was a programming task: subjects were invited to write PROLOG expert systems which would be capable of commonsense reasoning about motion. This method of knowledge elicitation was inspired by developments in cognitive science research. The basic idea here is that the subject plays the role of a teacher and the computer that of a learner. In the course of teaching the computer, the subject is obliged to externalize his/her ideas. Once these ideas and intuitions have been externalized, they become in some sense accessible to reflection. The resulting expert system encapsulating the externalized ideas thus 'freezes' the subject's ideas at a certain instant, allowing him/ her to use it to explore the logical consequences of the externalized knowledge and at the same time to support easy modification when this is deemed necessary.

6.2 Expert Systems

In cognitive science research, it is common to use production systems to build cognitive models of human reasoning in various domains (Newell & Simon, 1972). Production systems are rule-based systems that generally comprise the following elements:

* A set of rules (productions) of the form

  If condition 1 and
  condition 2 and
  ...
  ...
  then action 1
  action 2
  ...

  The '*if*' part determines the applicability of the rule and the '*then*' part describes the action to be performed if the rule is applied.

* One or more databases that contain whatever information is appropriate for the particular task. Some parts of the database may be permanent, while other parts of it may pertain only to the solution of the current program.

* A control strategy that specifies the order in which the rules will be compared to the database and a way of resolving the conflicts that arise when several rules match at once.

(For an introduction to production systems, see Rich, 1983 or Winston, 1984)

Apart from being used extensively to model human problem solving by information processing psychologists, production rules are also very popular as a form of knowledge representation in the creation of expert systems. The DENDRAL project (Feigenbaum et al, 1980) and the MYCIN project (Buchanan & Shortliffe, 1984) are two of the successful examples. An expert system is essentially a computer program capable of performing some tasks normally performed by a human expert (e.g. programs that can perform medical diagnosis or geological analysis). These tasks require a great deal of specialized, and very often empirical, knowledge that normally only experts have access to and most people do not possess. Such programs are useful because there is usually a shortage of qualified human experts. Users normally consult an expert system when they have certain specific problems in mind. They put queries to the expert system, asking for advice or analysis as the case may be. When necessary, the expert system may query the user on certain pieces of information about the specific situation of the current consultation if such are needed for determining the answer to the user's query. Some expert systems also provide an explanation facility on how the answer to a query is arrived at.

Technically, expert systems are usually built through a joint cooperation between a knowledge engineer and a domain expert. The knowledge engineer tries to elicit the knowledge that the domain expert uses when solving problems and to put this knowledge into the form of a program that can answer queries from the user and if necessary generate queries to the user about specific information relevant to the problem in hand. The expert system as an operational program thus comprises two kinds of knowledge: the domain knowledge characteristic of experts in the specific problem area and the capability to make inferences and deductions from the given knowledge base, an ability characteristic of general human intelligence.

Every one of us can reasonably be called an expert in the domain of commonsense operations in the world of everyday motions: The fact that we have been faced with myriad different types of motion for years and been able to cope successfully with them justifies such a claim. What features can we find in the knowledge structures of these motion experts? How do such knowledge structures differ from the general rules and principles of Newtonian mechanics?

The present investigation tries to elicit children's intuitive ideas about motion through asking them to write expert systems reflecting their own knowledge about motion, and observing their interactions with their own knowledge so expressed.

It was anticipated that the secondary school students participating in this research would have little or no programming experience (it would be extremely undesirable to use programming experience as a criterion for selection of subjects for the purpose of the present research). As the subjects are expected to write expert systems reflecting their own expertise in the handling of everyday motions, they have to play the roles of both the domain expert and the knowledge engineer. To play the first role should present no problem, but to play the second role requires the skills of both knowledge elicitation and programming. In order that this task will not present insurmountable difficulties to the subjects, the choice of a suitable programming environment is of critical importance.

The software used for this task was the micro-PROLOG programming language (McCabe et al., 1985) together with an expert system shell, APES (short for Augmented Prolog for Expert Systems, Hammond & Sergot, 1984). PROLOG is different from the more familiar programming languages like BASIC or LOGO in that it is a declarative language whereas the latter are procedural ones. Programs written in procedural languages specify actions to be performed. They are geared to the description of what is to be done in order to achieve the desired result. Programs written in a declarative language on the other hand describes a set of relations or functions to be computed. A PROLOG program can be seen as a set of propositional statements containing facts and rules which then constitute a 'world' from which consequences can be generated when the program is executed. A PROLOG program is executed when a query is posed to the system. For example, a small PROLOG program describing the state of certain objects and the concept of support may read as follows:

```
X rest-on Y if
      X on Y

X rest-on Y if
      Z on Y and
      X rest-on Z

X supports Y if
      Y rest-on X and
      X strong-enough-for Y

book on table

table on ground

Mary on chair

chair on ground

ground strong-enough-for table

table strong-enough-for book

table strong-enough-for chair

chair strong-enough-for Mary
```

It can be seen that the first three statements in the above program are rules (with the first two defining a recursive relation 'rest-on') while the rest are facts. To find out from the program whether the book is supported, a query can be posed in the following format:

```
is (table supports book)
```

and the program would respond *YES.* Alternatively, the program may also be asked to give a list of objects that are supported by the ground using the following query:

```
which (x: ground supports x)
```

and the program would respond with the following answer:

book
table
Mary
chair
No (more) answers

A distinguishing feature of the PROLOG language is that answers to queries like those described above can be computed without explicit instructions inside the program about how the solution has to be computed. The reason is that PROLOG has built-in goal-resolution mechanisms based on first order predicate calculus and may thus be visualized as a logical inference engine (Clark & McCabe, 1984). Consequently, it is capable of performing logical inferences based only on input "knowledge" in the form of propositional statements. Furthermore, the processes of program development and execution can both be done as a form of dialogue: adding statements and querying the program can be done at any time.

An inspection of the above program shows that the rules represent a relatively general way of looking at support whereas the facts statements are just specific data concerning a particular situation. Presumably the same program can be used to answer queries about support for other situations provided the relevant facts are entered into the program. In many expert systems, the program has the ability to ask the user about pieces of empirical information required during the process of computing the answer. Another attractive feature of many commercially available expert systems is that they can explain how they have arrived at the answers put forth by the program.

APES is a program written in micro-PROLOG which extends the programming environment so that both the query-the-user facility and the explanation facility are available. A sample dialogue in the APES environment about the same problem of support can go on as follows:

User: X rest-on Y if X on Y.
X rest-on Y if Z on Y and X rest-on Z.
X supports Y if Y rest-on X and X strong-enough-for Y.

User: confirm (ground supports book)

Computer: Is it true that book on ground?

User: no.

Computer: Is there any Y such that Y on ground and book on Y? If so, name the Y.

User: table.

Computer: Is ground strong-enough-for book?

User: yes.

Computer: I can confirm that ground supports book.

User: why?

Computer: I can confirm that ground supports book because book rest-on ground and ground strong-enough-for book.

User: why (book rest-on ground)?

Computer: I can confirm book rest-on ground because book on table and table on ground.

User: why (table on ground)?

Computer :You told me (table on ground).

From the above dialogue, we can see that the main advantage of using APES is in the execution of the program. When a Prolog program is used in APES environment, the computer is capable of asking the user for missing information or facts required for the inferences demanded. For example, a medical expert system should be able to ask the patient for specific information in order to arrive at a diagnosis. Furthermore, the computer is capable of explaining how it arrives at the inferences made using the facts and rules put in.

APES belongs to one type of programming environment often referred to technically as an expert system shell. It is called a shell because the environment itself does not hold any explicit knowledge about the world whether in the form of facts or rules. However, upon entry of declarative information about a particular domain in the form of logical propositions, it can perform the functions of an expert system.

There are a number of advantages in using expert system development as a knowledge elicitation and cognitive modeling task. Firstly, it forces the subjects to make explicit their ideas as to be able to put them into a formal system. Secondly, the resulting knowledge base is explorable by the subjects and can thus provide a feedback to them on the cognitive consequences of their own elicited models. Thirdly, this provides an opportunity for the subjects to modify and test their ideas in a non-threatening context, and at the same time leaving a traceable record of the changes that have taken place. Moreover, the task requires the subjects to work for a number of one and a half hour sessions over an extended period of several weeks, thus providing a good context for the researcher to probe into the reasoning structures employed under different situations.

6.3 The Pilot Study

Building an expert system is a novel task for a secondary school student and this may thus be able to sustain their interest for a considerable period of time. However, it is a novel technique that has not been employed for similar purposes, so the feasibility of such a technique poses serious problems. A pilot study was carried out to answer the following questions: (i) Can a secondary school student with no programming experience learn to write such an expert system? (ii) What kinds of ideas about motion can students write when given such a task context?. Would this task prove too restrictive to the students for describing everyday conceptions of motion? (iii) What are the difficulties that one would encounter in this process? If such a task is practicable at all, what procedures should be adopted for it to be useful as a means of scientific enquiry into the subjects' conceptions and reasoning?

All of the students who participated in the pilot study of the classification task were invited to take part in the programming task as well. As these were all form six students (from the same class), they represent the most intellectually capable section of the population in secondary schools. If the task proved too difficult for them, there would be no hope of success with subjects coming from lower down in the schools. As it was near examination time, two out of this group of four students declined this offer though they did think the task interesting. Of the two students who participated in this pilot study, one was the best student in the class and so the examination posed no threat for him. The other student was amongst the weakest in the class, but as he had already made up his mind at that time to quit school after the summer vacation, the examination again posed no hindrance.

As the task was intended to elicit the ideas held by individual students, each of the two students worked separately. The students were told that the purpose of the exercise was to assess the friendliness and power of the expert system shell: whether it can be used by people with no training in computing to teach the computer some commonsense everyday knowledge. They are the ones to give the computer system the test by trying to teach it something about motion in the everyday world of human beings. After the exercise they would be requested to give their assessment of the system. Thus, the task as described to them focussed on the programming environment. This was to ensure that the participating students would not feel that they were being challenged or judged on their Physics knowledge. However, as the task was to build an expert system capable of reasoning like the student himself/herself about motion, it provided a context for serious introspection on the part of the student.

To start with, they were given an introductory session on programming in PROLOG. They were told that PROLOG programs consist of rules and facts. Only the infix format of PROLOG statements (available in APES) was introduced as this bears closer resemblance to natural language. To help them understand more about the working of the system as an expert system shell and the possible forms of interaction, a simple expert system on biological classification was used as an example. They started writing their own program in the second session. No guidelines were given to the students as to what they should write and they all found it difficult to pin down any ideas to write about at the start. To help them break the ice, it was suggested to them that they might like to start by writing rules about the ideas they had used for classifying the motion comic pictures.

The students worked on the programming task for five sessions, each lasting for about one and a half hours. Several observations were made in this process:

i. The students found it easier to first jot down their ideas on paper before attempting to type them in as PROLOG rules. This was especially true for the weaker student.

ii. As Taylor & du Boulay (1985) and Dean (1986) pointed out, the close resemblance of the PROLOG syntax to natural language is deceptive. Students found it easy to transform their ideas into PROLOG syntax. However, they tended to forget that what the computer does is actually pattern matching, a purely mechanical task, and that the computer does not *understand* the contents of the rules the way humans do. They did not appreciate the need to structure the rules to use exactly the same relation name when the meaning is the same. For example, they wrote 'motion needs a-force' on one occasion but 'running needs-a force' on another. To leave the rules exactly as they wrote them is not desirable because they could not see the full consequence of their ideas when queries were executed, since not all relevant rules could be invoked. However, intervention from the researcher was not always desirable. In many cases, it was difficult to tell whether the similar but different relations actually referred to slightly different ideas and intervention at this stage could unduly affect and shape the students' ideas.

The student with the higher academic performance picked up the importance of structuring his rules very soon, and also began formulating hierarchically related rules. The weaker student never realized the crucial difference between PROLOG statements as a formalism and ordinary natural language statements.

iii. The students seemed to be more interested in expressing their ideas than in testing them out. Since the kinds of queries they intended to ask were on the whole simple ones, the restrictions of the PROLOG syntax did not pose a serious hindrance to them.

iv. Both students found no serious difficulty in starting to work in PROLOG after the initial session. Sometimes errors arose which they could not debug by themselves. However, as the researcher was present all the time during the sessions, such difficulties could be sorted out without too much problem and no undue proportion of time was spent on the technicalities of the programming language.

v. PROLOG forces the students to express their ideas in the form of general rules. It may thus be argued that this distorts commonsense ideas about motion because they are by nature fuzzy, informal, very often situated, context dependent, and may even be tacit. However, the PROLOG formalism forces them to reflect more deeply about what are the key issues and parameters concerned in different situations. Furthermore, the resulting program gives a relatively 'objective' record of the students' thoughts, an objectivity that is normally unavailable by probing through other means. The objectivity here refers to the fact that the elicited ideas are not just something written up by the researcher, but something written by the subject, open for inspection, reflection, exploration and which is capable of providing feedback on the operational consequences of the knowledge base, so that s/he can determine whether such a representation is satisfactory or not.

vi. PROLOG syntax is sparse and inexpressive. Sometimes, it is difficult to appreciate the full meanings the students wanted to express just by reading the program. However, the programming activity provided a good context for communication between the students and the researcher and it allowed the researcher to probe deeper into the thoughts of the students under the pretext of trying to find out whether the rules do actually reflect their own ideas.

vii. The contents of the programs written by the students reflected everyday, commonsense concerns which are not the subjects of discussion in school physics lessons, indicating that the programs do offer an insight into the students' intuitive beliefs about motion. At the same time, both students' programs also incorporated terms and ideas learnt from school physics, indicating the possibility that knowledge from these two domains are interacting, and perhaps becoming integrated.

viii. There were qualitative differences between the final programs produced by the two students. At the structural level, as mentioned earlier, one was much more hierarchical, forming an integrated whole, while the other program displayed much less inter-relationship between the rules and the whole program stood as an assorted set of statements. At the content level, the two programs reflected different levels of adherence to common misconceptions about motion, with one of the programs incorporating some Newtonian ideas. However, it is very difficult to directly compare the two programs as they wrote about different motions with different emphasis.

The above observations from the pilot study results indicate that the task of asking students to write expert systems in PROLOG is technically feasible. The programs do provide an alternative means of probing into students' conceptions of motion, and that this technique offers a special kind of 'reflexivity' as described above which may be valuable to our understanding of commonsense conceptions of motion. The pilot study also provided valuable experience in the kinds of problems and difficulties encountered in employing such a technique, and demonstrated the need to adopt more formalized procedures in order to analyze and compare the elicited ideas from different subjects.

6.4 The Programming Task in the Main Study

6.4.1 Objectives of the Investigation

As the report on the pilot study in the previous section indicates, expert system development in PROLOG is a potentially valuable experimental technique for probing into students' conceptions of motion. The technique was employed in the main study with some further specific objectives:

i. To find out how features of the schemas for commonsense interpretation of motion as elicited in the classification task relate to the schemas elicited when a student is required to rationalize about the logical relations between different aspects of motion.

ii. To find out how learnt physics concepts are incorporated into such schemas.

iii. By investigating into the above two problems to seek insight into the difference in knowledge structures between high achievers and low achievers in physics and between two age groups. Hopefully, this may also yield interesting information about the possible learning processes involved in the learning of mechanics. Differences in performance between the two age groups may be due to intellectual maturation with age or the amount of physics instruction received. No attempt is made in this investigation to explore these two factors as independent variables.

6.4.2 Subjects

The two variables that are of interest in this present investigation are the amount of physics instruction (especially mechanics) the student has received and the general ability of the student. Twelve students participated in the programming task for the main study: six 18-yr-olds and six 15-yr-olds. The former were sixth form students studying science, and each of them had gained a grade C or above in O-Level Physics. The latter were third form students who had not received any instructions in mechanics yet. The six students in each group (three girls and three boys) were representative of the high, average or low abilities in their respective classes.

6.4.3 Administrative Details

The following procedure was formalized and adhered to for conducting the programming task in the main study:

i. Each participant was asked to participate in the classification task described in the last chapter. This generated some parameters/features to act as starting points for the building up of the expert systems.

ii. Each group of students attended an initial session which introduced PROLOG programming and interaction with APES. This was the only group activity they had and it was done in this format to save time. Each student was given a handout on how to start working with the system (App. 5).

iii. At the first programming session, each student was given a summary of the classification types used by that group of students during the classification interviews (App. 6). This provided some starting ideas for the students to develop their programs on, making it unnecessary to suggest ideas otherwise and thus eliminating the possibility of the researcher unknowingly biasing the students towards certain ideas.

iv. The students were encouraged to put notes on paper about what they wanted to write before typing them into the computer so that they did not have to worry about the syntax at the first instance. For particularly complicated statements, the researcher sometimes helped the students to transform them into PROLOG syntax on paper first, and asked them if that actually represented their thoughts, before asking them to type the rules in.

v. In order that the results from the programming task could be easily compared, it would be desirable if they were all concerned with roughly the same aspects of motion. However, it was considered not suitable to ask the students to write on any particular theme right from the start so as to allow them greater freedom in the exploration of the system and to arouse interest. From about the third session, when they had finished with their first set of ideas and began to think of new topics to write about, it was gently suggested to them that they might like to write some rules on what affects motion, causes of motion or control of motion. Which of these three topics they did write about was entirely up to them. Each student worked for six to eight sessions on the programming task, each session lasting for about one and a half hours.

vi. During the programming sessions, whenever there was a discourse between the researcher and the student clarifying the contents and meaning of the rules and how the students actually thought about motion, the conversation was audiotaped.

vii. At intervals, when the students had finished typing in a set of rules, they were encouraged to test the programs using different concrete situations. This posed no serious problem because, normally, the students would already have in mind some situations when they first made up the rules: Syntactically the rules were general rules, but often they were actually anchored to very specific situations when they were first formulated. Thus they normally did the tests using the anchoring situations and these usually produced acceptable results. It was considered desirable to confront students with situations which produce conflicting answers to queries so as to provoke students to reconsider seriously what they had already written. When the students became reasonably pleased with the answers from the test questions, the researcher would suggest different, and possibly, more stringent test situations for the programs.

viii. Most of the activities during the programming sessions were initiated by the students. However, it was sometimes necessary for the researcher to act as initiator, e.g. when she spotted a syntax error or when she wanted to suggest a test situation. Notes were taken to record the initiation of all the activities.

ix. Screen dumps of all computer interactions during the programming sessions were recorded. Brief notes were also taken during all the sessions to record the development of events and the main ideas expressed. These were then re-organized to produce the diary notes on each of the sessions shortly after they had taken place (see next section).

x. There was a slight technical difference in the organization for the younger group of students. Instead of using APES as the expert system shell, another shell written in PROLOG Professional (McCabe et al., 1985), MITSI (Briggs, 1984), was used. MITSI is a shell very similar, slightly less powerful but much more friendly and easy to use when compared with APES. It also supports the infix format, query-the-user facility and offers explanations for answers. This change in software did not affect the comparability of results between the two groups of students and the change was necessary in order to make the task less technically threatening for the younger group of students.

xi. In the original planning, after the students from one age group had completed the programming task for some time, each of them would be invited to test all the final programs from their own group with standard test objects and give a written assessment of the performance of each of these programs, including their own. The purpose of this exercise was to find out how public, that is, how generally acceptable the reasoning elicited from these programs are. However, because of difficulties in time arrangements, only the 18-yr-olds completed this exercise, and only on a very limited scale.

6.5 Data Collection and Preparation

For each programming session, there were four sources of data collection:

i. Written notes by students about ideas they would like to put into the program.
ii. Screen dump of all computer interactions.
iii. Notes taken by the researcher during the session, recording the development of activities, the initiators of the individual activities and the main development of ideas.
iv. Transcripts of parts of the conversation between the student and researcher for clarifying the ideas they were writing on or sometimes about how they reasoned about everyday motions.

In order to facilitate systematic analysis of data, these four forms of data needed to be collated. Upon inspection, it was found that each programming session could be subdivided into episodes, there being essentially four kinds of episodes: writing notes on ideas, addition of rules into the knowledge base, testing of the program or modification of rules. It was decided that diary notes for each of the sessions containing a description of what happened, the development and exposition of ideas, collated from the researcher's written notes and transcripts of conversations, alongside a record of the program development would be a clear and useful format to adopt for the data preparation. Such diary notes, an example of which is shown in table 6.1, were compiled shortly after each programming session from the above four sources of data, and these formed the basis for the further analysis of data.

Table 6.1 Sample diary notes for one of the programming sessions.

DIARY NOTES FOR PROGRAMMING SESSION 2
BRENDA
September 29th, 1986.

EVENT	PROGRAM DEVELOPMENT
1. She asked me about the key words confirm and find - how they should be used in the program. I then tried to clarify the three different hats that she is wearing alternately, stressing their relationship and when she should wear which - the first hat is that of the expert, when she is trying to extract her own commonsense knowledge about motion; the second hat is that of the knowledge engineer/ computer programmer, when she is trying to put her own thoughts into a Prolog program; and the last hat is that of a user who wants to solve some problems using the finished expert system,when she is testing the program with "find" or "confirm".	
2. She tested the rules she put in last using the apple as the testing object. She seemed to find the response acceptable.	TESTING find (_force: apple experience _force) ->air-current (reason:moving upwards) ->gravity (reason:travelling parabolically) -> no more

EVENT	PROGRAM DEVELOPMENT
3. I then suggested using the tape-recorder as the object. The response - that the tape-recorder does not experience gravity is obviously wrong to me, but she was not bothered.	TESTING confirm (tape-recorder experience gravity) -> No
4. She looked at the screen for a moment and then went on to test the program with a kite.	TESTING find(_force: kite experience _force) -> air-resistance -> centripetal-force -> no more
5. After she finished, I asked if she is happy with the result. She only said that it is too small to really know much. I asked if she is satisfied with the respon-ses so far from the computer. She said she think they are o.k. I then referred back to the response to the case of the tape-recorder and asked if she thought that was reasonable. At that point she said it was wrong because tape-recorder should also experience gravity. I asked her again if she is really sure that the tape recorder does experience gravity, to which she responded to the affirmative. I then asked her to make amendments to take account of this and she added a further rule in.	ADDITION _object experience gravity if _object has weight and _object is still
6. Because "is", a reserve word for PROLOG, was used in the program, a syntax error occurred and we made the appropriate amendment and tested it with a book. That was o.k.	MODIFICATION the relation name in the last condition was changed and the amended rule reads: _object experience gravity if _object has weight and _object stay still TESTING confirm (book experience gravity) -> Yes (reason: book has weight and stay still)
7. I then suggested we should give more tests on the program and suggested to use the case of the kite that she used earlier. She proceeded to do so, without actually anticipating that the result would be that it does not experience gravity. After she has read the explanation, she immediately said that it was wrong and that she should cross out the condition saying that the object should be still from the	TESTING confirm (kite experience gravity) -> No (reason: not travelling parabolically and not stay still)

EVENT	PROGRAM DEVELOPMENT
rule. I asked if there would be any difference whether the object is still or not and she said that there should not be as far as gravity is concerned. I then asked her why she added that condition in at the first instance.	
B: Actually all objects are affected by gravity. But I wanted to write a rule specially for still objects. I wanted to specify that it stays still. But after I have put that in, then it gives me a result that I don't agree with. But if Inow add in a condition for kites, that it can move and still have gravity. That seems to be rather contradictory. What should I do?	
I: When you put in the condition about stay still, what was in you mind? What did you want to express?	
B: I was all along talking about things that are moving. So I wanted to put in a rule that can talk about objects that stay still.	
I: But actually, would there be any difference between moving and non-moving objects?	
B: No difference. For gravity. As long as it has weight, then it would be affected by gravity.	
I: But you must be thinking of something when you added that condition in. So there should not be no difference at all. Or else you wouldn't have added that in.	
B: Yes,..	
I: So, what were you thinking then?	
B: .....It was talking about the tape recorder in particular. I wanted to specify that it is staying still.	
8. She then decided to eliminate that condition. Then she tested the changed program with the kite again.	MODIFICATION eliminated last condition on rule to become: A object experience gravity if _object has weight

EVENT	PROGRAM DEVELOPMENT
9. She then went on herself to test the program with the case of a bird. I couldn't quite understand why the bird was moving upwards, in a circle as well as parabolically. *B: It sometimes moves in a circular fashion, sometimes it would move down. There are different cases. For somebody coming along to use the system, he would look at the case of the bird very generally, so it would some-times move like this, sometimes move like that, and so on.* At the end, I asked her to comment on the behavior of the program. She again said that it is a bit too small, but the responses are quite reasonable. The task for the next session was agreed to be working on control of motion or cause of motion in the form of _object motion-controlled-by ______ if .... or _object motion-caused-by ________ if	TESTING find (_force: bird experience _force) -> air resistance (reason: travelling upward) -> centripetal force (reason: moving in a circle) -> gravity (reason: travelling parabolically) -> air-resistance(reason:same as above) -> gravity (reason: has weight)

CHAPTER 7 ANALYSIS OF THE PROGRAMMING TASK RESULTS

7.1 A General Description of Performance

As was found during the pilot study, the students in the main study did not find it difficult to write PROLOG rules about motion. Though the programming environment forced them to write their ideas down in a formal syntax, their interpretation of the task at hand was much less formal. This was especially so with the younger group of students. They tended to see this as an opportunity to express their own ideas and seemed to have stronger confidence in the rules they wrote down. As was explained in the previous chapter, there are basically four kinds of episodes taking place in a programming session: writing notes on ideas, addition of rules into the knowledge base, testing of the program and modification of rules. Table 7.1 shows the number of instances of each of the three kinds of programming episodes: addition, testing and modification of rules, for each of the two groups of students. The episodes of notes writing were left out because it was essentially a pre-writing exercise, trying to formulate ideas.

Table 7.1 Statistical counts of the different activities during program development.

(a) Distribution by student

Episode	Students											
	F. 6						F. 3					
	wong	evon	lin	bren	sand	fong	cheng	schan	ychan	cheung	leung	wat
Addition	27	18	3	37	12	41	50	34	48	51	36	35
Testing	16	39	39	31	37	26	3	29	37	28	16	5
Modification	11	19	13	12	9	22	2	10	12	14	7	6

(b) Distribution by student group

Episode	Student group	
	F.6	F.3
Addition	148 (35%)	254 (59%)
Testing	188 (45%)	128 (30%)
Modification	86 (20%)	51 (12%)
Total	422 (100%)	433 (100%)

Results from the table show that the F.6 students tended to be more reflective, devoting about half the total number of episodes to the testing of rules, whereas the F.3 students devoted 60% of the episodes to the adding in of new rules. The greater reflectiveness of the F.6 students also resulted in many more modifications made to the rules. This also resulted in a sense in a higher "productivity" of the F.3 students, producing 67% more rules than the F.6 group, though the total number of episodes for both groups were roughly the same.

Besides a general difference in reflectiveness, the older group also displayed greater systematicity in their program development. Students in the F.6 group all started working by adding in rules describing various features and parameters about motion. Once they had taken up the suggestion to develop rules on the cause or control of motion, they would work on the problem with focussed attention, trying to work out various possibilities under the same theme before switching to another. None of the F.3 group displayed the same order of systematicity. Even after they had agreed to write rules on cause or control, their pattern of working was not much different: The themes for the rules kept changing all the time, switching back and forth, suggesting that they were trying to put down whatever they thought about at that time. They did not display the same level of mental discipline as was evident in the older group.

7.2 Organization of the Analysis

As mentioned in the previous chapter, the PROLOG programming environment offers a medium for the students to externalize, reflect on and explore their ideas. The analysis of the programming task data is divided into two main parts. The first part of the analysis focuses on the programming task as an externalization process. As such, it can also be viewed as a knowledge elicitation process, offering a means to look into the students' conceptions of motion and to compare the results with those obtained in other researches on intuitive physics, employing other methods of elicitation. The second part of the analysis focuses on the testing and rule modification processes. Because the written rules are 'runnable', the externalized knowledge base was open to exploration, reflection and modification. The analysis examines the students' behaviour during the testing and modification episodes, how the students faced conflicts presented in the program testing, the cognitive strategies employed in the resolution of such conflicts and how such a process may affect the genesis of the students' ideas.

The rest of this chapter will be devoted to the first part of the analysis as described above: an analysis based on the content of the rules developed by the students. This will be done first by looking at the content and structure of the rules: what the students were writing about. The rules will also be analyzed to find out the sort of physics of motion they reflect and thus to explore one of the main objectives of this research: what the programming task data can tell us about intuitive conceptions of motion, and how the amount of formal physics instruction affects such understanding. Next, a deeper level of analysis will be attempted: the rules will be classified into groups according to the kind of explanation or prediction that they offer, so as to be able to make inferences about the level of the students' cognitive functioning in their performance of the programming task.

The second part of the analysis will be discussed in the next chapter.

7.3 A First Level Analysis: Contents and Structure of the Rules

At the content level, the rules written by the students fell into four clearly identifiable areas:
i) rules relating different attributes and aspects of motion,
ii) rules relating to causes of motion,
iii) rules about control of motion and
iv) rules providing definitions for terms.

The first group constitutes the largest single group of rules for both the F.3 and the F.6 students, accounting for nearly half of the total number of rules. The figures in table 7.2 indicate that the relative distribution of the different broad content areas of the rules are very similar for the two groups of students. However, a careful scrutiny of the finer distribution amongst the different rule structures in each broad rule group shows that the important differences between the two age groups are such that a more careful analysis is warranted.

Table 7.2 Content distribution of rules.

Content type	**Student Group**			
	F.6		F.3	
About motion	59	(40%)	124	(50%)
About cause	33	(22%)	60	(24%)
About control	28	(19%)	11	(4%)
Definitions	15	(10%)	19	(8%)
Miscellaneous	13	(9%)	33	(13%)
Total	148	(100%)	247	(100%)

Most of the PROLOG statements written were in the form of production rules with the following syntax:

```
goal statement (conclusion)if
     condition 1 (premise) and
     condition 2 and
          ....
          ....
```

This raises two problems. The formalism is very rigid, and in addition the whole set of rules should form a logical system. These students had had very little experience of working with logical propositions, and none had had the chance to do so in the context of building up a formal system about a specific domain. Thus, for example, what relations should go as premises and which should go as conclusions were not clear to them. Out of the total of about 400 rules written by the students, only about 10 distinctly different kinds of relations can be identified. Most of these relations can be found taking up positions of premises in some rules and conclusions in some others. Sometimes we may even be able to find both

p(x) if q(x) and

q(x) if p(x)

as two separate rules in the same program, with p(x) and q(x) definitely not seen as equivalent relations from the associated discussions with the students concerned. Very often, the logical implication was only treated as a means to indicate a high reliability of co-occurrence, and not a causal sequence.

Table 7.3 lists the distribution of rules for each student, with each broad content group subdivided into finer subgroups taking the structures of the rules into account: each subgroup having one particular composition of relations in the premises and conclusions ends of the rules. It can be seen that the older group of students exercised much greater caution in their choice of formats and that they showed a better understanding of what a logical proposition is. Of the 25 finer subgroups of rules, the F.6 students' rules were distributed amongst 13 of them, whereas the F.3 students' rules were distributed over 20 subgroups.

7.3.1 Rules about aspects and attributes of motion

The largest content group of rules is the set relating to different aspects of motion, titled as 'about motion' in tables 7.2 and 7.3. This group can be subdivided into 8 smaller subgroups of rules. The two most popular subgroups are of the format

```
motion type if
     observable physical conditions
```

and

```
motion attribute (abstract) if
     observable physical conditions.
```

Table 7.3 Fine Classification of Rule Structure and Content

Rule Content		Student Group												Rule type				
		F.6						F.3										
		wong	evon	lin	bren	sand	fong	cheng	schan	ychan	ahlun	leung	wat	EMP	SPC	HYP	CLS	DEF
About the motion																		
m. type	<- obs phy cond	11		1		2	4	6		6	4	6	3	39	4			
m. type	<- abst att							1					1	1	1			
m.att(conc)	<- motion type							4			4	2		6	4			
m.att(conc)	<- obs phy cond	2								3	4			4		5		
m.att(conc)	<- nature of obj											4		4				
m.att(conc)	<- abst att							1	5	3	3	4	2	1	2	15		
m.att(abst)	<- obs phy cond		7	3	9		16	18	10	3	17		10		76	17		
exp traj	<- cond.		2			2								2		2		
About cause																		
c.agt(conc)	<- obs phy cond				15	3			6	3			2	11	17	1		
c.agt(conc)	<- nature of obj							1				9			9	1		
c.agt(conc)	<- force req		3	3												6		
c.agt(abst)	<- obs phy cond	5	2			1							2	1	5	4		
c.agt(abst)	<- force req			1											1			
m.att(conc)	<- c.agt									7					7			
m.att(abst)	<- c.agt									3						3		
abst class./des. of agents								4								4		
enumeration of agency								5		5		6		5	6		5	
enumeration of forces								4					3				7	
About control																		
ctrl(conc)	<- obs phy cond	6		2	7	4	1							4	13	3		
ctrl(conc)	<- force req		2						2							4		
ctrl(force)	<- obs phy cond			1	5										6			
ctrl(state	<- nature of obj								1	4		2	2		4	5		
Definitional		2	1		1		11	1	5	3	6	1	3					34
Miscellaneous																		
concrete relations		1		2				2		1	4	2	5	17				
abstract relations			1				9	2	4	2	4	2	5		8	21		
Total		27	18	13	37	12	41	49	33	43	46	38	38	95	163	91	12	34

The former subgroup tries to deduce the kind of trajectory type from the observed physical conditions of the motion. Some of these express concerns very similar to those found in kinematics in the study of classical mechanics, for example:

> *motion type falling if*
> *motion due to gravity and*
> *motion moved downwards and*
> *motion will reach ground and*
> *motion speed increasing.*

More often, however, the first subgroup of rules represents empirical correlations between motion type and physical observables like whether something is supported or the type of object that is moving. The popularity of this kind of rules may be taken as another indication that knowledge about motion events is stored in the form of encapsulated complex schemas, or scripts, as described in chapter 5. Each rule in this category tries to express the encapsulated knowledge in the form of reliable coexistence of parameters, that is, default values for the slots, for a prototypical motion script. Two examples of such rules are:

> *thing will fall if*
> *not (thing supported by other thing).*

> *thing move upward if*
> *thing is living thing and*
> *thing has wings.*

The latter subgroup tries to make deductions about some abstract attributes of the moving object, for example about force or energy, from the observable physical conditions. This coincides with the concerns of dynamics. An example of this kind of rule is:

> *thing move upward and then downward if*
> *thing has limited energy in upward direction.*

These two most popular subgroups had rules from both F.3 and F.6. Of the rest, one subgroup had rules from F.6 only, and all rules in the other subgroups from F.3 only with the exception of one rule from the F.6 group in one of them..

The group to which only F.6 contributed is the subgroup

> *expected trajectory if*
> *some conditions.*

This subgroup tries to deduce whether the subsequent motion/trajectory of the object can be anticipated or not, and this is the only subgroup which is absent from the F.3 students' work. As noted in the analysis of the classification task data in an earlier chapter, to be able to anticipate the subsequent development of a motion is one of the most important pragmatic concerns in the everyday processing of motion related information. Some of the older students thus exhibited the ability to externalize this pragmatic concern, a concern which is commonly held implicitly. In the formal study of kinematics, students would be required to work out the variations in displacement and speed with time. This formal physics training might also have prompted the older students to put in rules of this kind.

As mentioned above, the other subgroups are essentially only found in the work of the younger students. A careful examination of these five subgroups suggests that this difference in their production also relates to the two groups' different background in formal physics training. These five subgroups will be discussed in turn below.

i) *motion attribute (concrete) if motion type*

It is understandable that if the type of motion is known, many features of the motion can be predicted. However, to the older group, such features or attributes reside already in the definition of the type of motion and such a rule does not portrait anything interesting.

ii) *motion attribute (concrete)if observable physical conditions*

Again for the older group, the primary problem was to deduce from the observable physical conditions the type of motion involved. Once that is determined, the motion attributes would be known.

iii) *motion attribute (concrete) if nature of object (animistic)*

The essential message in such rules were just expressions of animism, for example:

things cannot do more than one action at a time if
thing is a dead object.

iv) *motion type if motion attributes (abstract)*

The rules here express the belief that different types of forces would be required for different types of motion. For example:

thing path shape parabolic if
thing has bounce force.

The older group of students are more used to justifying assumptions about forces from observable physical conditions.

v) *motion attribute (concrete) if abstract motion attributes*

These rules essentially express the belief that larger forces (or sometimes energy) lead to higher speed, greater movements, while friction would slow things down. For example:

thing distance moved big if
thing force great.

Such statements are explicitly in opposition to what the older group of students have learnt in mechanics. However, their not putting them down in the form of explicit rules does not mean that they do not hold such beliefs implicitly.

7.3.2 Rules about causes of motion

The 10 subgroups of rules concerning causes of motion (refer to table 7.3) broadly fall into three types according to their function. Five of these subgroups try to predict the causes of motion for some given conditions. The next two subgroups try to predict the attributes of a motion when the causal agent is known. The last three subgroups try to enumerate and describe possible causes of motion.

An examination of the conceptions of causation reflected in the rules brings out several striking features. Firstly, the younger students saw two basically different kinds of motion: autonomous and non-autonomous motions. The former were seen to be performed only by animate objects and the rules in the subgroup

control agent i <- nature of object

reflects such ideas. An example of such a rule is

thing caused to move by itself if
thing is a living object.

The older students' work on causation did not reflect such a distinction and they did not have rules in this subgroup.

Another prominent feature of these rules is that all the rules on causation, with the exception of the ones about autonomous motions described above, conform to one single general form, which is expressed diagrammatically in fig. 7.1. Such a general form can also be labelled a "gestalt" in the sense used by Lakoff and Johnson (1980) and the general form depicted here is similar to Andersson's (1986) 'gestalt of causation'.

In the rules for the prediction of causal agency, the F.3 students did not make explicit reference to any necessary mediation for causation, (that is, there were no rules identifying forces as causal agents or necessary conditions for motion to occur,) and their rules expressed a pragmatic, a theoretical correlation between likely causal agents and observable

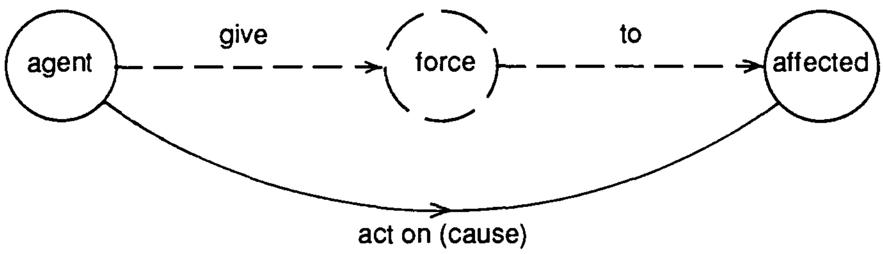


Fig. 7.1. Diagram showing the general form (gestalt) of causation expressed by the students' rules.

features of the motion. These rules again seem to be anchored on specific stereotypes of motion and the premises in the rules specify prominent features of that type of motion. An example of such a rule is:

motion caused by wind if
motion occurs in sky.

Some of the older students gave more elaborate descriptions of the causal mechanism in their rules. Many of their rules refer to forces being required as a necessary mediation for agency, and some identified specific forces as the causal agents concerned. Thus they produced rules of the following forms,

concrete causal agent i <- *force required for the motion*

and

abstract causal agent(force) <- *observable physical conditions.*

Examples of these two types of rules are:

object-A motion caused by object-B if
object-A initially in contact with object-B and
object-A force supplied by object-B.

object-A motion caused by reaction force if
object-A involved in collision and
object-A recoiled after collision and
object-A collided with object-B
not (object-A embedded in object-B).

These two forms of rules taken together express a model of causation as described in fig. 7.2.

The F.6 students also distinguished between two possible kinds of agency, physical agents and circumstantial factors such as the origins of the forces causing motion. The following two rules are examples of such a distinction.

object motion caused by reaction force if
object involved in collision and
object recoiled after collision and
not object embedded in the object it collided with.

object motion caused by gravity if
object moved in gravitational field and
object moved downwards and
not controlled by object in motion and
affected by mass of object in motion and
object speed increasing and
object in motion will reach ground.

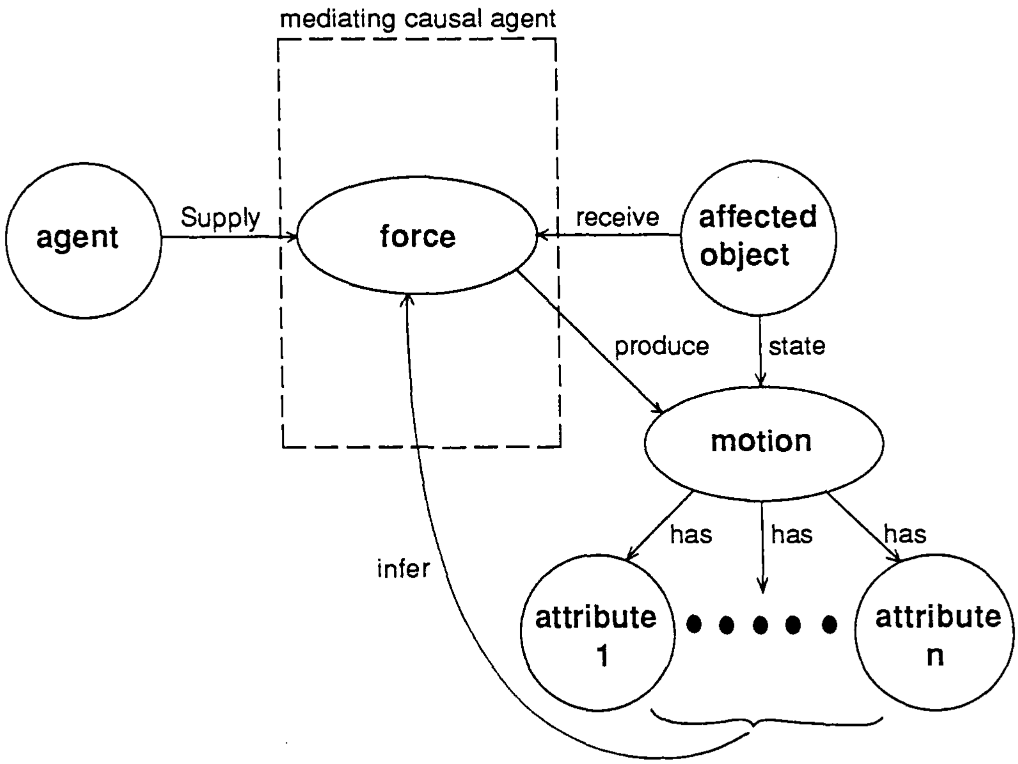


Fig. 7.2 The general form of causal mechanism as evidenced by the rules written by some F.6 students.

Though most of the F.3 students did not label forces directly as agents, they share the same fundamental general form or gestalt as depicted in fig. 7.1. The important role force plays in their causal "gestalt" of motion can be seen in the last four subgroups of rules on causation. These include rules like:

thing given force by other-thing if
thing movement helped by other-thing.

thing caused by move by something if
something is a kind of force.

thing caused to stop if
thing is a living object and
thing has no force.

The main difference between the two age groups in their conception of causation is that for the older group, force plays a much more central role in the causal process and this is reflected by their labelling force as the cause of motion, whereas the younger group separates force out clearly as the necessary ingredient for motion, and the status of agency is attributed to physical objects only when discussing particular motion situations.

7.3.3 Rules about control of motion

Ideas about control seem to be more varied across individuals than those about causation. Four main conceptions of control can be identified.

First, there are those who used control more or less as a synonym for cause, writing exactly identical rules for both cause and control:

object motion controlled by water current if
object moving up and down in water,

object motion caused by water current if
object moving up and down in water.

A second conception of control was to link control with those causal agents that are capable of exercising intentions and will-power, for example,

motion controlled by object in motion if
motion deliberately done by object in motion and
motion can change on will of object in motion.

A third conception was to attribute control to the dominating factor amongst a number of influences on the motion. An example of such a rule is

motion is controlled by self if
object moving is B and
external force acting on B and
external force less than maximum force exertable by B and
B can respond to stimulation and
B can be aware of occurrence of motion.

Curiously, all the above three conceptions of control of motion were absent from the younger students' rules on control. They did not look for agents of control. Instead, when they did write about control, they wrote about control as a possible *state* of a moving object, and they attributed control only to animate objects. Examples of such rules are

thing can control own motion if
thing is living object,

thing motion under control if
thing is dead(meaning inanimate) object and
thing used by animal.

Here again, we can see an explicit expression of concern for autonomous motion as a separate category by the younger students.

7.3.4 Rules offering definitions

Most of the students, at one point or another, felt the need to clarify the meanings of certain terms to ensure against ambiguities in the programs. Most of the 'definition' rules were of this type. However, some of these rules can also be seen simply as programming devices. For example, one of the rules in this group was

thing state at rest if
not thing is moving.

This rule was put in not for clarifying the rest state to the human user, but for the benefit of the computer in its goal resolution process. Such rules were normally put in when the students realized that the computer did not understand natural language and could not resolve certain queries because some statements carrying the same meaning had been entered in a different syntax.

7.3.5 Miscellaneous rules

Those rules not classifiable in the above four categories roughly fall into two groups. They are either highly particular descriptions of features of particular classes of objects or interactions not directly relating to motion, or they are more abstract relations about force and energy. These are respectively labelled as concrete relations and abstract relations in table 7.3. Examples of these are:

thing may be damaged if
thing crash with another thing.

thing force increase if
thing energy increase.

The ideas reflected by these rules that are relevant to motion are found in the other types of rules already.

7.4 Patterns of Agency

Analysis of the programs as elicited knowledge would not be complete without going above the individual rule level. Each of the students' programs paints a conception of the world of motion and it would be valuable to try and map what they seem to be. This section attempts to achieve this by looking at the conceptions of agency described by each program and to compare their features.

As Piaget (1974) has pointed out, the primacy of action and motion as an essential part of our experience from the beginning of life provides the experiential basis for the construction of causal concepts. It was thus no surprise that the notion of causality was a consistently important theme coming up in the classification task protocols, and that most students found little difficulty in writing rules about causes of motion. Most of the programs had the main chunk devoted to the explicit elaboration of conditions for determining the cause or control of motion. However, one of the programs did not contain any rules explicitly on causes of motion -Fong explicitly refused to write about the cause of motion, giving the explanation that the task was too difficult. Nevertheless, he did embark on building a program around the problem of control of motion.

On closer examination, it is found that most of the programs could not distinguish clearly between the two concepts 'cause' and 'control'. Further, irrespective of whichever of these two concepts they were writing on, the ultimate concern was with agency for the motion. Thus it was considered appropriate that a representation of the causal conceptions be extended to include the notions of control. One of the programs, Cheung's, did not explicitly use words like cause or control in the rules. However, an examination of this program reveals that the ultimate concern for the whole program was still the problem of agency for the motions. The present analysis thus tries to compare the main stories the programs presented by looking at the different 'gestalts' (patterns, structural forms) of agency for motion they reflected.

7.4.1 Development of a formalism for representation

In order to extract the conceptions of agency from the programs, it is desirable to develop a scheme of representation that can highlight the causal relations between entities. This imposes two constraints on the design of such a scheme. First, despite differences in the detailed contents, there must be some features that are invariant among causal conceptions. It would be highly desirable for the representation to retain these invariances. Another constraint is that not all the contents of the rules should be represented. Only those ideas relevant to the present focus of concern, agency of motion, ought to be represented to eliminate unnecessary distractions.

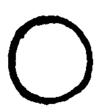
As was pointed out in chapter 3, knowledge representation formalisms can be classified into two broad categories: semantic net types of representation which are essentially very open, and can be used for representing any structure with entities (nodes) and relations between entities (links); and encapsulated knowledge structures like scripts and frames that are used for representing specific kinds of knowledge having well-defined structures. In the context of the present task, the former type of formalism, namely semantic nets, is more appropriate as it allows for the flexibility needed to accommodate the different ideas from all twelve subjects. What is needed here is a representation that can handle entities and different types of relations, and encapsulated script type structures may be imposing unnecessary (and possibly undesirable) restrictions onto the knowledge to be represented.

The first step in the development of the network formalism was to identify the entities/ concepts to be represented. Two entities must be there by definition: object and its motion. Three other major items are cause, control and force/energy (force and energy are collapsed into one entity here since those students mentioning both in their programs did not distinguish between them in any clear way). For ease of comparison, it was decided that these major entities should occupy fixed locations in the representations. 'Motion' and 'object' are respectively placed at the top and bottom of each network. 'Cause' and 'control' are respectively placed on the left and right hand side of the net when they do appear in the program. The position of the entity 'force/energy' is left floating, to be placed at different positions in the representations for different programs in order to reflect more appropriately the actual role played by it in each instance. The network is constructed by recording relations directly concerned with cause, control or force/energy. Relationships concerning the moving object or the motion are also recorded if they relate to the issue of agency for the motion.

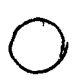
7.4.2 The network notation

Nodes

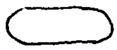
There are five different node notations.



The thick circle is reserved for the four main entities motion, object, cause and control.



Thin circles are for representing events, concrete objects or relations. A thick oval is used to represent the main concept 'force' or 'energy'.

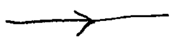


Thin ovals are used for representing abstract attributes and instances of force/energy.

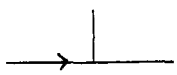
A small circle attached to a line is used for representing properties of a certain entity.

Links

There are six kinds of links in the network.

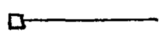


This represents a process relating two entities, e.g. production, participation. The exact nature of the process would be labelled against the arrow sign.

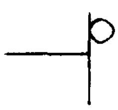


The line coming out of the process arrow is used to link the process with entities that are produced in the process.

This is to link instances of a concept to the concept itself. The small circular terminator end of the link resides in the higher level concept.



This is to relate an entity with its subsets. The end of the link terminated with an empty square resides in the superset entity.



This is to link entities with their properties.

This is to indicate a correspondence, a relationship, between the two links attached to each of the two entities at the two ends of the dotted line.

7.4.3 Main structural forms of the agency gestalt

Several notable features can be identified from the networks. First of all, causality or agency is a main organization concept for motion and the students used it equally for all kinds of motion, from autonomous, to non-autonomous and to those motions where no observable physical agent could be identified, e.g. falling or rolling down an incline.

> *'... the simplest causal relationship between an agent and a recipient of the action presupposes the transmission of something - movement, force, etc.'*
>
> *Piaget (1974), p.12.*

The above statement lucidly describes the essence of all the networks depicting the agency 'gestalts' of the students. However, there is no uniformity in the nature (kind) of entities that are perceived to be the main causes of or agencies for motion. In some instances, the agents referred to were concrete entities like physical objects, natural phenomena (wind, water current), etc. It is difficult to tell, in these cases, whether any mediating transmission was pre-supposed. In most other cases, concrete agencies played a lesser role and some mediation (normally referred to as force or energy) was hypothesized for the causal process.

Another dimension of difference between the networks lies in the perceived roles of the agent and the affected. In most of the networks, there is a clear distinction between the two, the agent being the *active* party and the affected the *passive* one. This distinction in role is, a fortiori, to be found in all the networks where the causal processes depicted are immediate transmissions. Even in the networks where mediation is hypothesized, most perceive the agent as the provider of the mediating entity and the affected as the recipient of this same entity. The following rule from Evon's program is a good example of this line of reasoning:

```
object-A motion caused by object-B if
      object-A initially in contact with object-B and
      object-A force supplied by object-B.
```

However, in two of the programs, the mediating entity was not expressed as directly provided by the agent, but through an event, an interaction, where both the agent and the affected partake. In these cases, the role difference between the agent and the affected seems to dissolve to a large extent and the relationship between the two parties start to look symmetrical. Furthermore, it can be seen from these networks that the agent would also be affected, and that the process would be reciprocal. Fong's program is a good example showing this kind of reasoning and a few rules extracted from his program are quoted here as an illustration:

object-A has a motion controlled by object-B if
object-B move along path-C and
object-A plan path-C and
object-A has the power to influence motion of object-B.

object-A has the power to influence motion of object-B if
object-A directly influence motion of object-B.

...
...

object-A directly influence motion of object-B if
object-B attracted by object-A.

object-B attracted by object-A if
object-B is affected gravitationally by object-A.

object-B is affected gravitationally by object-A if
object-B has mass and
object-B near object-A and
object-A has huge mass.

It can be seen from the above program lines that though it is still possible to distinguish between the agent and the affected, both parties would be affected by the gravitational interaction and if object-A does directly influence object-B, object-B also influences object-A.

According to the two dimensions discussed above, that of media-tion and reciprocity, the networks can be classified into three distinct categories. The following analysis will treat each as a kind of 'gestalt', that is, an organized expectation about types of entity and relationship.

7.4.3.1 The unmediated agency

Three of the networks (fig. 7.3 - 7.5) made no explicit reference to mediated transmissions. The agents referred to in the rules were all concrete physical agencies, e.g. man, wind, etc. In Leung's program (fig. 7.4), though some mention was made of force and energy, these two entities were not seen as related to the causation process. The agency gestalt for Brenda's program is more difficult to classify. In Brenda's network (fig. 7.5), abstract terms like tension, gravity and elasticity appear as control agents. Further, events like 'collision', 'a push' also take up the role of cause agents. However, a careful scrutiny suggests that this network probably belongs more appropriately to this category. In the rules:

object motion controlled by elasticity if
object moving up and down vertically with decreasing amplitude

and

object motion controlled by tension if *object suspended when moving,*

though the agents involved are forces, elasticity and tension, there is no evidence from the program that she sees any difference between these and concrete physical agents. For example, one of her rules about physical agency was as follows:

object motion controlled by water current if
object fluctuating in water.

It seems she might well be using elasticity and tension to stand as a label for the concrete agents that possess particular physical properties. In any case, there is no indication of ascribing a mediating role to the abstract entities as something going between the physical agent and the object in motion.

Similarly, Bren's rules

object motion caused by collision if
object bumping into agent before moving

and

object motion caused by a push if
object initially at rest

indicated more of a wish to specify the necessary situation for the motion to take place than one of hypothesizing an event for mediation.

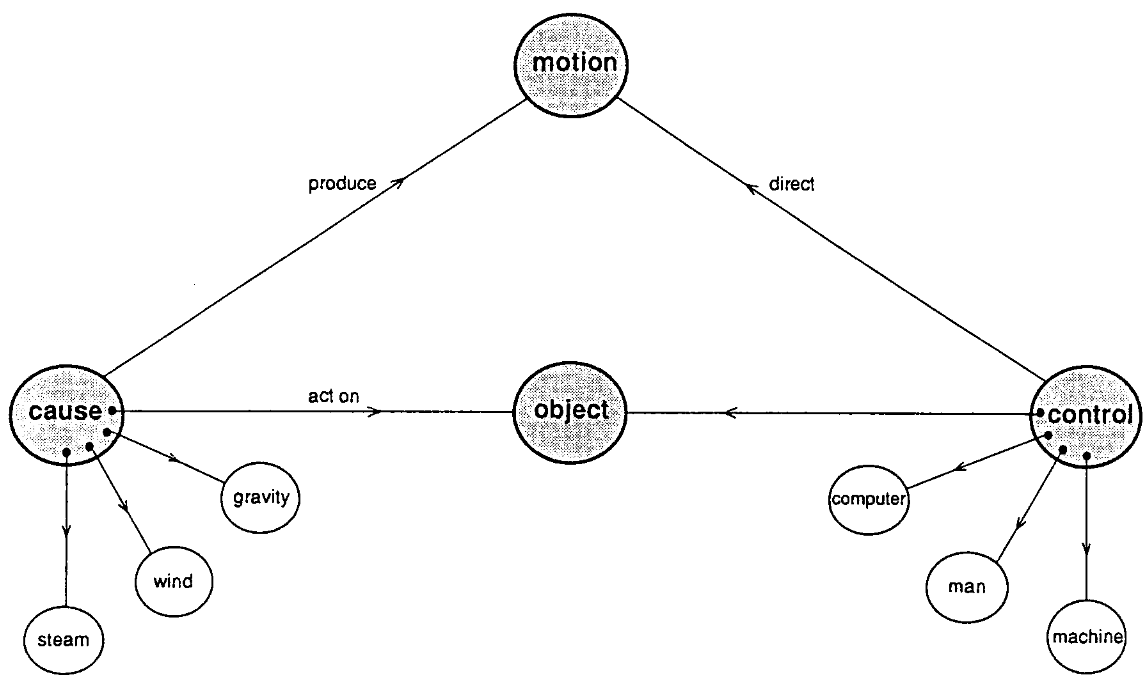


Fig. 7.3 The agency gestalt as evidenced by Sandra's program.

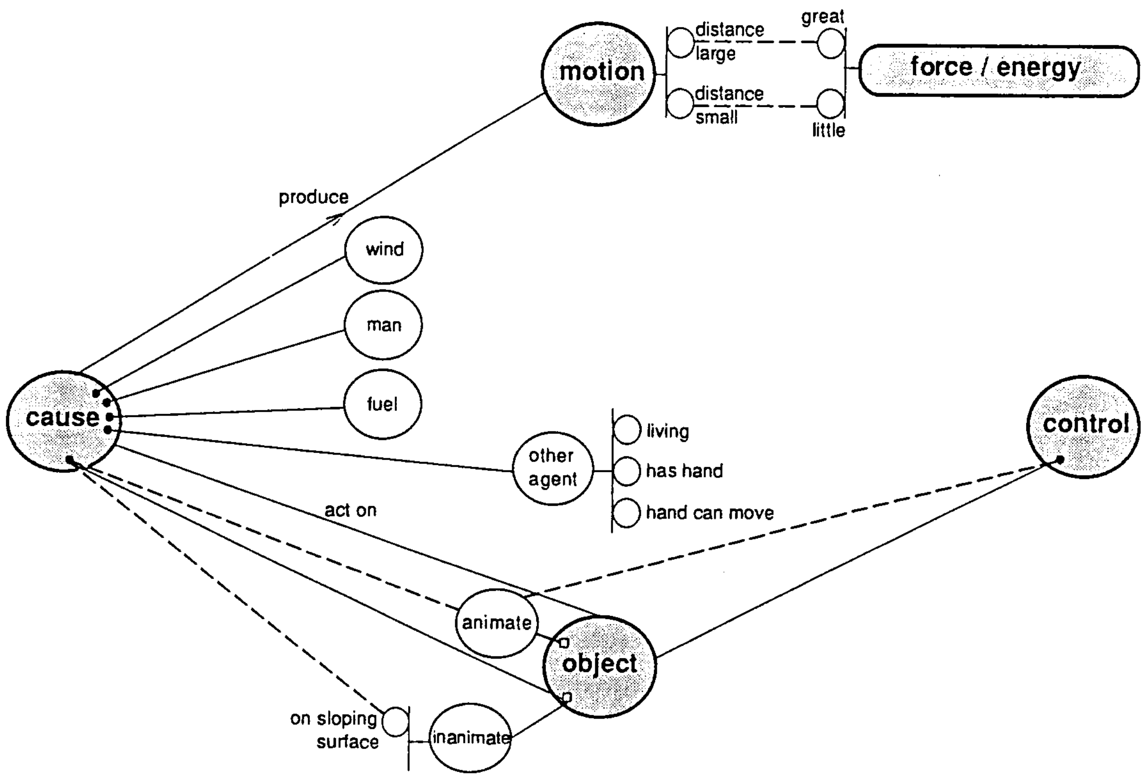


Fig. 7.4 The agency gestalt as evidenced by Leung's program.

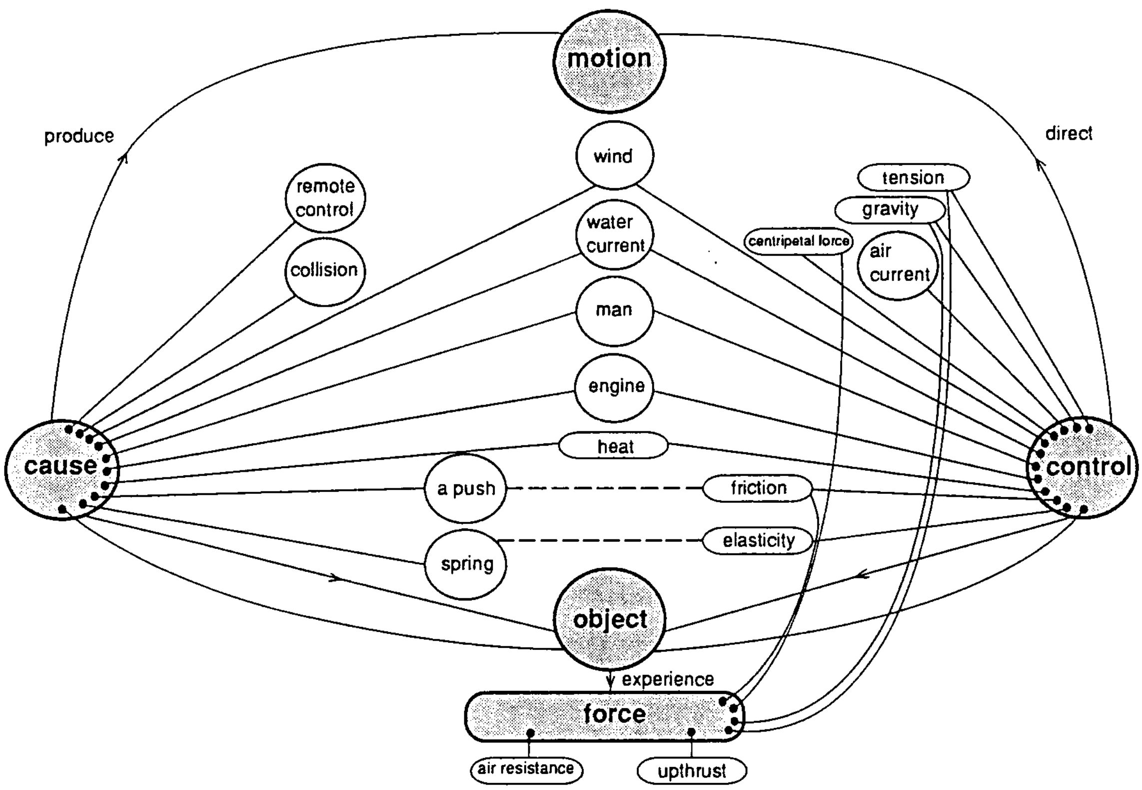


Fig. 7.5 The agency gestalt as evidenced by Brenda's program.

7.4.3.2 Mediation as a central feature of agency

More than half of the networks fall into this category (fig. 7.6 - 7.12). These reflect two different conceptions of mediation. One subgroup (fig.7.6 - 7.10) still sees the agent as concrete physical entities, but imposing a constraint that they be the providers of a mediation, normally force, which is a necessary ingredient for the sustenance of motion. One common feature of this subgroup of networks is that they tend to classify the mediations, attributing different categories of mediations to different characteristics of the resulting motions. For example, greater force/energy is attributed to greater motions or faster speeds. Similarly, 'external forces' are attributed as responsible for non-autonomous motions, while 'own force' or 'internal force' accounts for autonomous ones.

Another subgroup (fig. 7.11, 7.12) within this category discounted concrete physical agents altogether in their direct references to agency. The mediation, force or energy, is seen as the essence for the explanation of agency. It is interesting to note that Wong explicates a clear distinction between the two agency concepts 'cause' and 'control'. He attributed causal agency to force and agency for control to concrete physical entities. Furthermore, there is nothing conceptually deviant from the Newtonian framework in his agency network since he clearly confined the need for a cause to changes in motion rather than for motion itself.

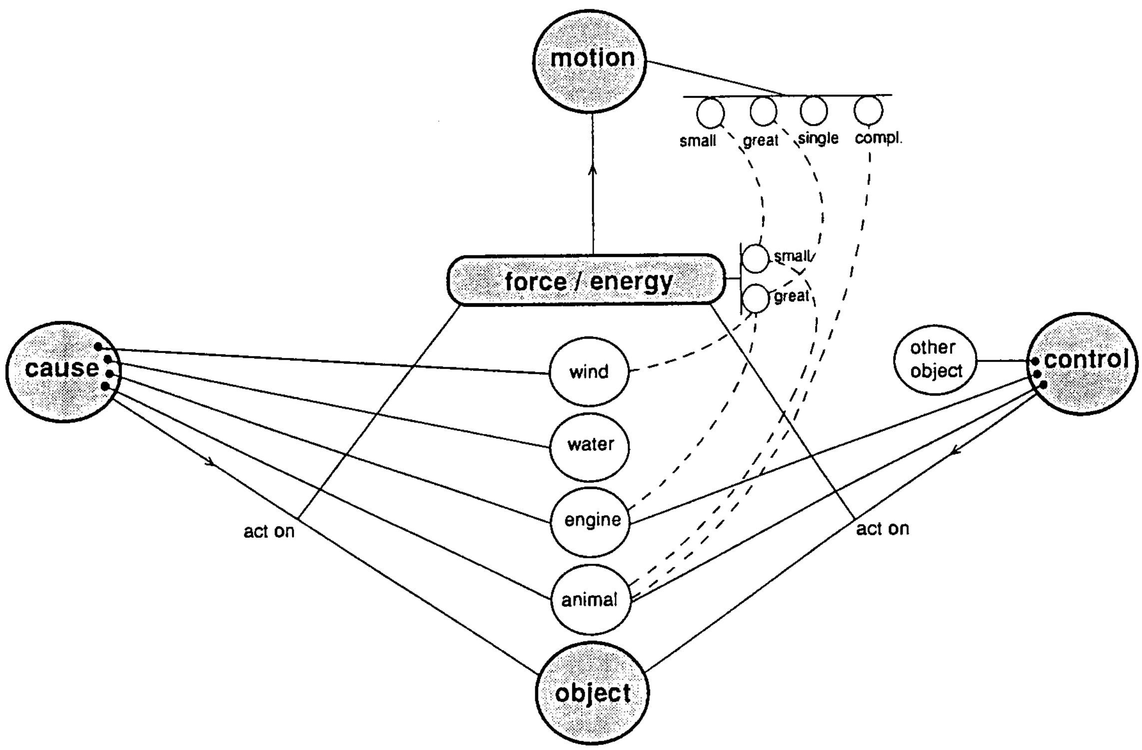


Fig. 7.6 The agency gestalt as evidenced by Schan's program.

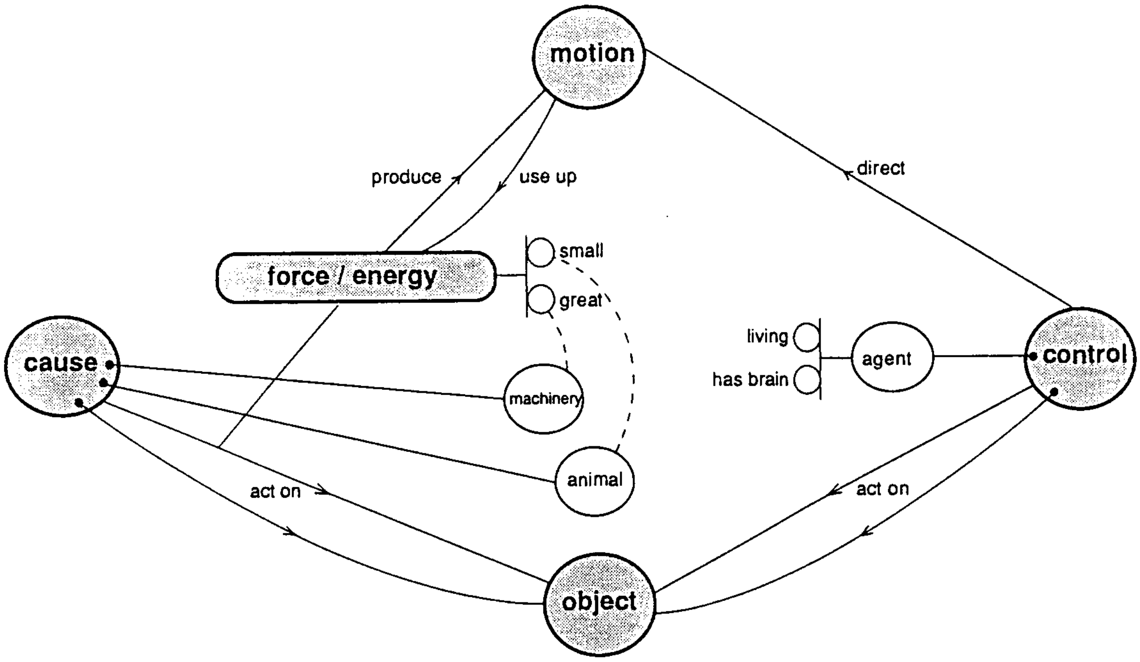


Fig. 7.7 The agency gestalt as evidenced by Ychan's program.

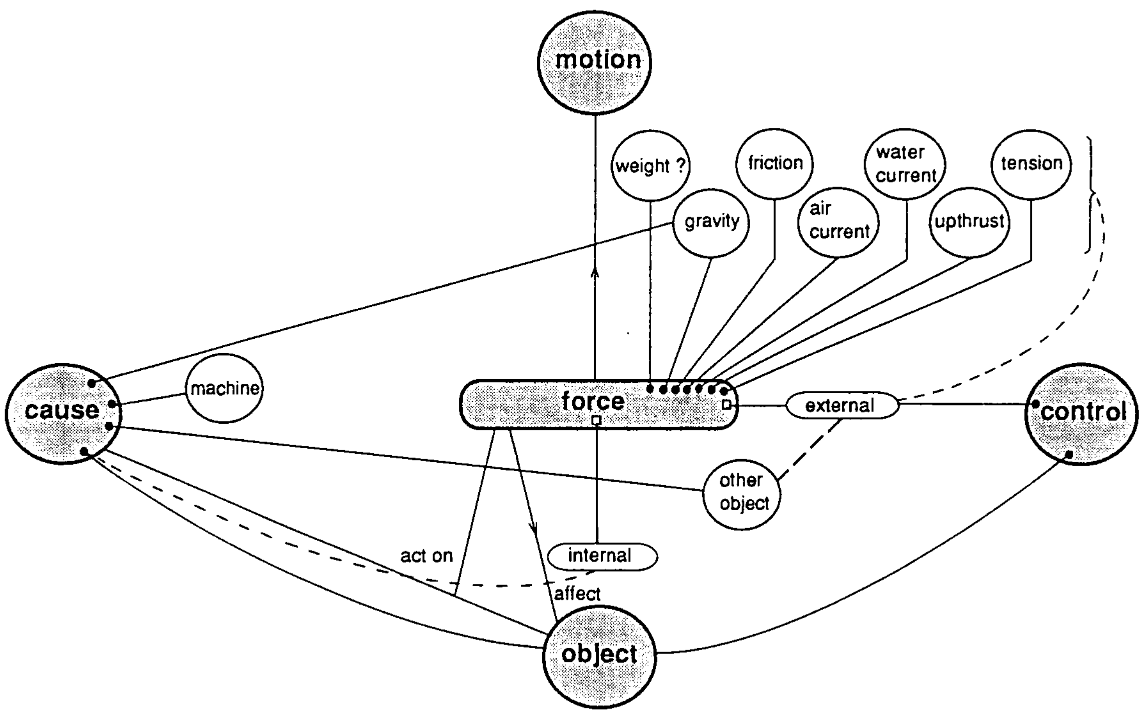


Fig. 7.8 The agency gestalt as evidenced by Evon's program.

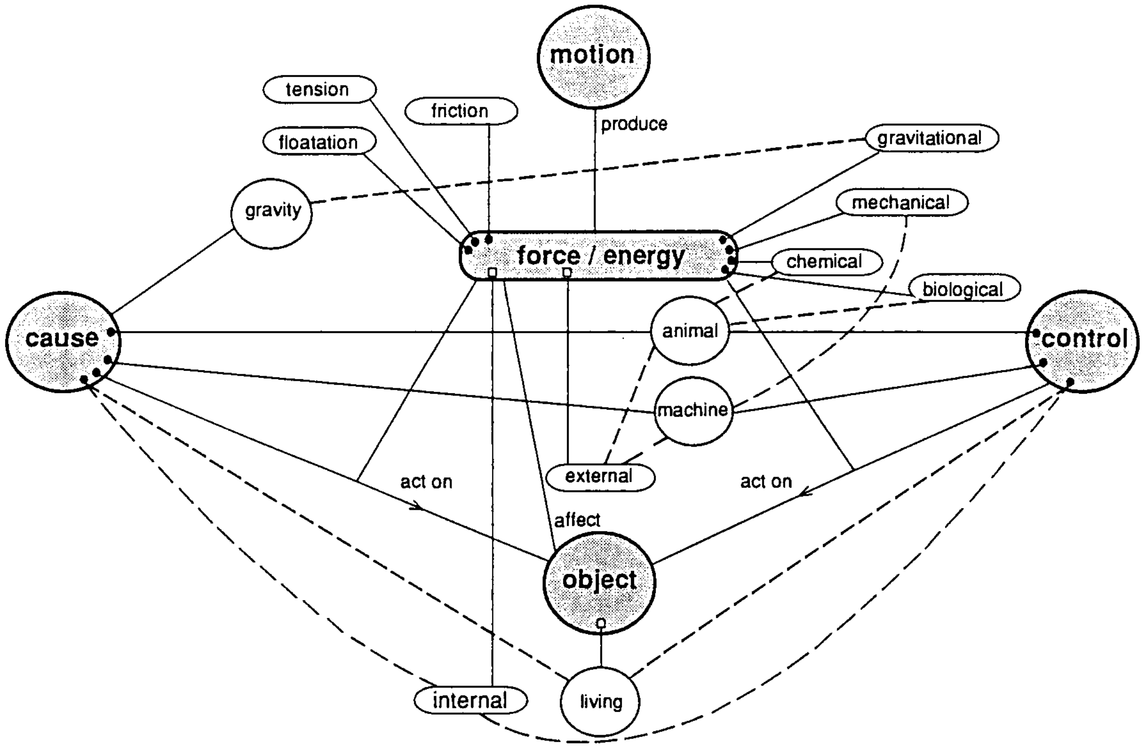


Fig. 7.9 The agency gestalt as evidenced by Lin's program.

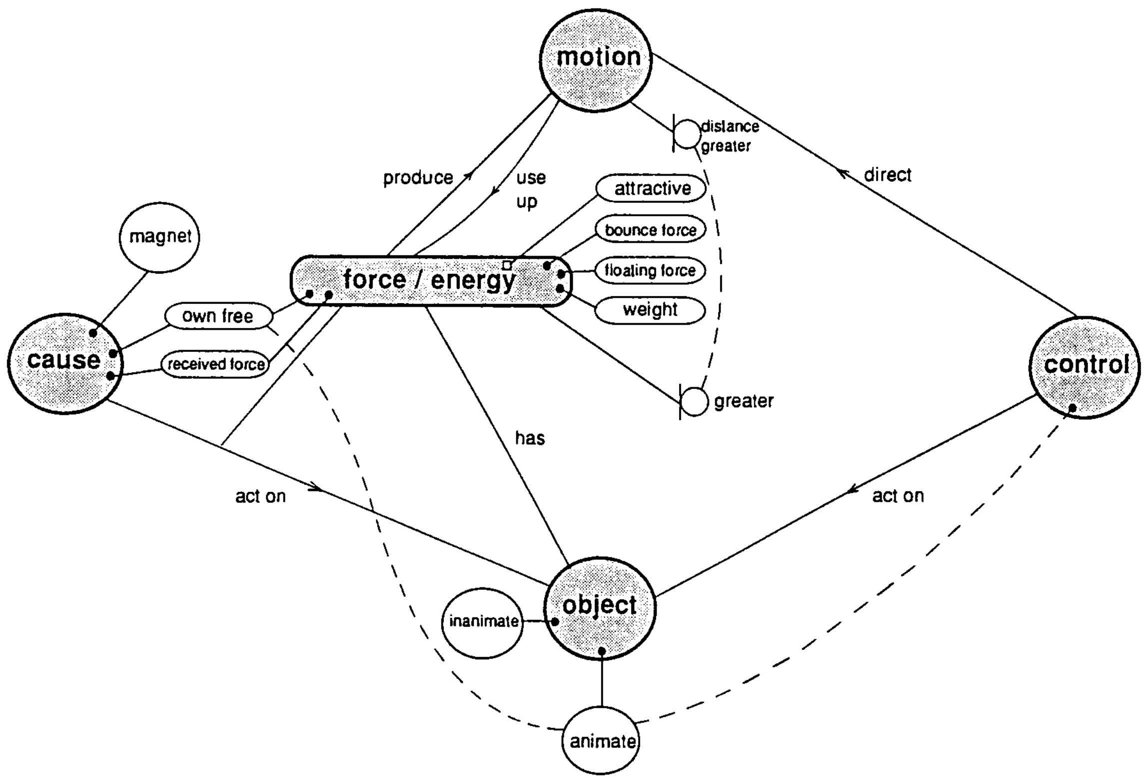


Fig. 7.10 The agency gestalt as evidenced by Wat's program.

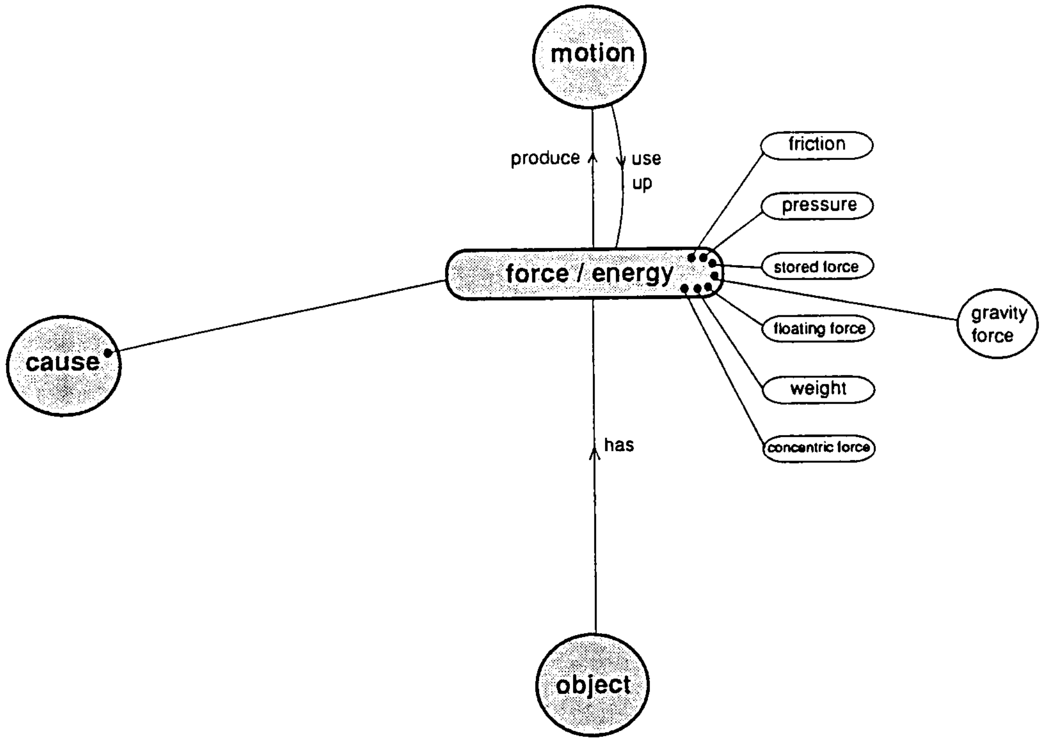


Fig. 7.11 The agency gestalt as evidenced by Cheng's program.

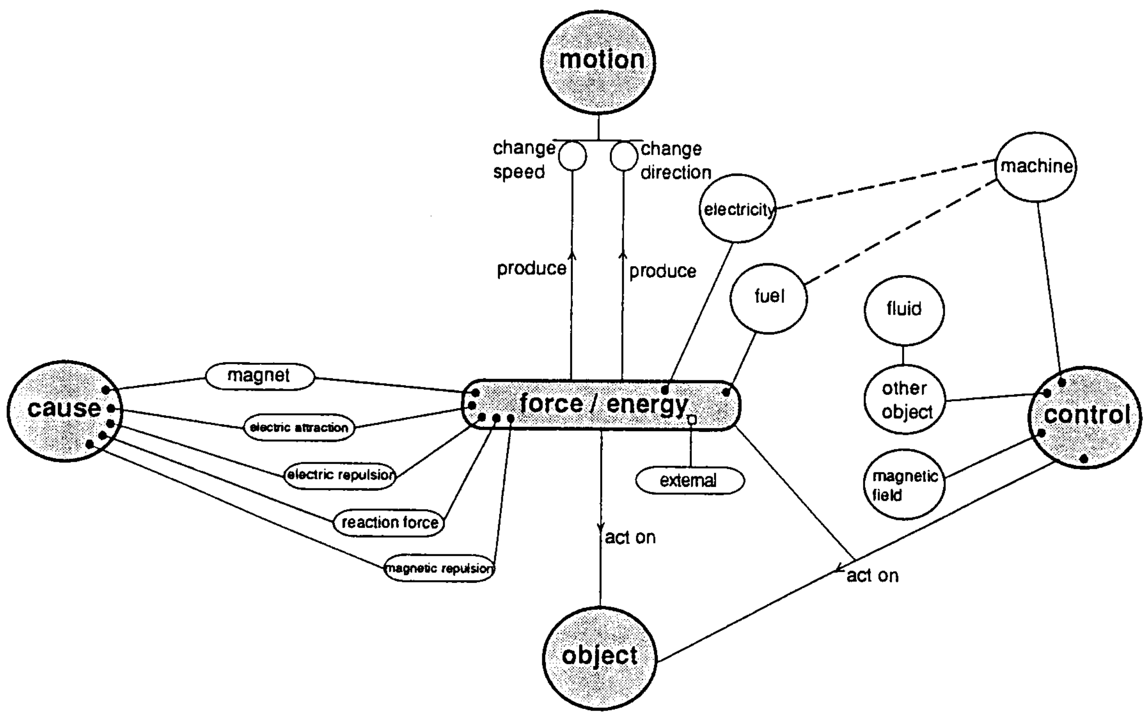


Fig. 7.12 The agency gestalt as evidenced by Wong's program.

7.4.3.3 The symmetrical agency

Two of the networks, one from each age group, are quite exceptional in that they present a conception of causation which encompasses a notion of reciprocity. This was accomplished, as mentioned earlier, through two steps. First, both assumed a priori the central role of a mediation. Second, both focussed on an event and held that as directly responsible for the production of the mediating entity, rather than the physical objects perse. One special feature of this conception of agency is that it encompasses the role of physical agents (as opposed to the second subgroup in the previous category, confining the causal conception to abstract entities), and at the same time did not have to attribute the mediation as something to be possessed by and resident in a physical object. This feature is evident in Fong's work as discussed earlier. The same is found in Cheung's program:

```
thing give force to other-thing if
      other-thing touch thing.

thing give force to other-thing if
      other-thing hit thing.

thing give force to other-thing if
      thing placed on other-thing and
      thing affected by earth.

thing give force to other thing if
      other-thing placed on thing and
      other-thing affected by earth.

other-thing give force to thing if
      thing hit other-thing.
```

Though Cheung still used words like "give force to", he clearly and explicitly identified the "giving" in each case to be in the context of an event, a situation, whereas in the other programs, the "giving" was mentioned purely as a transaction between the agent and the affected, paying no attention to the situation for the interaction.

Doubtless, the most prominent feature of these two networks is the encapsulated notion of reciprocity: it is implicit in the networks that the two objects involved in the event producing the mediation are both affected by the event. This change is in fact a big step forward towards a Newtonian conception of motion. Commonsense intuitions normally associate agency with a strong notion of potency and an active role. The symmetrical role of objects participating in an event producing the mediating entity for motion as depicted by these 'gestalts' exhibit a detachment of the agent from an active role. It is possible that such a conception of agency is more congenial to the possible adoption of Newton's third law of motion into the cognitive framework.

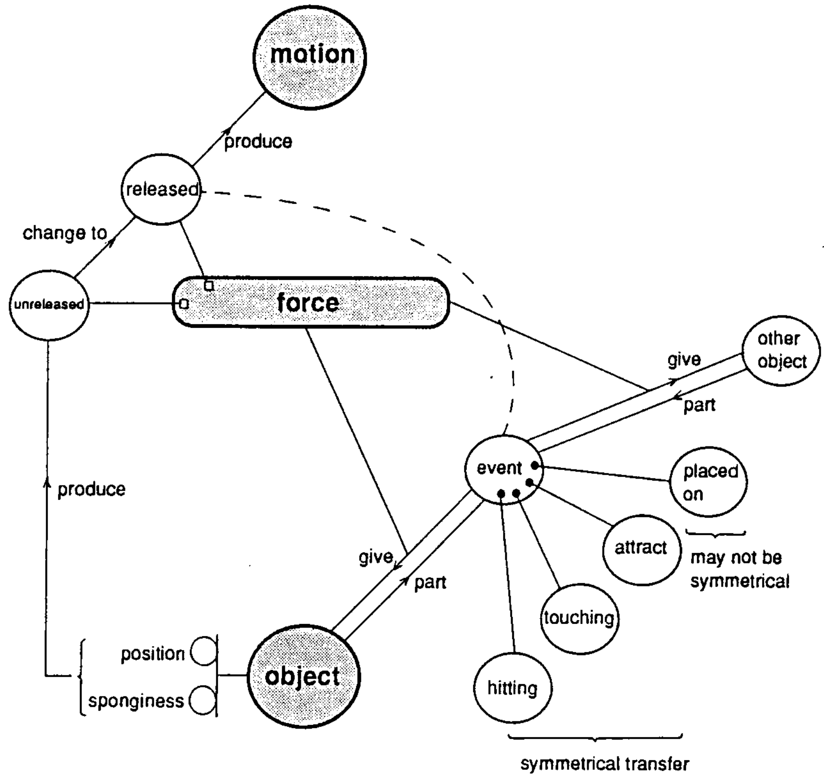


Fig. 7.13 The agency gestalt as evidenced by Cheung's program.

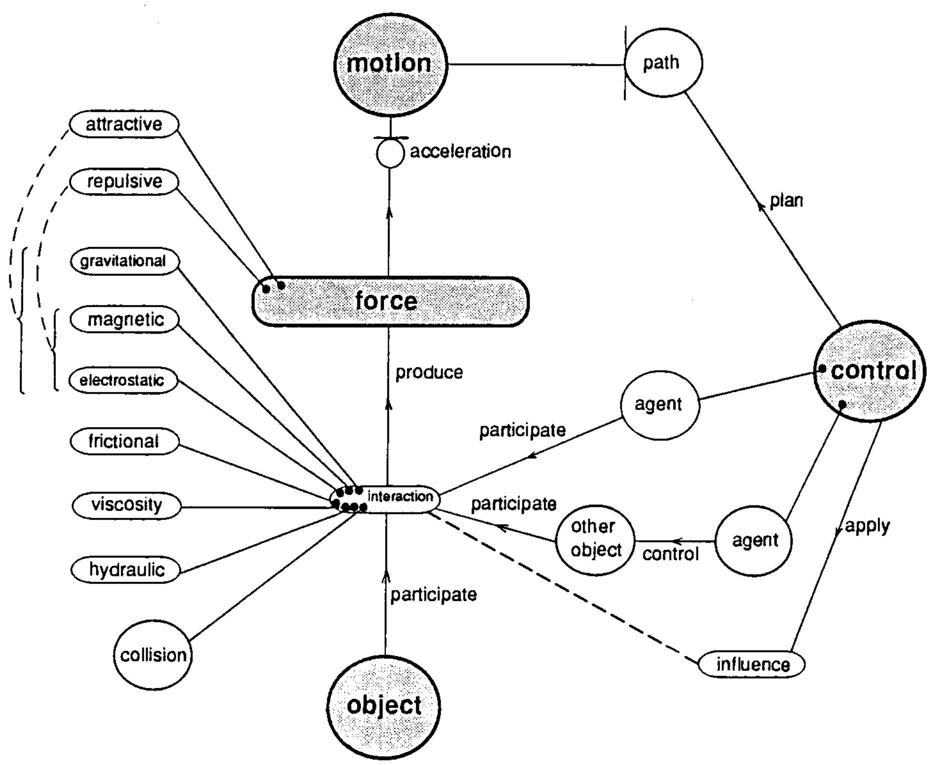


Fig. 7.14 The agency gestalt as evidenced by Fong's program.

The above mentioned feature of detaching the attribution of agency from a physical entity towards an interacting event is not completely accomplished in Cheung's network. There are some instances where the force held responsible was not derivable from a visible event, e.g. when an object becomes unsupported or when a compressed spring is released. In such cases, he attributed the force to be derived from properties of the object or the situation. Possibly because of his better grounding in mechanics, Fong had overcome this difficulty and been able to attribute different forces to different kinds of interactions. Again, there is nothing explicitly deviant from Newtonian mechanics in Fong's program. However, in comparison with Wong's program (fig. 7.12), Fong's network (fig. 7.14) is a much richer and closer representation of the Newtonian conception of motion.

7.4.4 Further remarks on the agency gestalts

It is interesting to note that each of the three agency 'gestalts' was held by students from both age groups. The structure of the agency gestalts thus seems to be independent of the cognitive maturity of the students or the amount of mechanics instructions they received. It is the claim here that the difference between the groups of students lies not in their general level of development but in their locus of attention when reflecting on causal events.

A possible explanation for some of the networks, not to mention mediation at all, could be that the students hold different views on what kind of knowledge is valuable. Irrespective of the exact contents of the networks, each provides some useful information about the world of motion. However, their utility differs depending on the purpose at hand. If the knowledge is to be used for solving the pragmatic concerns of everyday locomotion and prediction of tendencies, knowledge of the form imparted by the unmediated agency gestalts would be the most expedient. This is given support by the rating given to the performance of individual programs by the students after they had completed the programming task (details of this rating exercise are described and discussed in the next chapter). If we define efficiency as the time taken for a goal to be resolved and performance as the number of times the results turn out to be the expected (most reasonable) ones, then the best program in both respects turns out to be the one belonging to the unmediated agency group (Brenda's program, fig. 7.5). This 'objective' rating concurs with the participating students' 'subjective' rating. All but one of the students rated this particular program as the best.

It is thus understandable if the students whose networks belong to the unmediated agency group did not express concepts of mediation simply because they did not think that they were important. The concepts of mediation would only be important if one wants to do a more systematic analysis of motion events in general, venturing out of the pragmatic domain into the theoretical.

With reference to the notion of reciprocity, it is argued here that this notion as expressed in Fong and Cheung's networks is much more sophisticated than the notion of reciprocity Piaget discussed with reference to the children's conceptual development, so much so that they could justifiably be labelled as qualitatively different. The reciprocity that Piaget referred to was just an awareness that the recipient (or the affected) exhibits a resistance to the action, while the essence of the networks in fig. 7.13 & 7.14 was detachment of agency from the object to the event, thereby creating a symmetrical structure in the causal gestalt. This introduction of a symmetrical element is a very important conceptual jump which is not a necessary development of maturation. It is closer to the kind of conceptual jumps that is characteristic of scientific revolutions.

One pedagogical hypothesis coming out of this analysis is that one should try to explicitly relate Newtonian mechanics to the commonsense gestalt of causation and to anchor the relationship onto interacting events rather than the physical objects. This point will be discussed further in the final chapter.

7.5 Levels of Explanation the Programs offered

As a product of the knowledge elicitation process, what do the programs offer to the researcher besides a glimpse of what issues the students concern themselves with when considering motion? Is there anything interesting besides the actual contents of the programs? The answer to such a question depends on one's view of the nature of the programming process as a cognitive task. The process of program development is a creative one. It sets up a context very different from that of a problem solving task or the classification task described in an earlier chapter. The students have virtually complete freedom in deciding what they wish to write down. Even when they were encouraged to focus on certain themes, causation and control, in the later part of the programming sessions, they were not only free whether to take up that suggestion or not, but they were also free to write about these themes in any way they liked. Given such an unstructured task, how would the students go about structuring it? It is the view here that the programming task is essentially an explanation task. In trying to write down general rules about motion, they were in effect trying to offer explanations of how the world of motion works in general.

In explaining how something functions, there is no fixed criterion as to what counts as an adequate explanation. Depending on the depth of understanding the explainer has of the subject domain and the purpose at hand, the explanation may vary from the very specific to very general. For example, to explain how to make an object move faster, the advice given

Table 7.4 Distribution of rule types.

(a) Distribution for individual students

Rule Type	Student Group												Total
	F.6 group						F.3 group						
	wong	evon	lin	bren	sand	fong	cheng	schan	ychan	ahlun	leung	wat	
Empirical observation	12	0	3	0	12	0	14	4	9	16	14	10	94
	44%	0%	23%	0%	100%	0%	29%	12%	21%	35%	37%	26%	24%
Specific conjecture	11	7	5	36	0	26	20	10	13	13	15	8	164
	41%	39%	38%	97%	0%	63%	41%	30%	30%	28%	39%	21%	42%
General Hypothesis	2	10	5	0	0	4	10	14	13	11	8	14	91
	7%	56%	38%	0%	0%	10%	20%	42%	30%	24%	21%	37%	23%
Classification	0	0	0	0	0	0	4	0	5	0	0	3	12
	0%	0%	0%	0%	0%	0%	8%	0%	12%	0%	0%	8%	3%
Definition	2	1	0	1	0	11	1	5	3	6	1	3	34
	7%	6%	0%	3%	0%	27%	2%	15%	7%	13%	3%	8%	9%
Total	27	18	13	37	12	41	49	33	43	46	38	38	395

(b) Distribution for each age group.

Rule Type	Student Group			
	F.6		F.3	
Empirical observation	27	18%	67	27%
Specific conjecture	85	57%	79	32%
General Hypothesis	21	14%	70	28%
Classification	0	0%	12	5%
Definition	15	10%	19	8%
Total	148		247	

can be very specific such as giving it a push, or a general one, such as to find a means of applying a net force to the moving object along the direction in which the object is moving. As the programming task was an extended one both in time span and duration, and the focus was on writing general rules, one might hope that the resulting programs could reflect the deepest level of cognitive understanding the subject has in the domain of every-day dealings with motion. The analysis in this section tries to examine the PROLOG rules written by the students in terms of the degree of generality they offer as an explanation.

Excluding those offering classifications and definitions, the PROLOG rules written by the students can roughly be divided into three levels of explanation along the dimension of generality. At the lowest (most specific) level, there are the *empirical observations*, describing specific situations or concrete entities. Next are the *specific conjectures*, hypothesizing the role of abstract entities in specific situations. At the highest (most general) level come the *general hypotheses*, postulating general relationships between entities, abstract or concrete, that would be applicable to motions in general. This part of the analysis tries to classify the PROLOG rules into these three levels of explanation and to examine and compare the performance of the two groups of students according to this perspective.

7.5.1 Empirical Observations

These are empirical rules describing relationships between concrete entities or attributes for specific motions or specific situations. There are no abstract quantities or attributes involved. The most frequent candidates for such empirical rules are rules of the following formats:

i) *motion type* <- *observable physical conditions,*

for example,

```
thing will rebounce if
      thing hit other-thing and
      not (thing can stick to other-thing);
```

ii) *concrete causal agent* <- *observable physical conditions,*

for example,

```
thing motion caused by wind if
      thing moving unsteadily and periodically in air;
```

iii) *concrete relations about specific events/situations,*

for example,

```
thing may be damaged if
      thing crashed by other-thing.
```

These rules require very little in terms of cognitive processing to be created. They demand only direct recall of prominent features of particular events. As these rules are so specific, they are of little utility unless those specific types of motion or situation were under consideration. Thus the catalogue of motion types described in such rules might be seen as important prototypes of motion for the students' everyday processing of motion information. As explained in chapter 5, the encapsulated knowledge structures (scripts) for such prototypes would play the role of allowing easy recognition and hence facilitate analogous handling of the situation once the mapping has been made. The most popular types of motion appearing in this category of rules bear a close relationship with those appearing in the descriptive D2 schemes in the classification task. These include motion types like falling, colliding, flying, bouncing, etc.

The results shown in Table 7.4 suggest that the younger group of students were much more willing to write down rules at this level. Rules in this category constitute about one third of the total number of rules written by the F.3 students exclusive of classifications and definitions, while making up only one fifth of the F.6 students' corresponding total "production". But such a gross description is not adequate. The F.3 students also displayed a much more uniform tendency towards offering this level of explanation: all the students in this group had written some rules in this category, the minimum of which constituted 12% of the individual student's total output. The F.6 group's behaviour is extremely diverse. Half of the group did not write down any single rule in this category while one of them had all of her rules in this category.

7.5.2 Specific Conjectures

At the next level, there are rules which postulate the existence or functioning of some abstract entities, or possible causal or control agents, but with reference only to specific types of motion or specific situations. The most popular formats for such rules are as follows:

i) *abstract motion attribute* <- *observable physical conditions,*
for example,

thing affected by tension if
thing attached to elastic string;

ii) *concrete causal agent* <- *observable physical conditions,*
for example,

thing motion caused by a push if
thing initially at rest;

iii) *concrete control agent <- observable physical conditions,*

for example,

motion is controlled by self if
object moving is B and
external force acting on B and
external force less than maximum force exertable by B and
B can respond to stimulation and
B can be aware of occurrence of motion.

The F.6 group produced more rules belonging to this category than the younger ones. 57% of the older groups' total rule "production" were in this category, whereas rules at this level make up only 32% of the younger group's total production. At this level, there is much more uniformity within each of the two groups, though one student in the F.6 group did not write any rule in this category.

A possible explanation for the F.6 group's greater tendency to produce these 'specific conjectures' is that the older students are primed to do so through the formal physics instruction they have received. In dynamics, the students are taught to analyze the nature and magnitude of forces acting on different objects in specific settings. They may thus be more prone to seeing this as more acceptable and meaningful. Further, they are equipped with a better vocabulary for writing specific conjectures, as reflected in the technical jargons they employed in these rules.

7.5.3 General Hypothesis

The highest level of explanation a stand alone rule can offer is a general hypothesis. Such a rule needs to be a generalization that applies to motions in general. To write such a rule, there must be some guiding principles or theoretical assumptions that the person holds which underlie their whole conception and understanding of motions in general. They thus involve a high level of abstraction and structuring of facts and beliefs. The more popular formats of rules in this category were:

i) *concrete motion attribute <- abstract attributes,*

for example,

thing speed increase if
thing force increase and
thing energy increase;

ii) *abstract motion attribute <- observable physical conditions,*

for example,

thing has force if
thing state moving;

iii) *abstract relations about theoretical attributes,*

for example,

thing can produce force if
thing already has energy, and
thing can receive force.

One very curious and important observation to note here is that the older group is producing a lot less rules at the general hypothesis level compared with the younger ones: a total of 21 rules (14% of total rule production) for the former group to 70 rules (28%) for the latter. However, as the subsequent analysis will try to explain, this does not mean that the younger group has higher abstraction or generalization ability. Furthermore, at this level there is again a large diversity in the performance of the students within the F.6 group, with two of them producing nothing at all in this category.

7.5.4 Comparison of the two Age Groups

Diversity versus relative uniformity is the main feature when one compares the distribution of rules written by the two age groups amongst the three levels of explanation, vis-a-vis empirical observations, specific conjectures and general hypotheses. There are two aspects to this diversity. Firstly the distribution of the rules among the three levels for individual students was much more uniform for the F.3 students than for the F.6 group. Most of the rules written by the latter group tend to concentrate at the level of specific conjectures. A second diversity lies in the variation in distribution of rule levels across students. Two of the students effectively wrote rules at one single level of explanation. It would be useful to explore the possible reasons for such a phenomenon.

The main difference between the two groups of students with respect to the task at hand lies in the fact that the older group had received two year's instruction in Physics at O-Level and were still receiving instructions in Physics at A-Level while the younger group had not received any instruction on mechanics. Is it possible that the differences in rule level distribution between these two groups are due to this difference in the mechanics instruction they have received?

The absence of rules at the empirical observation level is not too difficult to explain and could just be an indication of intellectual maturity. It would be understandable if the older students saw more value in abstraction and generality. A training in mechanics could have helped the older group to find more to write about at the specific conjecture level. The large number of technical terms, such as names of specific forces, appearing in this group's rules at this level offers support for this argument. The most puzzling feature in the results is the fact that the older group wrote a much smaller number of rules at the general hypothesis level, some even to the extent of producing none in this category. If anything, one would expect mechanics to be offering a view of motion that is more general than naive physics and it is the latter which is commonly characterized as fragmented and inconsistent (Claxton, 1985). To explore this phenomenon further, one needs to look at the contents of the general hypotheses written by the students.

At the content level, the general hypotheses written by the younger students are very similar, and most of them explicate ideas that would be deeply frowned upon by physics teachers. The following is a representative selection of these rules.

motion can be great if
    object has great force

thing can start moving if
    thing can produce force

thing motion direction unchanged if
    thing only change force

thing has force if
    thing is moving

thing1 give force to thing2 if
    thing1 is moving and
    thing1 touch thing2

thing given force by other-thing if
    thing movement helped by other-thing

thing can receive force

thing can control own motion if
    thing is living object and
    thing has brain and
    thing brain normal

thing motion caused by own force if
    not thing is dead object and*
    speed changeable by thing and
    amount of force changeable by thing

thing motion caused by received force if
    thing receive force from other-thing and
    not other-thing is dead object and*
    force direction changeable by other-thing

* dead object here refers to inanimate objects

It seems that the older students have learnt enough Physics to refrain from writing so profusely about force as the necessary 'fuel' of motion that can be possessed, passed along and used up, nor do they express such outright animism. Diversity in the performance of the F.6 students existed also at the qualitative level. Important individual differences are found in the contents of the general hypotheses within this older group. They expressed different interpretations and levels of integration of learnt physics concepts into their commonsense framework of motion.

Wong wrote only two rules belonging to this category:

> *object will change its direction if*
> *there is external force applied on object,*

and

> *object will change its speed if*
> *there is external force applied on object.*

These can be interpreted as another form of Newton's first law. However, these two propositions have no direct relationship with the other rules in his program and they stand alone as isolated statements and would not in fact be called into play in the resolution of the other queries. Thus they exist in the program but are never used. Is this also the case in his thinking?

Evon and Lin's general hypotheses were very similar in the ideas they express: classifying causes as being due to two broad categories of forces, internal and external. They have conflated the scientific terms internal and external forces with their concept of autonomous and non-autonomous motions. They thus referred to internal and external sources of supply of force as the required ingredients for different kinds of motion. Examples of such rules are:

> *object cause of motion self if*
> *not (object needs external force in causing motion) and*
> *object source of force internal,*
>
> *object cause of motion machine if*
> *object make use of specially designed instrument and*
> *object source of force external and*
> *object type of force mechanical,*
>
> *object cause of motion animal if*
> *object source of motion external and*
> *object type of force biological and*
> *object source of energy chemical.*

The concern with autonomy, the distinction between animate and inanimate objects, and the belief that force is the fundamental cause of motion is still very evident and are very similar to that of the F.3 students. However, they are wrapped up and integrated with what they have learnt in their formal physics - some kind of marriage of the two theoretical frameworks (there is a discussion of the effect of these two frameworks interacting within the context of testing and modification of the PROLOG programs in the next chapter).

Fong's work is distinctly different from the others and indicates a high level of integration of Newtonian concepts into his commonsense framework for the resolution of the pragmatic concern of control of motion. His four general hypotheses were:

object1 has a motion controlled by object2 if
object1 move along path P and
object2 plan the path P and
object2 has the power to influence motion of object1,

object1 has the power to influence motion of object2 if
object1 directly influence motion of object2,

object1 has the power to influence motion of object2 if
object1 indirectly influence motion of object2,

object1 indirectly influence motion of object2 if
object3 directly influence motion of object2 and
object1 has the power to influence motion of object3.

The rest of his program was mainly concerned with defining ways in which one object can directly influence another object's motion, and that was done through building specific rules about the interaction forces in each case.

7.6 Epistemic Levels of the Elicited Knowledge Bases

In the previous section, an analysis was made at the level of rules, looking at the level of explanation that each of the rules taken independently offered. It is possible to offer a deeper level of analysis by looking at each of the programs as an integrated whole. Each program, though finite and arguably limited in scope, when taken as a whole may paint a picture of a student's world view about motion. If we do see the programming task as essentially an explanation task, as suggested in an earlier section, the complete program can be taken as an explanation for the kind of questions the student is answering. Implicit in this elicited knowledge are the fundamental levels of generality the student was working with, and the kinds of questions s/he was answering, and these in turn reflect the kind of knowledge the student figured as most important for understanding the world, that is, the sort of *epistemic questions* s/he was concerned with.

This section explores the problem of what kind of epistemic questions were tackled by the students in the programming task by looking at how the different levels of rules as identified in the last section orchestrate together. We may not be able to tell too much about the main concerns of the authors of the programs, whether they are mainly interested in general issues or specific instances, just from looking at the percentage distribution of the different kinds of rules. If the rules in a program in any sense form an integrated knowledge base, then it is reasonable to expect that the importance of the individual rules are not all the same. Rules that stand in isolation are normally of limited value to the system and those whose influence proliferate through other rules would be much more central to the cognitive functioning of the whole system. The analysis in this section thus tries to look at the program as a complete system in order to find out the level of generality it is working at, and from there to the epistemic problems tackled by the students. It is also argued here that the different epistemic frameworks identified represent a progression that necessarily accompanies different phases of development in a progression towards the development of a scientific understanding or model.

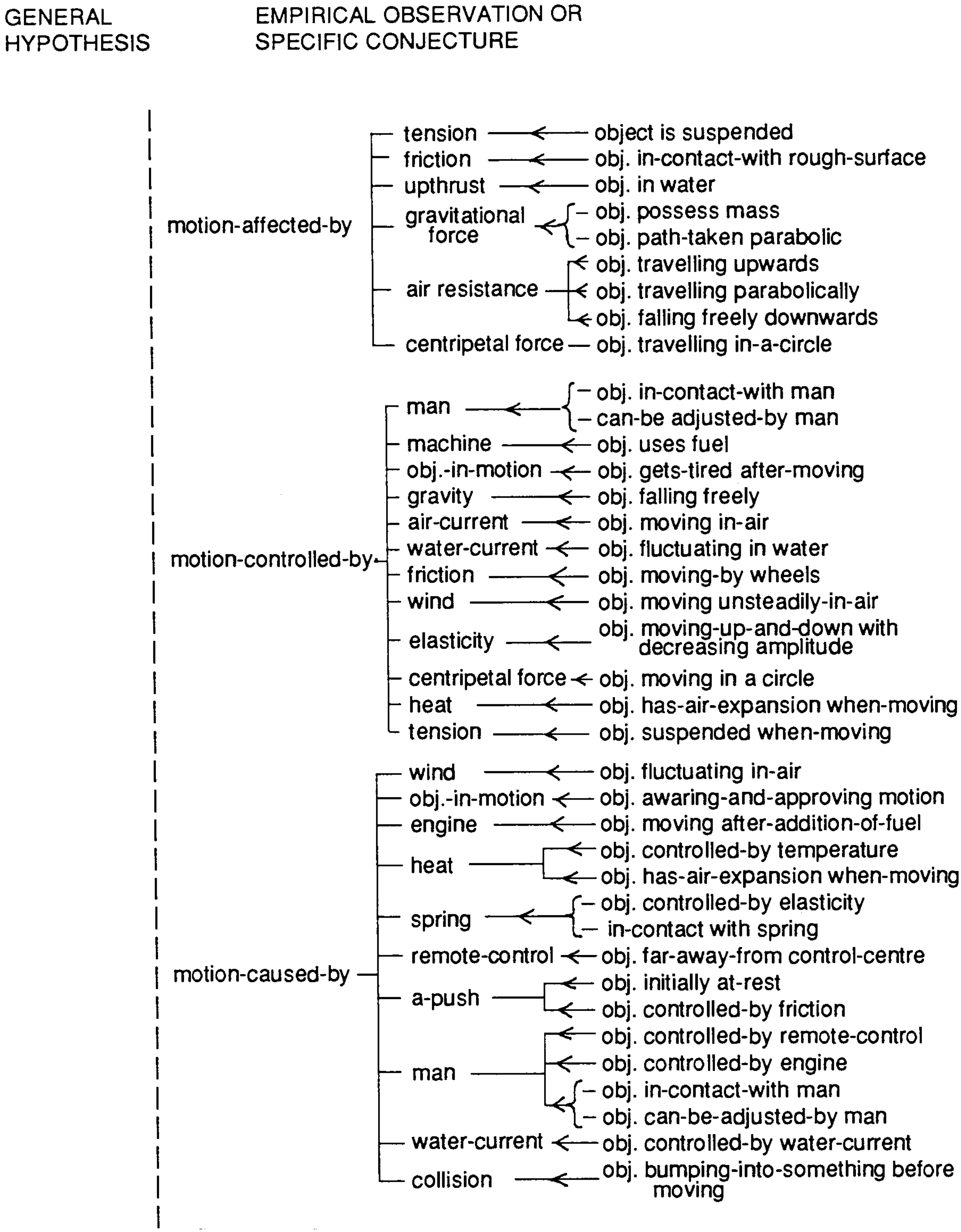
7.6.1 'Specific' epistemic frame

First, let us start by looking at the level of generality implicit in the different kinds of rules. For the category *empirical observations*, the level of generality is necessarily a specific motion or a specific situation since it describes the relations between different concrete parameters, e.g. velocity, trajectory shape, motion direction etc., for a specific kind of motion event or situation. The category *specific conjectures* also shares this same level of generality since it differs from the first category only in the inclusion of abstract parameters into the relations. The focus was still on a specific events or situations. Good examples of these are:

thing move upward if
thing is living object and
thing has wings (empirical observation
on specific situation)

thing path shape parabola if
thing has bounce force(specific conjecture
on specific motion type)

For a program that contains exclusively empirical observations and/or specific conjectures, as Brenda's or Sandra's was (see fig. 7.15 & 7.16), it seems fair to say that the epistemic question tackled here was at the level of specific occurrences. Each case was dealt with independently. It was a world in which the clue to the pragmatic concerns of predictability and regularity was sought from properties and attributes of the particular situation. This is not to imply that neither Brenda nor Sandra entertain analyses of a more general nature. The claim is that the programs reflect where their priorities, confidence and even values lie.



legend: ←< denotes an implication. Thus before '←<' is the goal statement and after that is the condition.

is taken to represent disjunctions. Thus the above figure represents 12 rules on 'motion-controlled-by and 2 rules on 'motion-caused-by heat'

represents a conjunction on the condition side of a rule.

Fig 7.15 Outline of Brenda's program

GENERAL HYPOTHESIS | EMPIRICAL OBSERVATION OR SPECIFIC CONJECTURE

- motion caused-by
  - gravity
    - move in gravitational field
    - move downwards
    - not(controlled by obj.-in-motion)
    - affected by mass of obj.-in-motion
    - speed decreasing
    - obj.-in.-motion will reach ground
  - wind
    - obj.-in-motion placed in open area
    - obj. has large-surface-area
    - affected-by weather
    - obj. material light
  - steam
    - motion need heat-energy
    - motion need water
    - causing-agent need high-temp.
  - object-in-motion
    - motion controlled-by obj.-in-motion
    - energy supplied-by obj.-in-motion
    - motion can-change on-will-of-obj.
- motion controlled-by
  - computer
    - control-agent need electricity
    - motion can-repeat many-times
    - control-agent can-store programs
    - control-agent has central-memory
    - information input manually
    - not(motion need manual-work)
  - man
    - not(controlling-action can last-long)
    - not(motion need electricity)
    - not(able-to keep uniform efficiency)
    - can-change on-will-of control-agent
  - machine
    - control-agent created-by man
    - not(motion can-be-done-by man)
    - control-agent need electricity
    - not(motion ctrl.-by obj.-in-motion)
    - control-agent can-do only-one-job
  - obj.-in-motion
    - motion deliberately done by obj.
    - motion can-change on-will-of-obj.

Fig 7.16 Outline of Sandra's program

The two programs suggest that they feel comfortable about rules and questions framed at the specific epistemic level, and that these figure prominently in their rational self-elicitation of their knowledge about motion.

7.6.2 'General' epistemic frame

For rules belonging to the category *general hypothesis*, the epistemic question is different. The focus is detached from any particular type of motion event and the rules try to look for relationships that are invariant for different kinds of motions. For example, the rule

thing use up force if
thing is moving

is proposing a relation that holds true for all motions - an invariance about motion. This implies that the author of the rule is asking a question at a qualitatively different epistemic level from the ones quoted in the previous section. Regularities and predictability is sought as invariances. The focus is no longer upon the specificities of the situation but one of abstracting generalities that hold across specific relations. It can be interpreted as a generalization where the specifics under consideration are the generalities (relations) for the 'specific epistemic frame'.

In the search for generalities applicable to all types of motion, invariances cannot be sought from concrete physical attributes like speed or trajectory shape. A system describing motion that contains concrete attributes only is essentially a kinematic system and there is simply no single kinematic system that is applicable to all motions. To look for invariances across all motions thus necessitates the invention of abstract entities and the attribution of roles and functions to such entities. The concern expressed by programs at this level is still very much that of examining the causal link and predicting possible developments. Many of the general hypotheses written by the students were hypothesizing the existence of a mediating entity, often referred to as force or energy, as the ultimate underlying cause. They elaborated on the nature and attributes of such an entity: that it is an essential ingredient for motion; that it can be passed from one object to another; that the passage of this quantity has to be from the agent to the affected in a causal relation; that the more of this entity one has, the more the quantity of motion (the parameter used for measuring this may vary); and that it would be used up in motions. The essence of the invariance lies in this invented entity.

However, the ten programs that contain general hypotheses (fig.7.17 to 7.26) do not reflect only one level of epistemic framework. The differences lie mainly in the structure and inter-relationships within the knowledge. In fig. 7.17 to 7.24 are systemic networks showing three kinds of rules from each of the student's programs:

(i) rules that are general hypotheses,

(ii) rules that are in themselves not general hypotheses but the relations they concern do possess some rules in the general hypotheses category, and

(iii) empirical observations or specific conjectures that have goal statements coinciding with the conditions used in the general hypotheses.

The figures have been laid out so that all the general hypothesis are on the left hand side of the page and empirical observations or specific conjectures are on the right hand side. An examination of fig. 7.17 to 7.24 reveals striking similarities in the structure of these networks. First of all, they are all structurally 'flat' - very few of their rules are hierarchically related to each other. This means that the general hypotheses, or invariances, stand alone from each other as independent generalizations, but that the system includes no explanation for why they are as they are, or whether and how they relate to one other. Furthermore, just these general hypotheses, unsubstantiated by hierarchically related empirical observations and specific conjectures, would not be very helpful in reasoning about concrete daily events when specifics of the situation would have to be taken into account. Another feature of these networks is that where more than one rule exists for a particular relation, these rules do not all pertain to the same level of explanation. An example of this can be found in Schan's program (fig. 7.21). He had three rules on the relation 'motion caused by', two of which are general hypotheses and one which is an empirical observation. Two of these three rules, one from each category, are quoted below.

```
thing motion caused by own force if
      not thing is dead object and
      moving speed changeable by thing and
      amount of force changeable by thing.

thing motion caused by magnet if
      thing stick to magnet and
      not thing can separate itself from magnet and
      thing speed uniform.
```

Besides the fact that these two rules reflect the employment of two different levels of generality, there is another significant difference between them. In the first rule the causal agent is an abstract entity, 'own force', while in the second the agent is a magnet, a concrete physical object. This suggests an oscillation between the two epistemic frames, specific and general. These students were not yet able to exhibit a clear preference for either levels of analysis. It is important to note that such a characteristic can be found in the work of both groups of students and seems to be independent of the scientific accuracy of the rules. For example, there were only two general hypotheses in Wong's program, and both of them were about the condition for an aspect of motion to change:

object will change its direction of motion if
there is external force applied on object,

object will change its speed if
there is external force applied on object.

Taken together, these two rules are just paraphrases for Newton's first law of motion. However, there were two other rules closely related to these (refer also to fig. 7.17):

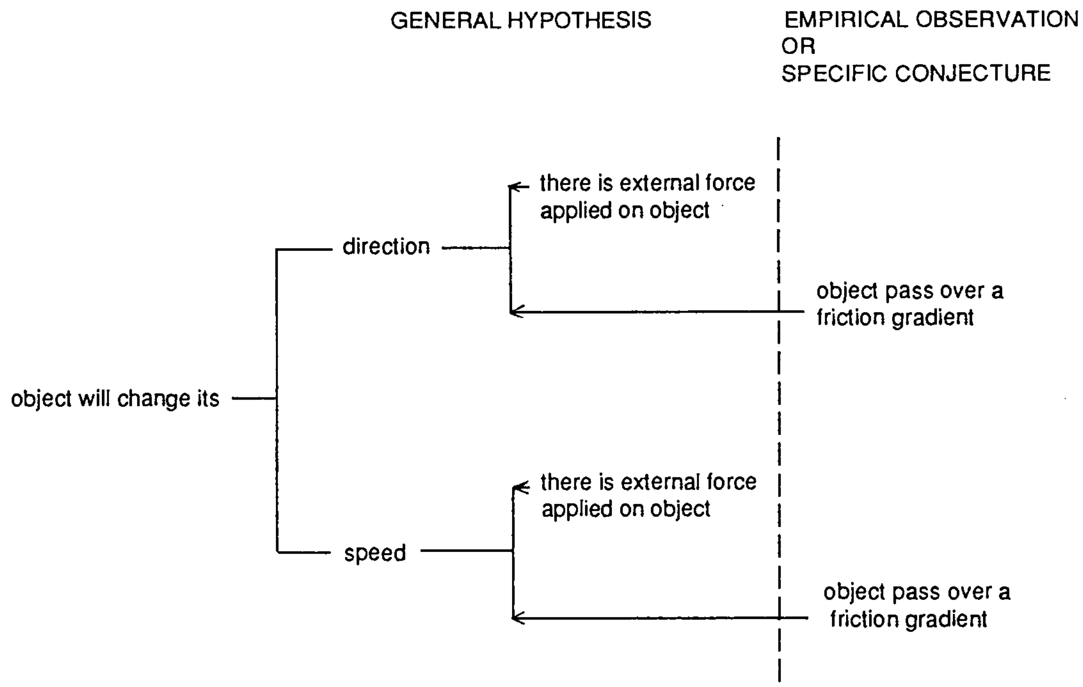
object will change its direction of motion if
object pass over a friction gradient,

object will change its speed if
object pass over a friction gradient.

These two sets of rules, though both pertaining to the same goal statements, operate on two entirely different planes, one rather specific and the other extremely general, and the knowledge base did not offer any clue as to how to traverse between these two planes.

However, the preceding discussion and the reference to 'oscillation' are not taken to imply that the students were conscious of such differences among the rules and trying to make choices between levels. The claim here is rather that when trying to explicate the same relation, there was an unconscious tension between writing about invariances which may be too general to be really helpful and relations which are very specific and thus limited in scope. The essence of the tension is between satisfying the pragmatic needs of everyday dealings with motion and the subconscious aspiration to provide a more general view of motion at a higher epistemic level.

The two structural characteristics of this group of programs belonging to the 'general' epistemic frame may be inter-related. The failure of the knowledge structures to form hierarchical structures relating general hypotheses to specific conjectures and empirical observations (i.e. the 'flat' structure) diminishes the utility of the invariances constructed. This may be an important contributing factor leading to the oscillation between these two epistemic levels.



* From this program onwards, empirical observations or specific conjectures are only represented if they are linked to a general hypothesis.

Fig 7.17 Outline of Wong's program

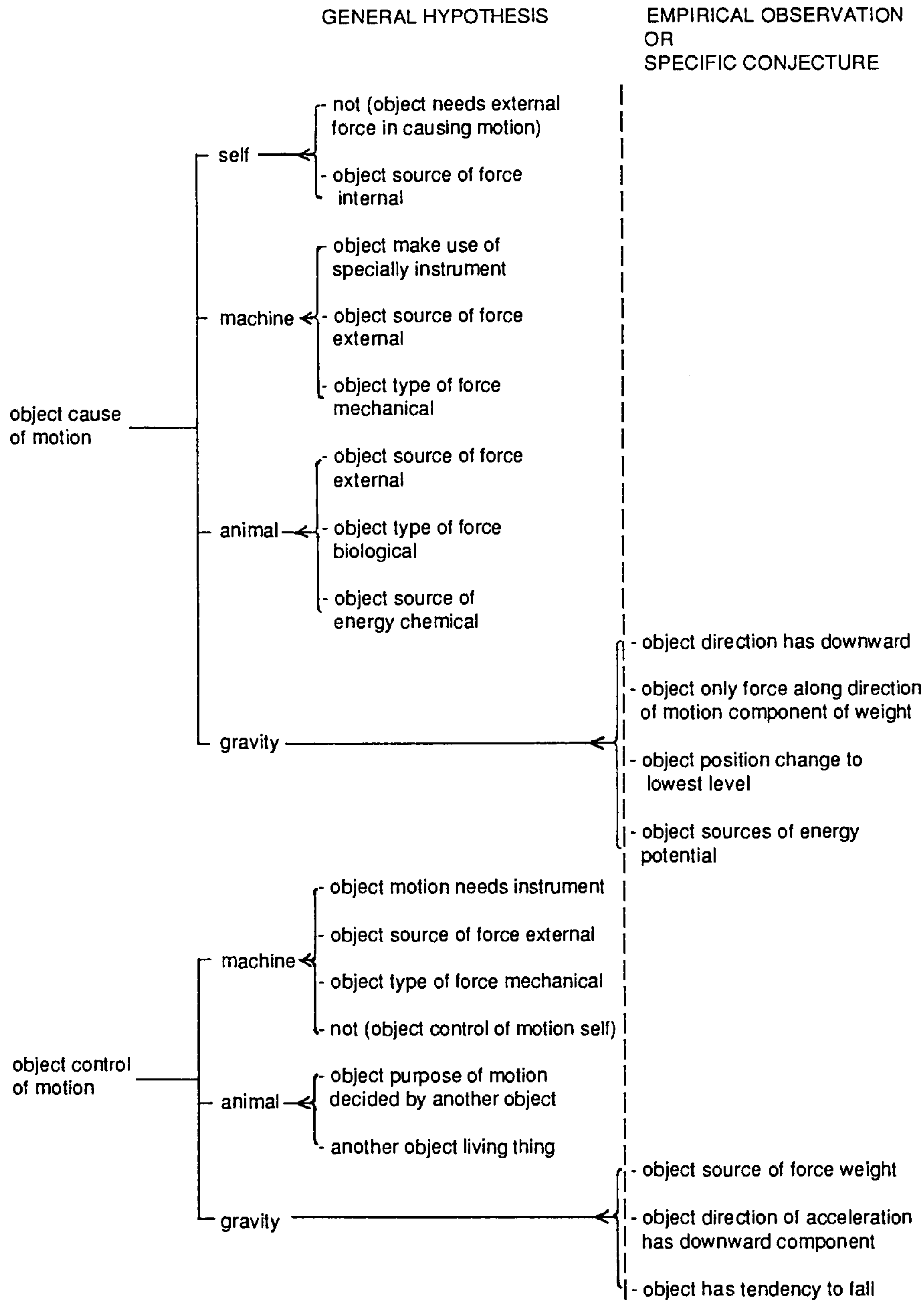


Fig 7.18 Outline of Lin's program

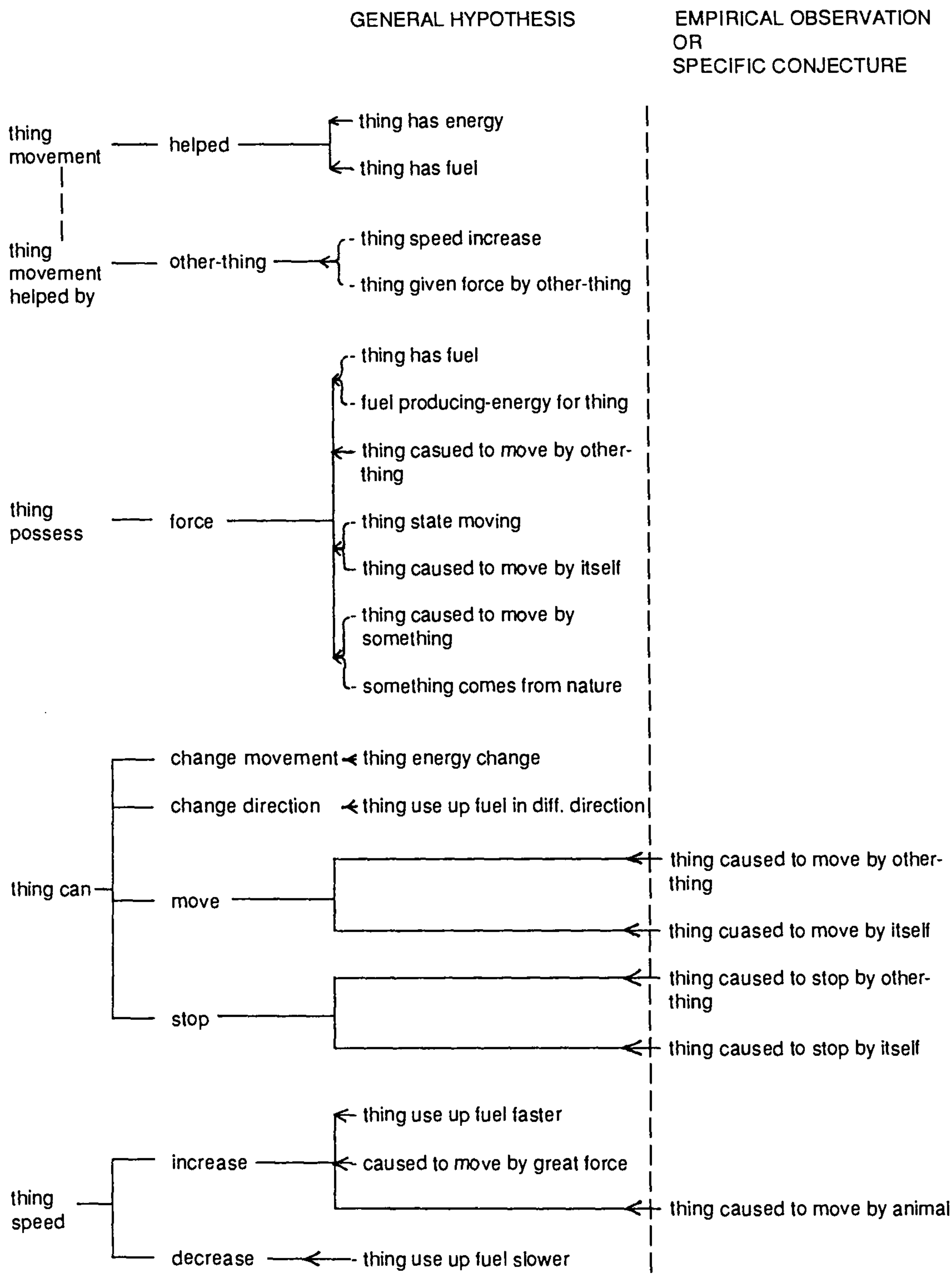


Fig 7.19 Outline of Ychan's program

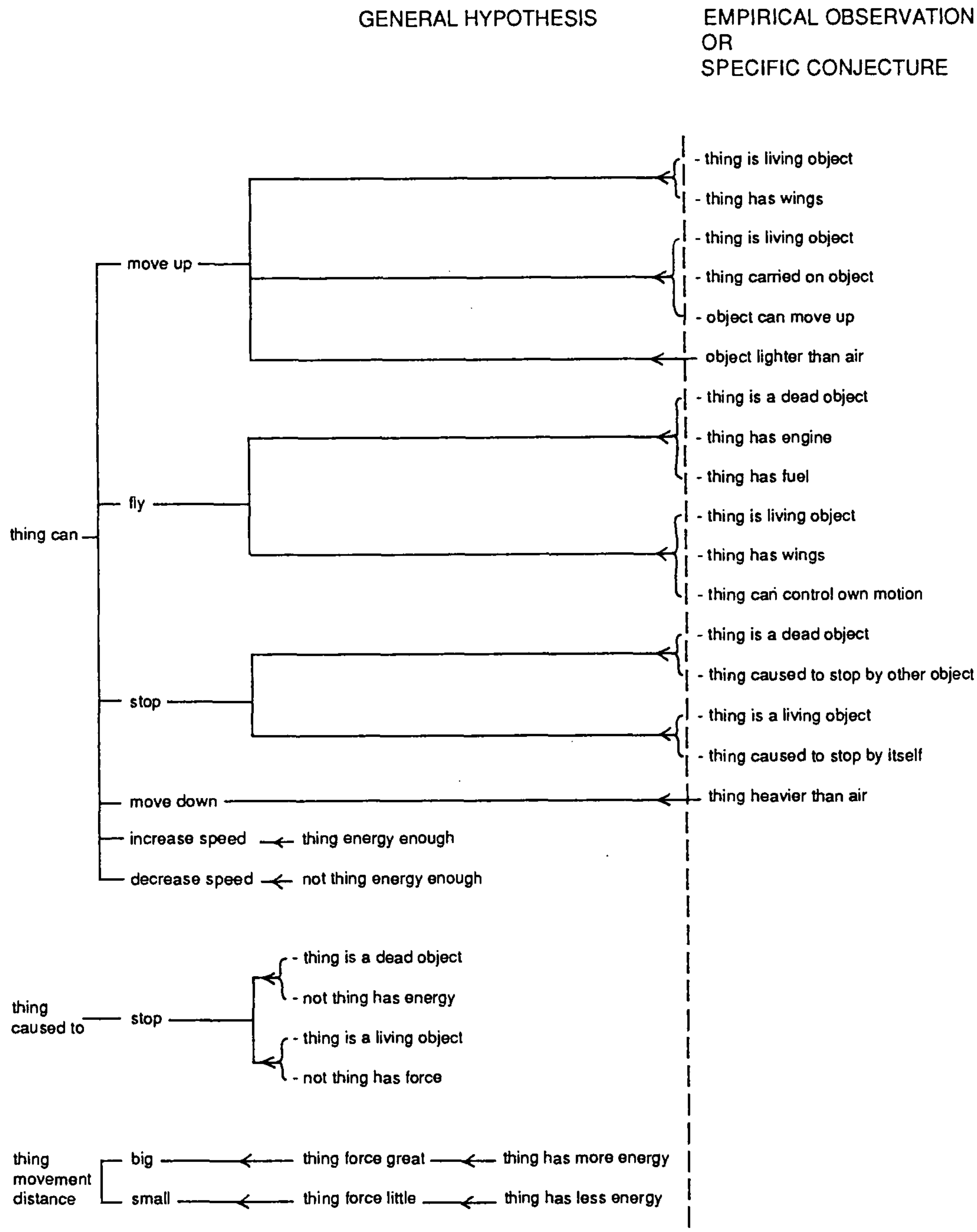


Fig 7.20 Outline of Leung's program

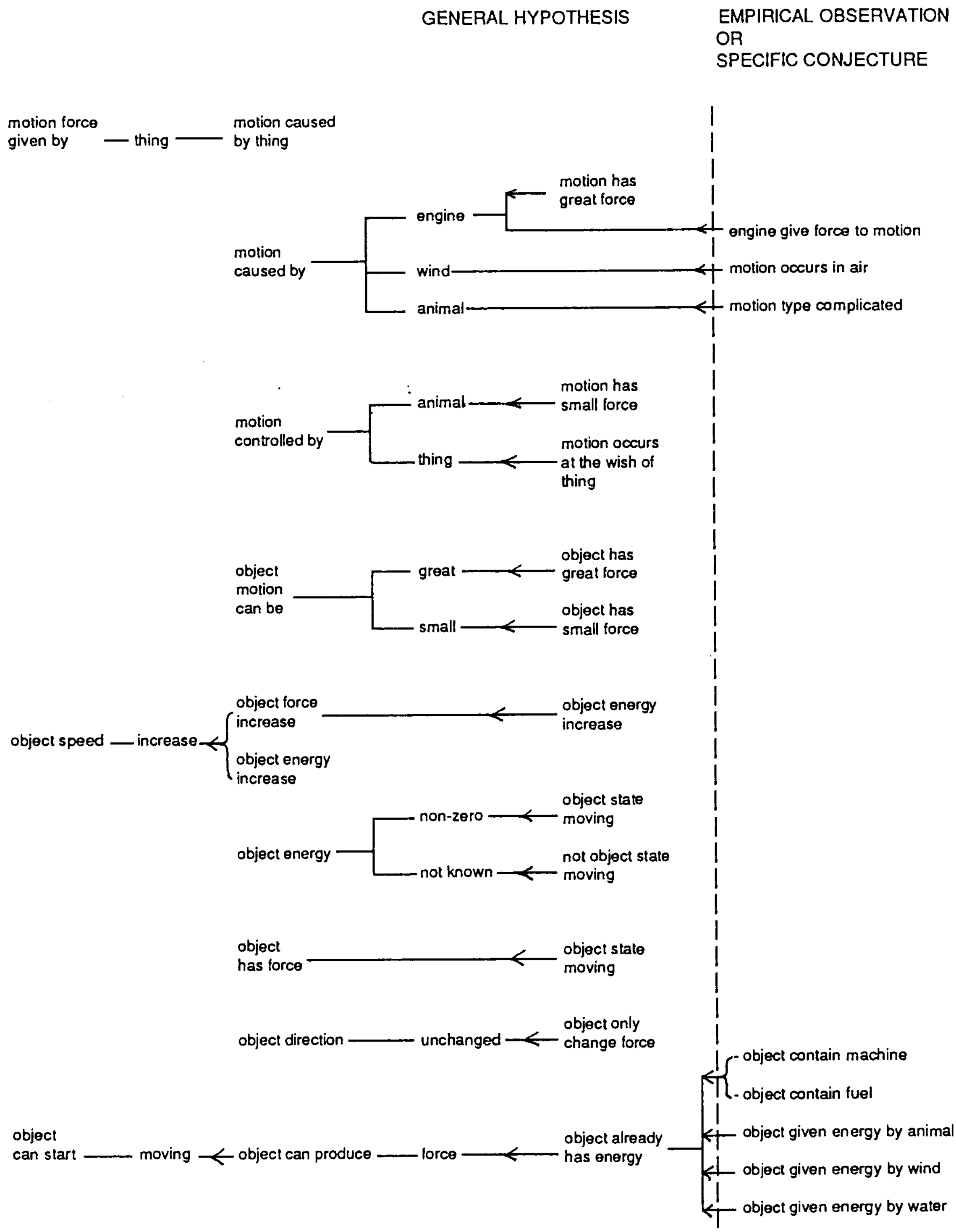


Fig 7.21 Outline of Schan's program

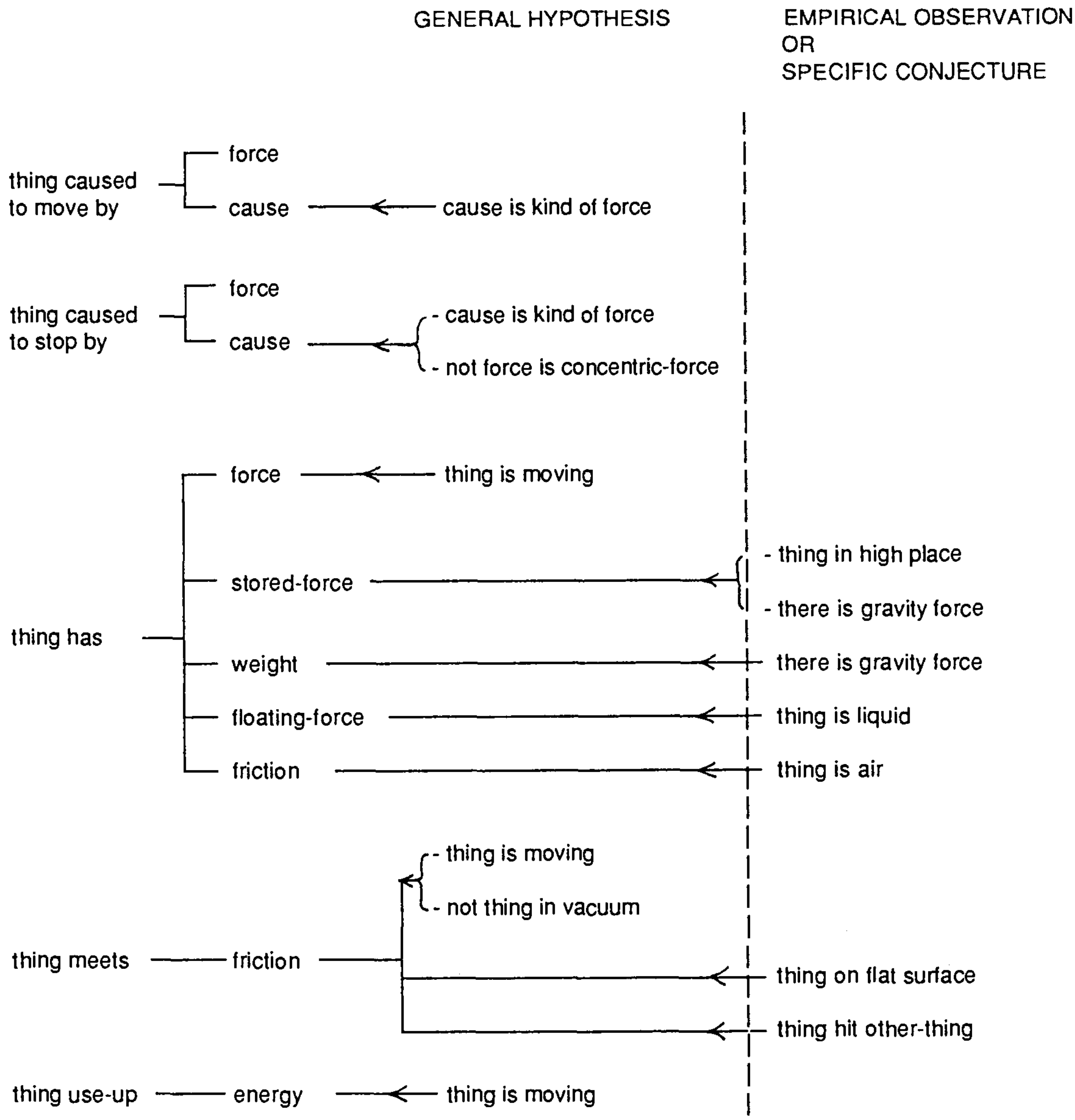


Fig 7.22 Outline of Cheng's program

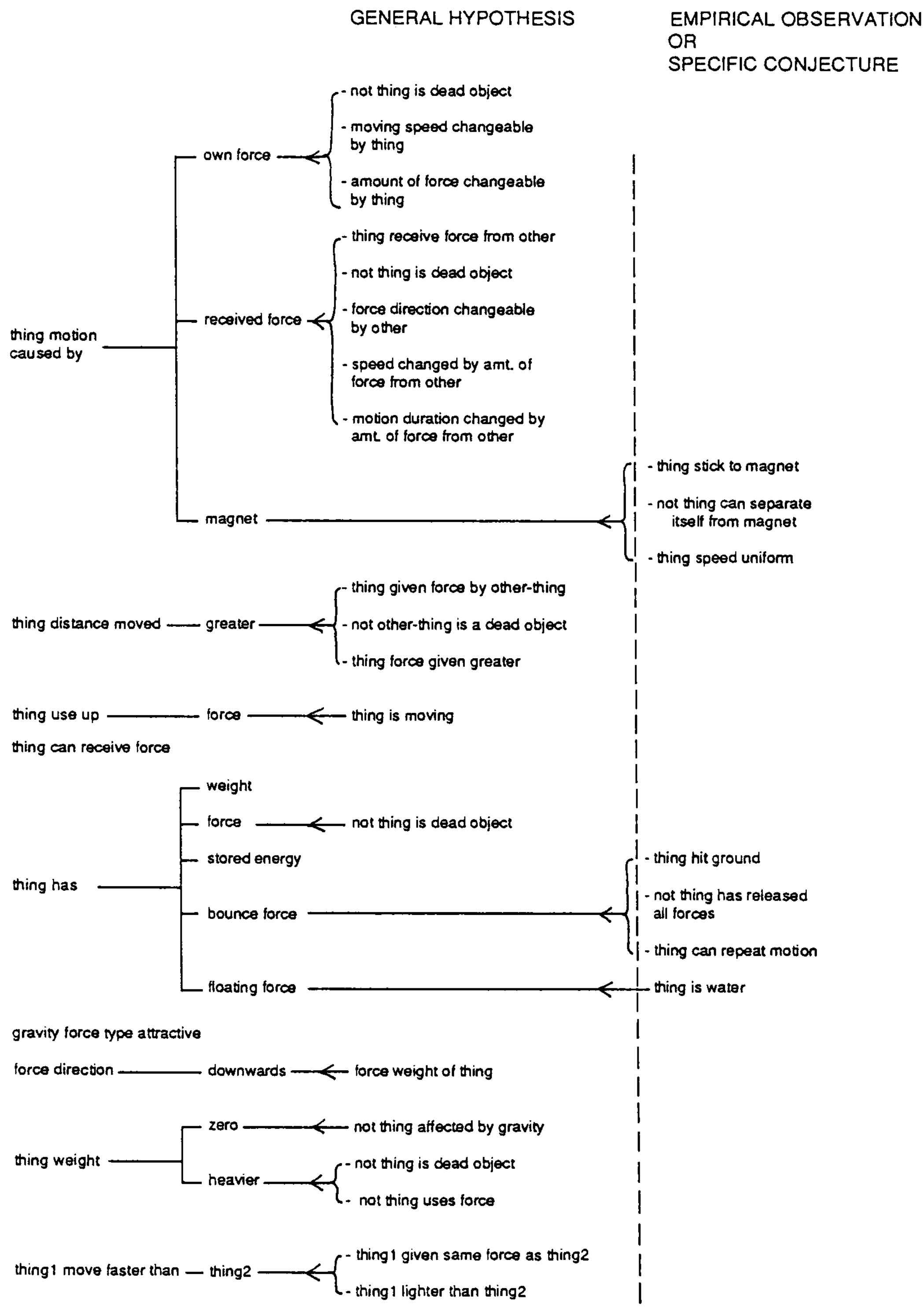


Fig 7.23 Outline of Wat's program

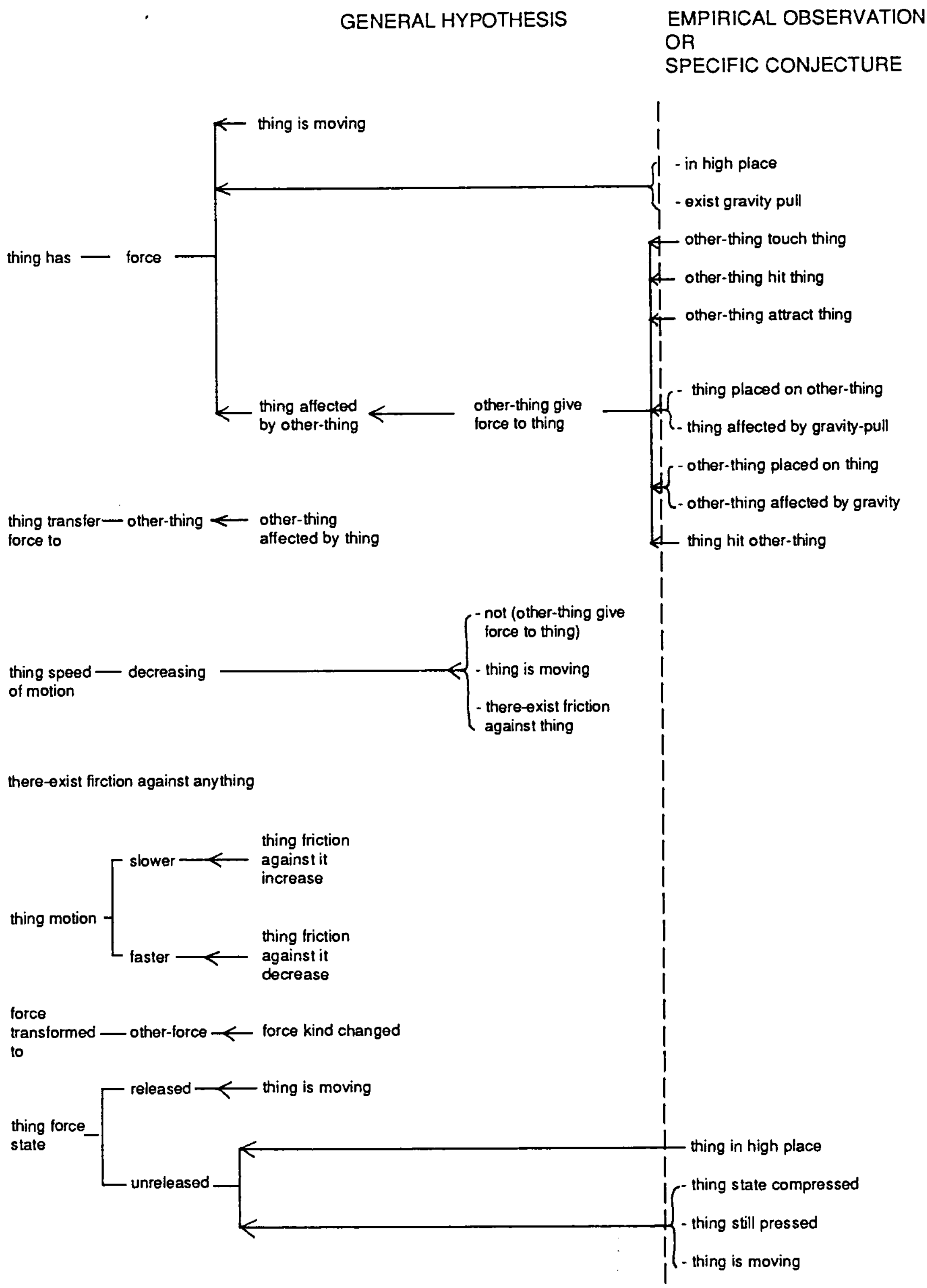


Fig 7.24 Outline of Cheung's program

7.6.3 'Structural' epistemic frame

Two of the students' programs had a notably different structure. They are the work of Fong (fig.7.25) and Evon (fig. 7.26). In these, the number of rules in the category general hypotheses was not necessarily large (only 5 in Evon's program), but they were all hierarchically related. The appearance of such a hierarchical structure implies a further level of abstraction above that required by operating within the 'general' epistemic frame. It represents an abstraction from the invariances. The epistemic question that Fong and Evon were answering was effectively this: "Is there a unifying framework that can explain the whole range of general hypotheses that they held valid?" In the context of the present programming task, such a framework should be able to relate the different facets of the hypothesized mediating entity(ies), in most cases force and/or energy, as depicted in the general hypotheses, and by so doing to highlight their utility for the specific conjectures. Thus each of these structured knowledge bases was built around a prominent theme, a theme which underlies, relates and explains the range of general hypotheses present.

Here, membership to the 'structural' epistemic frame is independent of the scientific validity of the contents of the rules. The theme of Fong's program was that agency has to be mediated through interactions between the agent and the affected, different forces being involved in different interactions, and the type of interaction for individual cases depending on the specificities of the particular situation. This is of course derived from and compatible with a Newtonian view of motion. However, Evon's program was built around a different theme, one which is a typical case of an alternate framework, a deviant from Newtonian mechanics. The theme of her program was built around the distinction between two different 'sources of force', internal and external. Internal forces are responsible for 'self-powered' motion, that is, the autonomous motions. All other motions (i.e. the non-autonomous ones) would be affected by external forces. She further linked this theoretical framework to the solution of the pragmatic concern to predict trajectories: only those objects powered by external forces would have predictable trajectories.

Another feature of these two programs that can be interpreted as a mark of a well-structured knowledge base is that at each level of the hierarchy, the entities sharing the same syntactic role in a relation are ontologically similar, and the nature of the conditions used for the resolution of goals of the same level were also similar. For example, in Evon's program (fig. 7.26), the entities that can affect motion are all forces: weight, tension, air current, friction, etc. However, in Brenda's program (fig. 7.15), motion can be controlled by forces (like tension, friction, gravity), or material properties (like elasticity), or energy (like heat), or physical agents (like man, machine, or the object itself). (For a more detailed analy-sis of this aspect of the programs, please refer to Law (1988b) in app. 3.) It is important to point out here that this is not just a programming feat. It is the consistent

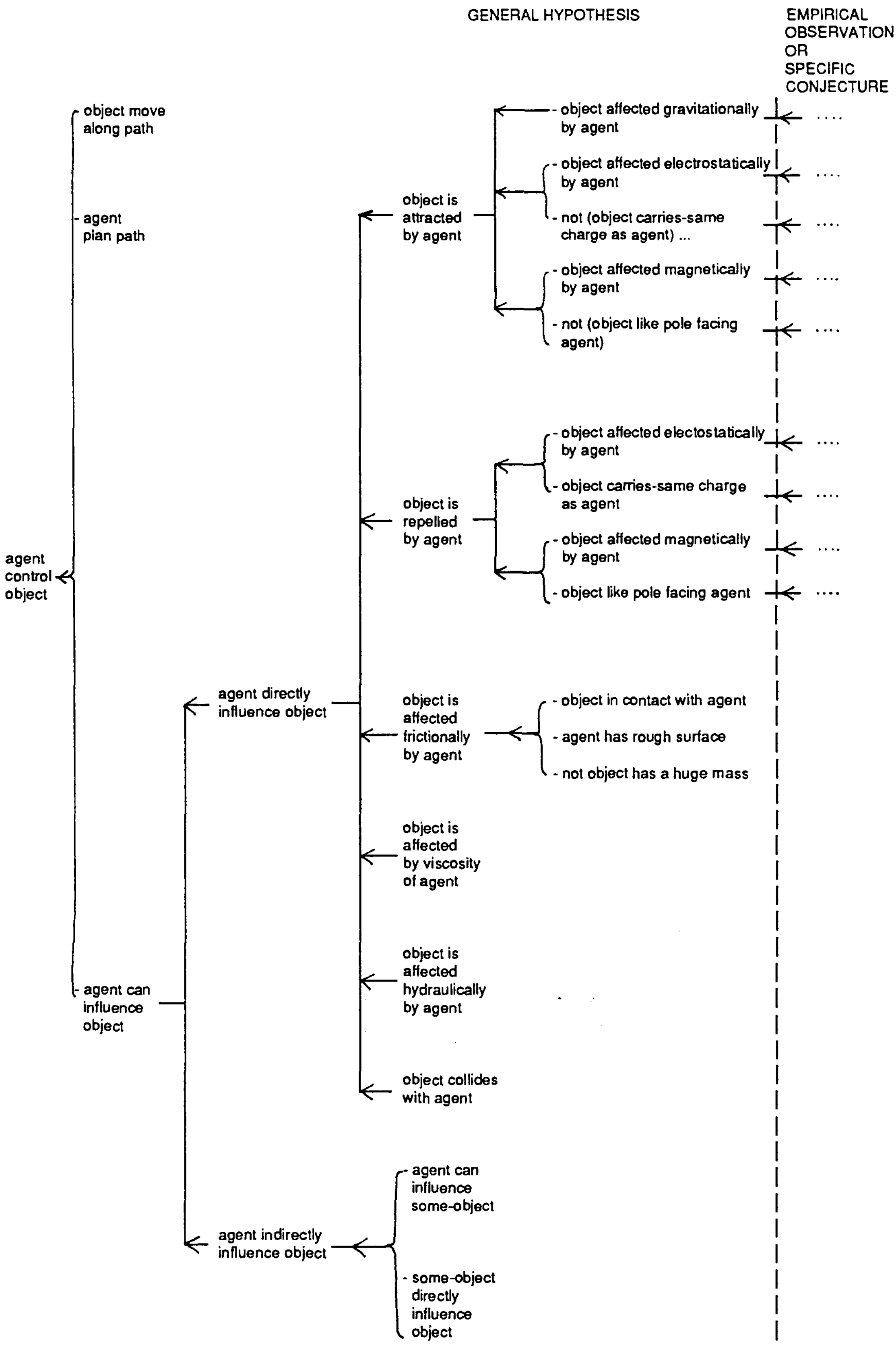


Fig 7.25 Outline of Fong's program

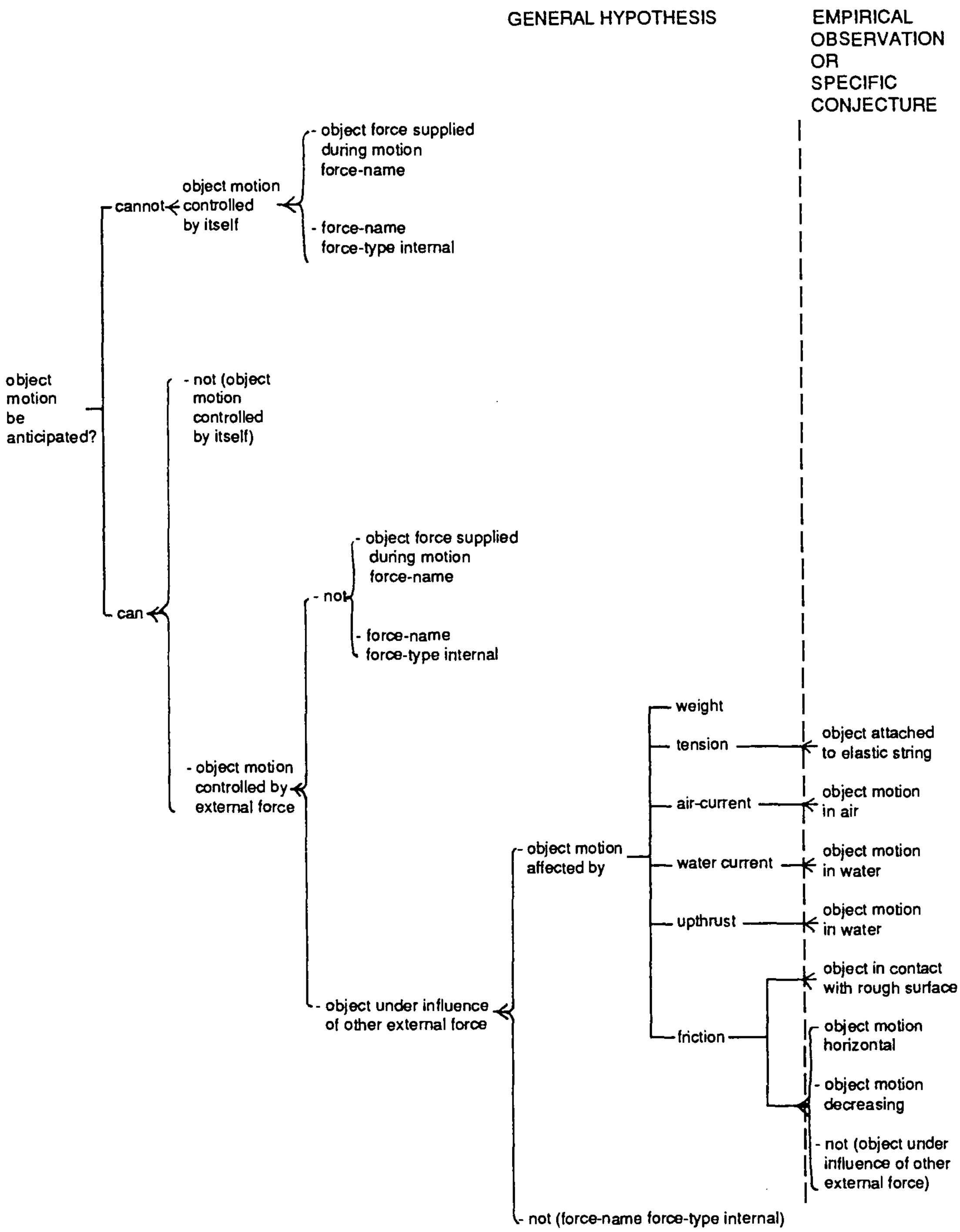


Fig 7.26 Outline of Evon's program

application of a central theme, a central hypothesis, to all considerations of motion that produces such ontological consistencies.

A related structural feature of these two programs, one which is hinted at above, is that many of the end nodes in the hierarchy of general hypotheses were linked to specific conjectures so that the whole set of rules together form an integrated system. In these two knowledge bases, the specific conjectures play a very important role - that of providing explicit connections between the concrete physical parameters for specific situations with the main theoretical theme. These structural links not only give life to the specific conjectures by enhancing their utility, but also provide a means for traversing between the general and the specific epistemic planes. These links thus remove the tension between the (aesthetic?) quest for generalities and the pragmatic need to cope with specific situations.

The programs written by Schan (fig. 7.21) and by Cheung (fig. 7.24) deserve special mention with reference to the present discussion. Both programs contain small fragments that are hierarchically related, and provide links from general hypotheses to specific conjectures. These two features are insufficient to claim that they were answering questions at the 'structural' epistemic level, as the general hypotheses touch on diverse themes and they lack a consistent central theme to relate to each other and from which to organize the links to the concrete specific situations of daily reasoning. However, these two programs may indicate the emergence of a progression towards the structural epistemic frame.

7.6.4 Concurrence with the Piagetian mechanisms for equilibration

Given the limited number of cases that the present research dealt with, it is remarkable that the three epistemic levels identified in the students' programs seem to concur with the mechanisms for equilibration proposed by the Piagetian school (Piaget, 1975). As was described in chapter 3, Piaget hypothesized that developmental changes in mental schemas take place via three "knowing cycles": within-scheme, between-scheme and totality cycles. Garcia and Piaget's work on the psychogenesis and sociogenesis of knowledge (Garcia, 1987, 1983) found that similar mechanisms are also at work in the development of theories in general. Garcia's (1987) paper referred to these mechanisms as intra, inter and trans mechanisms and the following extract gives a clear outline of their theory.

> *"....we found in several fields that the construction of a theory goes, in general, through three different moments: a moment when a number of isolated facts are known, identified and analyzed independently of each other; another moment when these facts are found to be connected by transformations that leave something invariant; and a third moment when you have a structure that explains the transformations and the invariance, and each particular case. We have called these three steps 'intra', 'inter' and 'trans'"*
>
> *(Garcia, 1987, p. 139)*

There seems to be a clear correspondence between the three Piagetian equilibration mechanisms and the epistemic frames identified here. The correspondence goes between specific and intra, general and inter, and finally structural and trans. It is pleasing that so general a point of view is useful in analyzing such data. It may give one some confidence that the method has elicited a real and in some sense profound level of students' thought and that the three epistemic levels identified correspond to a progression that necessarily accompanies different phases of development in a progression towards the development of a scientific understanding (model).

Chapter 8. TESTING AND MODIFICATION OF PROGRAMS

8.1 Program Testing and Modification as a Further Process of Knowledge Elicitation

One important advantage of using expert system development as a means of knowledge elicitation in the present research is that a subject can explore the logical consequences of expressed knowledge by issuing queries to the computer and as a result subsequently be led to modify it if necessary. For example, a student held the following interactions with the computer after putting in a few rules about the kinds of forces an object may experience:

Student : apple experience _what? (student asks the computer for possible values for the variable "what")

Computer : is it true that apple is moving in air?
(computer queries the user for information to determine the value of a condition in one of the rules)

Student : yes.

Computer : apple experience air resistance.

Student : why?

Computer apple experience air resistance because apple is moving in air.

Student : any more answers?

Computer : is it true that apple moving downwards?
(computer attempting to match another rule)

Student : yes.

Computer : apple experience gravity.
no more answers. (no more rules unresolved)

Student : tape-recorder experience gravity? (student started a new query)

Computer : is it true that tape-recorder moving downwards?

Student : no.

Computer : no, tape-recorder does not experience gravity.

Student : why?

Computer : tape-recorder does not experience gravity because it is not moving downwards.

At this point, the student was dissatisfied because she had learnt that an object placed on a table would still experience gravity. She thought about it and added in a new rule to take the case of the tape-recorder into account:

Student : _thing experience gravity if _thing has weight and _thing stay still. (student typed in a new rule)
Student : tape-recorder experience gravity? (previous query re-issued)
Computer : is it true that tape-recorder has weight?
Student : yes.
Computer : is it true that tape-recorder stay still.
Student : yes.
Computer : yes, tape-recorder experience gravity.

The researcher encouraged the student to make a new query.

Student : kite experience gravity?
Computer : is it true that kite is moving downwards?
Student : no.
Computer : is it true that kite has weight?
Student : yes.
Computer : is it true that kite stay still?
Student : no.
Computer : no, kite does not experience gravity.

The answer from the computer was contrary to her expectation and she wanted to modify one of the rules.

Student : edit second rule on "_thing experience gravity"
(she then eliminated the second condition on that rule)

The edited rule then reads:

_thing experience gravity if
_thing has weight.

The researcher then took this chance to ask her why she put in the second condition in the first place. She explained that she wanted to distinguish between things that move from those that do not, and to specify that the latter also experiences gravity.

It can be seen from the above example that the process of program testing offers a valuable chance for the researcher to probe deeper into the students' understanding even after the initial elicitation in the form of program development. It was found that the students were mostly thinking of very specific situations or 'prototypes' when they were writing down the PROLOG rules, so that though the rules took the form of generally applicable statements, it was usually not the students' conscious aspiration to do so and the rules normally fail to live up to such an expectation. Thus when a student finished writing some rules on a certain theme, his/her first test of the program usually used the prototypical situations s/he had in mind during the writing process.

One serious problem in knowledge elicitation is the possible mismatch between the intended meaning of the subject and the interpretation of the researcher. The explorability of the expressed knowledge as a PROLOG program at least offers the researcher a chance to find out whether such a mismatch exists, thus providing a qualitative validity check on the results. This is especially important as the programming task is new to all the students.

While the program testing process provides deeper insight into the ideas and concepts expressed by the program lines, the program modification behaviour of a student provides very valuable information on a variety of his/her cognitive characteristics that are very difficult to explore otherwise. When would a student find it necessary to modify a rule (a PROLOG statement)? Are there systematic individual differences in the criteria for modification? What are the main types of conflict/ dissatisfaction? How do the students face these conflicts? What kinds of program modification actually took place? Are there qualitatively different modification behaviours, and if so, do they bear any relationship to the qualitative differences in program structure discussed in the previous chapter? Answers to the above questions shed light on the kinds of possible cognitive conflict one may expect when one's beliefs are faced with challenges from daily phenomena, the cognitive strategies that may be employed to cope with such conflicts and the possible cognitive outcomes for different conflict resolution strategies.

8.2 Criterion for Satisfaction - Pragmatism v.s. Intellectual Aestheticism

As mentioned earlier, the PROLOG programming environment forced the students to put their ideas into the *format* of generally applicable laws, no matter whether the student thought along such lines or not. When the programs were subsequently tested with different situations, the only means available to the computer were the explicit rules written by students, and the computer will treat all rules literally as true generalizations. As was evidenced by the behaviour of the students, this was not the general mode they operated in. The testing and modification behaviour of the students thus highlights the mismatch between the functioning of a commonsense system and a formalized logical representation of it.

A variety of qualitatively different standards for satisfactory performance were set by the students for their programs. Many of the students had their main attention focussed on the acceptability of the answers given by the computer and would not bother to find out the reasoning used unless the answer turned out to be unacceptable. Even when the researcher purposely exposed the idiosyncrasy of the reasoning process used in the resolution of a particular query, most were very reluctant to make any changes as long as the answer was deemed correct. Furthermore, not all students demanded that the programs be able to deal with all possible cases correctly. There seemed to be a region of tolerance for the performance of the programs, and they were only worried if the output answer was outrageously wrong. For example, one of the students, Lin, wrote the following rule to decide whether a non-autonomous moving object is animate or not:

object nature living if
object use specially designed instrument and
object cause of motion machine and
object control of motion self and
object purpose of motion self-decided.

He then tested the rule with the case of a driver and the computer duly answers that '*it is true that driver nature living*'. The researcher then asked him to do a second test with the case of a passenger, thinking that in this case he would have to answer that the control of motion would not be the passenger. However, he obtained the correct answer '*passenger nature living*' from the computer by saying that the control of motion belongs to the passenger because he instructs the driver on where to go. The researcher then tried to put him into an even more awkward position by asking him to test the program with the case of a baby passenger. However, Lin was not moved at all and calmly responded that there are surely exceptions to every rule, and it is very difficult to make one that works for all cases.

The dissatisfaction with a rule is normally greatest when the answer appears to be far from commonsense expectations. For example, Brenda had been working in a way very similar to Lin, when she was absolutely shocked and dismayed by one of the responses from the computer. She wrote a rule on control of motion:

object motion controlled by engine if
object has the ability to move continuously without getting tired

and had been testing it with objects like car and fish. Then she looked out of the window and saw a falling leaf. She used that as a test case:

Brenda : leaf motion controlled by what?

Computer : is it true that leaf falling freely downwards?

Brenda : yes.

Computer : leaf motion controlled by gravity.
is it true that leaf fluctuating in air?

Brenda : yes.

Computer : leaf motion controlled by air current.
is it true that leaf gets tired after moving?

Brenda : no.

Computer : is it true that leaf has ability to move continuously without getting tired?

Brenda : yes.

Computer : leaf motion controlled by engine

The thought that the leaf's motion was controlled by an engine was so absolutely appalling that she was totally disappointed and requested to be allowed to go home to review the whole program all over.

However, the pragmatic standard was not employed by all students alike. A few of the students were not happy if the reasoning given for the resolution of a particular answer was not correct and they wanted to modify the programs. Furthermore, these students tended to go beyond just writing down some empirical correlations between the possible actual causal/control agents of motion and some describable features of the motion, but wanted to express higher principles - to build up some 'theories', i.e. systematic reconstructions, about motion, even though they could become very frustrated by serious obstacles while attempting to do so.

Of all the subjects participating in the program development task, Evon met the greatest frustration. She was one of those students not satisfied with just meeting the pragmatic standard. She wanted to express a 'higher principle' which she strongly believed in: all motions are either autonomous or non-autonomous, and the determining criterion for such is the supply of force. If the supply of force is from the moving object itself, then it is autonomous motion, and if the supply of force is external, then it would be non-autonomous. However, free fall is a very special case. According to her own conceptualization, a motion which is neither autonomous nor caused by external agents must be caused by gravity (or weight, as she could not quite distinguish the two and both seemed acceptable). At first she made rules so that any motion not affected by external forces would be caused by gravity. However, motion of a stone in free fall will not be caused by gravity if weight is classified as an external force. On the other hand, if weight is to be classified as an internal force, then free fall would have to be classified as autonomous motion - a very unacceptable result (Law N. 1988). Despite the inability to resolve this dilemma, she did not, at any moment during the eight weekly sessions on the programming task, give up working on rules relating to this abstract notion of supply of force being internal or external and fall back on writing isolated concrete empirical relations.

Thus we can see in the programming behaviour of the students two different approaches to describing the world of motion. One group, which may be called the 'pragmatic' group, describes the world in terms of prototypes, broadly categorizing motions according to prominent stereotypes and then trying to correlate the features of each stereotype. When the rules failed, their immediate response tended to be that the categorization used was not fine enough and needed to be tuned further to take into account motions further away from the stereotypes. These students did not encounter serious challenges or frustrations even when their programs failed. They never anticipated their categorization schemes to be infallible. The other group, less numerous in number, may be called the 'aesthetic' group. They tried to work out, through the labyrinth of different kinds of motions and characteristics, a systematicity over and above just a flat categorization. They looked for some general principles, laws, that would govern all motions. Inherent in such an approach is the belief that there are some hidden principles, some grand laws of nature, that govern the world of motion and give a systematicity to the myriad manifestations of different motion characteristics. For a student adopting an aesthetic approach to the programming task, the different rules s/he wrote down would then have different level of importance, the most central ones being the core hypotheses which give rise to systematicity. During the course of program testing, these students have met much greater frustration than that faced by the pragmatic students, if it is a core hypothesis that the failure challenges. Frustration originates from a basic belief of what the world is like - the describability and generalizability of the world around us.

For the purpose of dealing with everyday motion events, the pragmatic standard is very often sufficient and may be more efficient than comprehensive, systematic theories. (One empirical source of support for this is the rating of the various expert systems given by the students, to be described later in this chapter.) What prompts a person to take up an aesthetic approach is thus not pragmatic needs but a personal commitment to intellectual passions (Polanyi, 1958). This pragmatic/aesthetic distinction is also one of the fundamental distinctions between science and commonsense (Nagel, 1961).

8.3 Program Modification and Cognitive Strategies

There are two forms of query a user may put to the expert system developed:

(i) confirm the validity of a statement, e.g., "Is it true that car's motion caused by engine?"

(ii) give instantiations of a variable that would make a statement valid, e.g., "What causes car's motion?"

In cases where the computer failed to provide a satisfactory answer, there were generally three possibilities:

- the test case was wrongly included as an instance of the goal statement;
- the test case was wrongly excluded from being a successful instance of the goal statement;
- the computer failed to give the correct answer because no rules pertaining to the appropriate goal statement had been entered. For example, a student may think that the appropriate answer for question (ii) above should be that the car's motion is caused by engine, but no rule of the form '*object motion caused by engine if* ....' had been entered.

In analyzing the modification behaviour of the students, modifications due to typing mistakes and syntactic errors were not taken into account as these only reflect on the technical features of the programming environment and the programming ability of the students, but not on the actual ideas put forward by them. Twelve different types of program modification could be identified from the modification behaviour of the students (table 8.1a), which relates to five distinctly different cognitive strategies (table 8.1b) being employed when facing unsatisfactory outcomes from the computer.

8.3.1 Maintenance strategies

This is by far the most popular strategy, accounting for nearly 60% of all modifications undertaken by the students. The main characteristic of these strategies is that the modifications try to introduce minimum change while maintaining the basic knowledge structure. There are five different types of program modification belonging to this strategy category.

8.3.1.1 Addition of conditions

As was mentioned earlier, most of the students were anchoring onto stereotypical kinds of motion when they were drafting rules. Ideally, the rules written down should be necessary and sufficient conditions for the goal statements to be valid. However, many of the

Table 8.1a Distribution of types of program modification.

Subj.	Maintenance strategies					Evasion	Extension by categorization		conceptual reform.	Theory building	
	Reph-rase	add cond.	del cond.	repl cond.	comb. rules	delete rule	add case	add goal	reformulate rule	Syntactic restruct	Total
Sand	2	1	3		1		2	1			10
Bren	2	2	1	3			1	3			12
Leung	2	1		1			1	2			7
Schan	4			1	2		4	1			12
Ychan	1	4				2	6	3			16
Lin	3		1			2		2	1		9
Evon	3	1		1		2		2	2		11
Wat	3	1							3		7
Cheng		1							1		2
Ahlun	3	3					1				7
Wong	8	3									11
Fong	5	1				1		1		4	12
Total	36	18	5	6	3	7	15	15	7	4	116

Table 8.1b Distribution of modification strategies and epistemic level of the programs.

Subj.	**Modification Strategy** Maintenance	Evasion	Extension	Reformu-lation	Theory Building	**Epistemic Level**
Sand	7		3			Specific
Bren	8		4			Specific
Wong	11					General
Leung	4		3			General
Schan	7		5			General
Ahlun	6		1			General
Ychan	5	2	9			General
Lin	4	2	2	1		General
Cheng	1			1		General
Wat	4			3		General
Evon	5	2	2	2		Structural
Fong	6	1	1		4	Structural
Total	68	7	30	7	4	

students failed to appreciate the nature of the programming task as such and worked according to their own understanding of how best to achieve the task, and different outcomes and coping strategies developed according to whichever interpretation of the nature of the PROLOG rules they employed. Some of them, e.g. Bren and Ychan, tended to use the rules to relate the most prominent feature(s) of the stereotypical motion they had in mind. Understandably, these rules very often posed conditions which were too loose, leading to the first kind of failure described above - non-instances being wrongly identified as satisfying the goal statement. To cope with the situation, the easiest strategy would be to add in conditions that would exclude the particular kind of non-instance at stake. The following is an example of this type of program modification:

Ychan had written the following rule about autonomous motion:

```
thing motion under control if
      thing is living object and
      thing has brain.
```

The program survived his first test successfully when given the query:

"walking man motion under control?"

The researcher then suggested another query:

"falling man motion under control?"

As is expected, the computer responded that the falling man's motion is under control. However, Ychan believed strongly that by definition no falling object can have its own motion under control. He modified the rule on motion under control by adding in a further condition so that the rule then read:

```
thing motion under control if
      thing is living object and
      thing has brain and
      thing move on ground.
```

It is evident from the above example that the modification was not made after an overall review of the validity of the rule but made just to cope with the failed instance at hand. As this is a rather defensive coping strategy, it belongs to the group of maintenance strategies.

8.3.1.2 Deletion of conditions

Another common misinterpretation of the proper functioning of the PROLOG rules was to treat them as if these should act as detailed descriptions of the stereotypical cases (Sandra's and Lin's program's are typical examples of such a misinterpretation). As such, the rules are bound to contain conditions that are typical characteristics of valid instances for the goal statements, but which were not necessary for the goal statements to be satisfied. A student under such circumstances normally interpreted the failure as arising from conditions being too stringent and would thus modify the rules by removing some of the conditions so as to include the appropriate instances. The following is an example of this type of program modification.

Sandra had written the following rule on autonomous motion:

thing motion caused by itself if
thing motion controlled by itself and
thing energy supplied by itself and
thing motion can change on will of itself and
thing can stop any time by itself.

The computer had been giving satisfactory answers to several test cases for this rule. Then the computer was given the query,

"high jump motion caused by itself?"

The answer obtained was that high jump was not controlled by the object in motion because the object involved could not stop the motion any time it wanted. To cope with the situation, Sandra removed the last condition, '*thing can stop any time by itself*', from the above quoted rule.

Again, this kind of program modification was also a coping strategy, just to handle the failed case in hand.

The above two modification methods, addition and deletion of conditions, represent the typical failure situation and the nature of the strategies employed in all maintenance strategies. The other three methods, rephrasing of conditions, replacement of conditions and combining rules are just minor variations of these two versions.

8.3.1.3 Rephrasing of conditions

Sometimes, a student may want to rephrase a condition so as to clarify his ideas. However, if such a rephrasing is prompted by a problem test case, the rephrasing would very often result in making the condition (and subsequently the rule concerned) more or less stringent, as the case may be. For example, Sand wrote the following rule on control of motion by gravity:

object motion controlled by gravity if
object source of force weight and
object direction of acceleration downwards and
object has tendency to fall.

However, this rule would not be able to classify the motion of a parachutist as controlled by gravity as Sandra would want it to. She thus relaxed the second condition to '*object direction of acceleration has downward component*' to cope with the case of the parachutist.

8.3.1.4 Replacement of conditions

When the PROLOG rules were employed to relate a stereotypical motion to the most prominent/important feature of the situation, the particular feature chosen was normally one that made the strongest subjective impression on them, and may not be a necessary or sufficient condition for the goal statement. Sometimes when the test results evidently show that this was the case, they decided to give up that particular feature and replace it by a completely different one. For example, Brenda wrote several rules about control of motion and the following rule about motions controlled by engines was one of them:

> *object motion controlled by engine if*
> *object has ability to move continuously without stopping.*

Later on, she tested these rules with the case of a big balloon (she had mistaken that such big balloons need an engine for motion). However, when queried by the computer about whether balloons can move continuously without stopping, she replied 'no' as she believed that the balloons can only go up to a maximum height. After thinking the situation through, she changed the rule to

> *object motion controlled by engine if*
> *object has ability to move continuously without getting tired.*

At a later stage of her program development, she had further problems with this amended rule. She then replaced the condition a second time:

> *object motion controlled by engine if*
> *object moving only after addition of fuel.*

This oscillation of the chosen condition from a description of the duration of motion to an animistic one of tiredness to a more physical one of energy source cannot be described as a reformulation because there is no fundamental belief of the student that is being challenged or altered in this process. This is just a trial and error process of trying to locate a more suitable condition and should thus be classified as a maintenance strategy only.

8.3.1.5 Combination of rules

A combination of rules occurred normally when two rules developed at different times with somewhat different wording were later found to be essentially referring to the same situation.

8.3.1.6 Maintenance as episodic accommodation for specific situations

As can be seen from the above description of the five types of modification belonging to this strategy category, the students normally started off with rules anchoring closely onto specific stereotypical motions, and when the rules fail, the modifications were developed so as to cope with the specific failures, and not as a consequence of inspecting the rule for general failures. Thus the knowledge structures as represented by the PROLOG rules were modified, (or, using Piaget's terminology, accommodated), to adjust for the local test situation, without bringing about any substantial modification to the knowledge structure. Furthermore, judging from the nature and context of these adjustments, it is not evident that the students have strong commitments to the modified rule and it is reasonable to expect similar adjustments (or even modifications in the reverse direction) to take place when further failure cases are met. These modifications are probably episodic accommodations - adjustment of the same knowledge schema to the constantly changing contents to which it is applied, rather than permanent changes to the schemas concerned.

8.3.2 Evasion Strategy

Of all the rules that the students wrote, only seven rules had been deleted without replacement. One of these was deleted by Fong because he thought it was too trivial after he had tested them. Thus this deletion was made because the rule could not satisfy the author's own quality standards, though it was nevertheless correct. The other six deletions, two each by Ychan, Lin and Evon, were made because they were found to be wrong and no viable alternatives could be found. Normally, when a problem arose, the tactic employed was to try to patch it up - through one of the maintenance strategies listed above. In these six cases no satisfactory maintenance strategy could be located and the authors had to resort to an evasive strategy - deleting them altogether.

8.3.3 Extension of Knowledge by Further Categorization

When the computer failed to provide a satisfactory answer to a query, it may not necessarily be that any one particular rule was wrong, but just that the rule base developed was too small and no rule yet developed could deal with the case appropriately. Under such circumstances, the normal course was to further develop the rule base so as to be able to handle the failed case in hand. There were two different kinds of program extension, depending on the nature of the rule base deficiency found.

8.3.3.1 Addition of cases

As mentioned earlier, many of the students wrote the rules with a clear stereotypical motion in mind. It may happen that the goal statement can be satisfied by motions very different from the stereotype it was originally written for. In such cases, new rules with the same goal statement would be put in. For example, Ychan wrote a rule on downward motion caused by another object:

```
thing caused to move downward by otherthing if
      otherthing change the position of thing to new-place and
      new-place cannot support thing.
```

Obviously he was thinking of the case when something fell to the ground because the support it rested on moved away. Later on, it was suggested to him to test the rules with the case of a child playing on a slide. Is the child caused to move down by anything? He reckoned that the answer was yes the child was caused to move down by himself. However, throughout the motion the child is supported by the slide. He resolved this dilemma by adding in one more rule:

thing caused to move downward by otherthing if
otherthing change position of thing to new-place and
new-place can support thing and
new-place sloping downward.

Thus the difficulty was overcome by creating a tailor-made rule for another stereotype - sliding motions, motion of objects going down an incline.

8.3.3.2 Addition of goals

Program extension may also result when no rule with the appropriate goal statement has been written for the test situation. For example, in trying to deal with cases concerning causes of motion, Lin had written two rules:

object cause of motion self if
not (object needs external force in causing motion) and
source of force internal.

object cause of motion machine if
object make use of specially designed instrument and
source of force external and
type of force mechanical.

Later he asked the computer about the cause of motion for the case of a bicycle and the computer failed to give any possible answer. He inspected his rules and decided that he should add in a new rule on cause of motion by animals (since bicycles can be ridden by men as well as other animals like monkeys and bears):

object cause of motion animal if
source of force external and
type of force biological and
source of energy chemical.

This rule thus extended the possible categories of causes for motion.

8.3.3.3 Extension of knowledge base as accretion

The above two types of programming activity resulting from failed queries undoubtedly lead to an extension of the rule base (or knowledge base implemented on the computer) and may thus be classified as learning from failure. However, this kind of learning does not result in new conceptual structures being constructed nor existing ones being modified or replaced. This was because the rules that were being tested were either empirical observations (as in 8.3.3.1) or specific conjectures (as in 8.3.3.2) which did not set up any general model for prediction of behaviour nor explanatory models of why things move in certain ways. The rules that were laid down were essentially descriptions of specific situations. The failures were thus taken as failures to take note of yet more categories of specific situations. No dissatisfaction with the initial rules arose from the failures and the new rules were just extensions of the original knowledge structure to take more specific cases into account. This type of 'learning' activity is similar to what Rumelhart and Norman (1978) described as accretion:

> *"One basic mode of learning is simply the accumulation of new information. We analyze the sensory events of our current experience, match them with some appropriate set of schemata, form a representation for the experience, ... The newly created data structures are instantiations of the previously existing ones, changed only in that the representations for particular aspects of the current situation have been substituted for the variables of the general schema.*
>
> *This is learning by accretion: learning by adding new data structures to the existing data base of memory, following the organization already present."*

8.3.4 Reformulation for better performance

In a few instances, the program modification involved a complete reformulation of the rules concerned. A deeper mode of learning seems to be involved, as the changes evidenced a conceptual reformulation, either in terms of taking on a different perspective, or in changing previous limits of validity. Since these are the more interesting cases from the cognitive point of view, and there are only seven of them, a description of each is given below.

8.3.4.1 Regression back to pragmatic specificity

Only one of the seven cases falls into this category. Lin started working on rules about control of motion and was trying to build up a theoretical model of motion that categorized motions into two main types: those that are controlled by animate objects and those that are controlled by inanimate objects. Again, in writing down the rules he was thinking of specific stereotypical cases - inanimate objects being driven by animate ones, with the latter being the source of power for the motions concerned. However, he began by writing down a general hypothesis that built on more general concerns, namely source and type of force, indicating his intention to build up a more general theoretical model that could transcend the specificity of particular situations:

```
object control of motion animal if
      object source of force external and
      object type of force biological.
```

When he first tested the rule with the case of a moving bicycle, the outcome was quite satisfactory. The source of force was considered to be from a cyclist, which was thus external to the bicycle and the type of force from the cyclist was biological as the cyclist is animate. However, there was a problem when he tested the rule with the case of a moving car. He considered the source of control to be undoubtedly animal - the driver. However, the car's motion was not powered by the driver, but the car's engine, which is inanimate. He immediately saw that none of the maintenance or extension strategies described above could be used to retain the basic form of this rule. It required a basic reformulation for it to handle the test cases satisfactorily. Instead of attempting another theoretical model for the control of motion, Lin actually replaced the rule by an empirical observation:

object control of motion animal if
object purpose of motion decided by something and
something is a living thing.

This regression from a general hypothesis to a situation specific rule involving animistic conditions is actually another form of evasion. As evidenced in the discussion above, there is no major conceptual confrontation when an empirical observation or specific conjecture fails. On the other hand, it would be much tougher to defend a general hypothesis.

8.3.4.2 Advancing towards generality

Two other cases of reformulation took just the opposite bent. Evon was classifying all influences on motion as derived from either internal force (basically referring to cases of autonomous motion) or external force (for cases of non-autonomous motion). She wrote the following rule:

object under the influence of external force if
object affected by something and
not (something EQUAL weight).

She tested the rule and was satisfied with the output from the computer. However, she pondered further on the rule and decided to make it a more general one:

object under the influence of external force if
object affected by something and
not (something force-type internal).

This is a very interesting case because here the modification is not directly prompted by any need to handle failure cases. The dissatisfaction seems to be with the specificity of the second condition. She seems to value the possibility of classifying forces into internal and external types and prefers general statements to singularities.

Another similar case was found in Cheng's work. She, like a number of other students in the 15-yr-old group, had been quite impressed by an experiment that they saw in the school laboratory. This was the rule she wrote about the situation:

thing take same time to fall as otherthing if
thing fall from same height as otherthing.

However, when she was challenged by actually testing the rule out using a pencil and a piece of paper, she changed the rule to the following upon seeing the experiment results:

thing take same time to fall as otherthing if
thing fall from same height as otherthing and
not (thing made of paper) and
not (thing made of feather).

This modified rule produced results consistent with her observation, but she was not too satisfied. She asked if there would be any more exceptions to the rule besides paper and feather. She got no hint from the researcher, but this question seemed to worry her a lot. She came back the next session and proposed the following replacement for the earlier rule:

thing take same time to fall as otherthing if
thing fall from same height and
thing in vacuum and
otherthing in vacuum and
there is gravity force.

This rule obviously did not come out as a result of intuitive reasoning, and the fact that the experiment she saw in the school laboratory was done in a vacuum was not clear to her when she wrote the first version of the rule. When asked how she came up with such a modification, she said that she was not happy with the rule and wanted to find out more about the school experiment. The rule summarizes what she found out from the school textbook. Again, the interesting point here is that she consulted the textbooks because of dissatisfaction - dissatisfaction with specificities.

8.3.4.3 Change in limits of applicability

The other four cases of reformulation were all concerned with a re-definition of the limits of applicability of the rules.

Two kinds of such changes were found. One involved a complete removal of all conditions so that the rule changed from a specific conjecture to a general hypothesis that applies to all objects. This normally involved a recognition of the need for rigor and consistency in the use of terms. For example, Evon wrote the following rule on weight, intending it to mean that weight would only have effect on the motion if the motion has a vertical component to it:

object affected by weight if
object has change in height.

When questioned whether things moving on a level surface, for example a bus, would have weight or not, she had to answer in the affirmative. Though intuitively she still thinks there is some difference in the effect of weight on motion, she could not explicitly clarify the difference. In the end, she proposed to change the rule to:

object affected by weight.

This is a significant decision in that she did not try to formalize her intuitive feeling of distinction by creating a separate category for horizontal motions. This is in line with her preference for generalization as indicated by the discussion about her other rule modification in the preceding section.

One of Wat's modifications was done in a similar vein. She wrote:

thing can receive force if
thing is dead object.

She intended to express the idea that animate objects can move because they possess force, while inanimate objects can only move when they 'receive' forces from other objects. However, when she came to a test case of an animate object that moved because it was pushed, she realized that this distinction which she had hypothesized between animate and inanimate objects could not be substantiated. She then decided to reformulate the rule to:

thing can receive force.

The other cases of change in limits of applicability actually concerned the incorporation of learnt concepts into the program. The two instances were both modifications made by Wat concerning the topic of weight. The initial version of the two rules were:

i) thing weight increased if
thing affected by gravity.

ii) thing has weight.

She believed that all objects have weight, which she saw as an intrinsic property of all objects. However, she also believed that gravity has something to do with weight - in places where there is gravity, the weight of objects would be increased. It seemed that because of the programming task, Wat became more sensitive to the issues she was writing about. Neither of the two modifications were made because of failures discovered in running the program but were made because subsequent to the writing sessions she met materials (TV and books)that were contrary to her initial understanding. She then proposed to modify her rules to be in line with the new understanding:

i) thing weight zero if
not (thing affected by gravity).

ii) thing affected by gravity if
thing on earth.

Superficially, Wat's modifications are uninteresting cases of rewriting factual statements upon reading/learning about statements to the contrary. However, why was she the only student to write about such aspects of weight when all the students in the same group had been exposed to the same teaching materials? Furthermore, what prompted her to be so aware of these points as to bring them up the next session for modification when she had not been challenged about them? The issue of limit of applicability has typically been a non-problem in commonsense reasoning (Nagel, 1961). The fact that she was working in the PROLOG environment might have helped her see the importance of defining applicability limits by providing a formal syntax for externalizing her ideas. However, she must also have placed a high value on rigor, on clear definitions of limits of applicability, a quality that is also valued highly by scientists.

8.3.4.4 Reformulation as fine tuning

Like modifications in the maintenance strategies category, the reformulations described above represent accommodations to the knowledge structures. However, the changes brought about by the modifications in this group are more significant. Here, the students are concerned about the general performance of the rules - how well the rules would work for a variety of situations. The problem with the rules is not that they are wrong, but that they need refining. The refining is done by changing the generality of the rules or their limits of applicability. It is anticipated that the rules after refinement would produce satisfactory results for a wider range of situations, with the likely consequence that the modifications are relatively stable. This category of modifications is similar to Rumelhart and Norman's (1978) schema tuning where the basic relational structure of the schema remains unchanged and only the constant and variable terms referred by the schema are modified

8.3.5 Theory building

As has been pointed out by philosophers of science, one main distinguishing feature between science and common sense is in the organizational structure of these two bodies of knowledge (Nagel, 1961). Science is an organization and classification of knowledge on the basis of explanatory principles. It seeks to discover and to formulate in general terms the conditions under which events of various sorts occur. Commonsense, on the other hand, is exclusively preoccupied with the immediate consequences and qualities of observed events, that is, the pragmatics of the particular situation, and has little interest in systematically explaining the facts it notes, nor the range of valid applications of its beliefs.

Thus science is typically a well structured body of knowledge which is hierarchically organized, with a few most general, widely applicable principles at the top (i.e. being the most important) and the local variations and empirical laws at the bottom.

The analysis of the students' programs in the foregoing chapter showed that their work differs from one another in the amount of theorizing involved. However, even in the case of the theorizers (those whose programs are hierarchically structured with the general hypotheses at the top), the structures only evolved slowly and not as a conscious effort of theory building. The program structures reflected the kinds of connections they saw amongst the different parameters of motion. There was one exception. Fong, a student from the 17-yr-old group, had a clear sense of building a hierarchically structured theory of motion. He was the only one who attempted syntactic restructuring of the rules - to streamline the relation names so that hierarchically related rules can be fired as such. This implies that Fong possessed two important qualities: an appreciation and respect for well structured knowledge, and a quickness to pick up the peculiarities of the programming language. His first syntactic restructuring attempts started as an effort to streamline his rules so as to reduce the amount of typing he had to do. One of the first rules he wrote down was

object will be falling if
  object is initially at rest and
  object is under influence of gravity and
  not (object is supported by something).

Then he thought about other possibilities for falling to take place and saw that these different cases would all involve three conditions. Furthermore, he recognized that two of the conditions would be common across all the rules, and the rules only differ from one another in the condition concerning initial conditions. He thus modified the above rule to

object will be falling if
  object is initially at rest and
  object may be falling.

He then added in the following rule:

object may be falling if
  object is under influence of gravity and
  not (object is supported by something).

After this, he made use of the rule on '*may be falling*' to build two further rules on '*will be falling*' with the initial conditions of '*object is initially falling*' and '*object is initially moving upward*'.

A later case of syntactic restructuring was in a way purely aesthetic, solely for the purpose of highlighting the hierarchical structure he intended for his program. His whole program was essentially about the effects of different interactions on motion, and he began his enterprise by working on gravitational interactions:

object-A affected gravitationally by object-B if
object-A has mass and
object-A nears object-B and
object-B has a huge mass.

After diverging to work on some other rules concerning interactions of object in direct contact, he wrote a rule on electrostatic interactions:

object-A is affected electrostatically by object-B if
object-A is a charged body and
object-B is a charged body and
object-A nears object-B.

He then very quickly referred back to his earlier rule on gravitation, decided that grammatically the second version is better and modified the previous rule accordingly to:

object-A is affected gravitationally by object-B if
object-A has mass and
object-A nears object-B and
object-B has a huge mass.

He subsequently wrote three rules of the form

object-A is affected ____________ly by object-B if
object-A ....

for electrostatic and magnetic interactions between objects. His modification of the program line quoted above could thus be understood as his demand that the syntactic structure of the program should reflect the hierarchical structure of the 'theory' that he wanted to build. Pragmatically, because of the very restrictive syntax of the PROLOG that he was allowed to use, he actually had to put in many program lines to highlight the structure, which would not have been necessary if all he wanted was for the program to work and answer queries according to his ideas (c.f. fig.7.26 for the structure of Fong's program).

Fong was the only student who was consciously tackling the programming task as one of theory building, and his program did deserve to be termed a theory both semantically and syntactically.

8.3.6 Program modification strategies, program structure and learning

The foregoing analysis of the program modification behaviour of the students is done at an individual rule level and yet table 8.1b indicates that there is a high correlation between the structure of the program written by a student with the strategies s/he employed in the modification. Part of this correlation is a feature built into the analysis scheme and so is not at all surprising. By definition, only rules in the category 'general hypotheses' would lead to modifications that may qualify to be called 'reformulation' or 'theory building'. Thus Sandra's and Brenda's work exhibited employment of the maintenance and extension strategies only. However, there are a couple of interesting observations one can make by comparing the program structure with the corresponding modification strategies employed.

First, general hypotheses tended to go unchallenged if they are not related to specific conjectures and empirical observations. Wong's program has all the three kinds of rules, but the two general hypotheses in his program stand entirely unrelated to other parts of his program (c.f. fig. 7.17). This isolation leads to the sterility of these two rules and Wong's modification strategies had been rather uninteresting. All his eleven modifications were in the category of maintenance strategies.

Another interesting feature concerning the evasion strategy can be discerned if one tries to look at which kind of rules seem to be most prone to the greatest challenges. If one makes the assumption that the most serious challenges were posed by failures of those rules that led to the employment of the evasion strategy, then these challenges arose only from failures of the general hypotheses category. None of the failures of the empirical observations or the specific conjectures led to a complete deletion of the rules concerned.

If we take the program testing and modification process to be representative of the possible processes of interaction of a person's knowledge schemas with inputs from the outside world, such interactions should lead to learning. It is widely accepted that there are different kinds of learning, some more significant than others (c.f. chapter 3). If we take the view that more significant learning is necessarily accompanied by structural changes in knowledge schemas, Rumelhart and Norman's (1978) classification of learning in accretion, tuning and restructuring provides a good framework for examining the learning processes underlying the program modification activities discussed above.

The modification strategies of maintenance and evasion as described in the previous section do not seem to lead to learning in any real sense. The former probably only leads to temporary modifications in the knowledge schema in order to cope with the specific situation in hand and such changes may easily be reverted. As for the latter, it is an indication that the student is avoiding the conflict altogether instead of trying to resolve it. Two thirds of all the modification activities belong to these two strategies.

The second most numerous category of modification strategy (accounting for 25% of all modifications) is that of extension by further categorization. As explained earlier, these modifications are attempts to extend the existing knowledge structures in order that they be applicable to a larger range of specific cases. It is thus appropriate to see them as accretion processes in the context of learning.

The most significant learning process as represented by the modification activities is that related to reformulations. These in fact rarely happened - only 7 out of 116 cases, and it is expected that they represent relatively stable structural changes in the knowledge schemas. It is important to note here that most of these cases could be interpreted as an indirect outcome of school learning. Furthermore, with the exception of one case (Lin's case), none of the modifications was necessitated by a failure in the testing process, but were considered moves after reflecting on the meaning carried by the program lines perse. Both Lin's and Evon's rules for conceptual reform involved the notions of internal and external forces. Both students had interpreted such a distinction to be the underlying difference between autonomous and non-autonomous motions. This interpretation results from their conscious efforts to integrate their formal physics learning (terms like internal force and external force) into their commonsense framework of reasoning. The modifications made by Wat and Cheng, both from the 15-yr-old group, were affected by their school learning in another way. The changes made were at a much less theoretical level. They have learnt certain facts about weight which were contrary to their commonsense expectations and they modified their rules accordingly to take these into account.

None of the modifications recorded brought about fundamental changes in the knowledge schemas. There is no evidence of learning at the restructuring level having taken place. This is not altogether surprising given the short time span involved in the administration of this programming task. However, a close inspection of all the conflicts and resolution strategies involved in the modifications shows that very rarely do the program failures pose serious cognitive challenges to the students concerned. There is no indication, with the exception of Evon's case which will be detailed below, that any further development in the programming activities would lead to structural changes in the knowledge schemas.

It is reasonable to expect that the nature of the cognitive conflict and what the student may gain from the interactions depend very much on whether the programs reflect any core theoretical hypothesis and its degree of structuredness. In the case when a program demonstrates a strong theoretical hypothesis, very often the failures arise from invalid core assumptions/concepts which may not be resolved by temporary local modifications using maintenance strategies or by means of fine tuning. Thus failures in programs belonging to the structural epistemic frame may bring about more serious conflicts than those in the other types of programs. Evon's program failures fall into this category and they are described in detail here:

Evon had written the following rules for determining the cause of motion:

object motion caused by itself if
object force supplied by object.

object motion caused by machine if
object force supplied by machine.

object-1 motion caused by object-2 if
object-1 force supplied by object-2.

object motion caused by gravity if
not(object under the influence of other external force).

She then set out to test the program with the case of a falling apple (thinking about the famous 'Newton's apple'). To her surprise, the computer did not come up with the expected answer. The reason was that in one of her earlier sessions, she included weight as an external (supply of) force, together with forces like friction, air current, etc. Faced with such a contradiction, her immediate response was to exclude weight from her list of external forces. As a matter of fact, in her scheme of intuitive physics, there was no clear orientation towards classifying weight as internal or external forces since she could find justification for either classifications. Weight arises because of attraction from the earth which is external to the object, so it makes sense to put weight in the class of external forces. On the other hand, weight is related to the mass of the object which is an intrinsic property of the object itself, so it also makes sense to classify it as an internal force. However, after changing weight to an internal force, the computer did not give her the expected answer - gravity, because according to one of her other rules, anything placed in air would be affected by air-current, and air-current is an external force. Now, excluding air-current from the list of external forces seemed to be too much a twisting of facts. She resolved the problem by replacing the fourth rule listed above by:

object motion caused by gravity if
not (object motion caused by something).

to avoid the problem with specific instances of internal or external forces. She did not review her underlying concepts about sources of supply of force.

Unfortunately, this was not the end of the problems for her. She again met similar problems relating to internal and external supplies of force when she explored her rules on control of motion, and she could not find a satisfactory solution after patching up various parts of the program via maintenance or reformulation. At the end of the programming task, she was beginning to raise questions about the more central assumptions.

Thus though knowledge restructuring has not been recorded, a fair conjecture based on the present results is that only when the knowledge about a particular domain is relatively well

structured, possibly to the extent of being classifiable as at the structural epistemic level, would conflicts possibly lead to restructuring. Restructuring would be used only as a last resort when all the other strategies fail to resolve the conflicts.
Another important observation here is that both Evon's serious cognitive conflicts and the fine tuning of schemas pertaining to rule reformulations arise because of students' attempts to integrate their commonsense knowledge about motion with the mechanics knowledge they learnt in schools. There thus seem to be two pre-requisite criteria for significant learning in the sense of changes in knowledge structure to take place. The first concerns the learner's epistemology, whether s/he sees value in the theoretical and abstract, and the second is the presence of external stimulus that gives rise to cognitive conflicts with the learner's original concepts, beliefs.

8.4 A Matter of Value - Performance Rating Given to the Programs

In the last chapter, the students' programs were compared in terms of content and structure. How would the students themselves compare their own programs? What criteria would they be using in such a comparison? Would there be great individual differences between students in the criteria they choose? Answers to these questions would provide further insight into the epistemological values of the students.

For the group of F.6 students, their program development work was all completed within the first term of the school year. At the end of the school year when the school examinations were over, they were invited back to give a performance rating to all the programs developed by the group. In order that their assessment would not be affected by their personal relations with and esteems for the respective authors, the authors for the individual programs were held anonymous except when they were assessing the performance of their own program.

The fact that all the programs contain large segments about causes and agencies of control for motion makes it much easier for a comparison of the programs to be made on their performance with respect to these aspects. Each student was asked to pose six queries to each of the programs being tested and to record their answers:

car motion affected by what?
car motion caused by what?
car motion controlled by what?
upward-moving basketball motion affected by what?
upward-moving basketball motion caused by what?

upward-moving basketball motion controlled by what?

Each student had to give for each program, except his/her own, a rating for each response from the computer and the overall performance rating, and these were recorded on a sheet of paper (fig. 8.1). At the end of the exercise, each student was interviewed to find out how they arrived at the rating scores. The interviews were audiotaped and a translated version of the interview was transcribed (as with the classification and programming task, these interviews were conducted in Chinese).

Because of the difficulties in arrangement, only five students from the F.6 group took part in this rating exercise. Furthermore, some of them could not complete the rating for all the programs. Thus the administration of this task is far from satisfactory and it is not possible to make any definite claims on the results. However, some of the initial observations are so interesting that they are included here as a corollary.

Despite the incomplete data, it can still be seen that the program with the highest rating so far is the one written by Brenda. With the exception of Fong, all assessed the programs in terms of efficiency (how few queries would the computer pose for the user before it can come up with an answer) and accuracy (how well the answer matches with their expectations). Though they were all allowed to ask the computer for explanations of how it has arrived at a particular solution, very rarely (except Fong) would they make use of such a facility unless the answer was reckoned to be unreasonable. It was as if how the answer was arrived at was of little value, only the actual answer really mattered. Because of the pragmatic outlook adopted in the rating exercise, Brenda's program was rated as the best by all the others except Fong. This result is very interesting and in a sense illuminating. As can be seen from fig. 7.15, one main feature of Brenda's program is that it is an essentially flat structure without any hierarchical relations. All the rules were either empirical observations or specific conjectures, and the whole program operates at the level of the specific epistemic frame. Further, the premises for the rules represent the most prominent or most highly correlated features of the stereotypical situations the rules were written for, and generally bear no causal relations to the goal statements. The following rules from her program illustrate these characteristics:

object motion caused by wind if
object moving unsteadily and periodically in air.

object motion caused by a push if
object initially at rest.

Program name		Car			Upward-moving-basketball			downward-flying-bird			falling-leaf		
pa		Answer	Rating	Comments (on explanation, questioning, etc)	Answer	Rating	Comments (On Explanations, Questionings, etc)	Answer	Rating	Comments	Answer	Rating	Comments
	Motion-affected-by	gravity friction	3. 5.	the effect of gravity is not so noticable. & the effect is indirect. power of engine ~~and~~ should also affect the motion?	air resistance gravity friction tension	5 5 5 1	but it must in contact with air. but it ~~is~~ is equal to air resistance. (the basketball is suspended in air when moving).	gravity. Tension friction	4. 1 5.	(- - - - -)	air resistance gravity tension friction	5. 5. 1. 5.	(- - -) (= air resistance)
	Motion-Caused-by	a push. man engine	2. 3. 5	it is only in a special case. motion of car is controlled by man but it is not necessary caused by ~~man~~, e.g. other animals can start the car motion e.g. monkey.	a push man	5. 5.	(Is it true that basketball far away from control-centre when it starts moving ?)	itself.	5.	(gravity ~~...~~ is also a factor that caused the motion).	a push	1.	~~the push~~. gravity or 'weight' is ~~...~~ the true ~~...~~ answer. ~~air weight~~
	Motion-controlled-by	engine. air current friction man	~~4~~ 5 2 5 5.	~~current~~ normally, air current can only affect the motion but not ~~controlled~~ controls it.	air current tension	4 1	~~it~~ basketball motion is affected by air current but not necessary controlled by it (basketball is suspended in air).	itself air current tension	5. 3. 1.	(. . . .).	gravity air current	5. 5.	air resistance can be the other factor that controls the motion.
		forRating: Use a 1 - 5 pt. scale, 1 - totally unacceptable, 2 - not reasonable, 3 - maybe, 4 - quite possible, 5 - good answer OVERALL: Rating — 4. Comments — quite good., the programme can apply to most ~~of the~~ cases and give the possible solution to the questions. (~~...~~ close to commonsense). (concrete st. forward reasoning) Name of evaluator CW Wong											

Fig. 8.1 A sample evaluation sheet for rating the performance of the programs.

Table 8.2 The performance rating given by the students# to the various programs.

program*	Rating@ given by			
	Brenda	Evon	Sandra	Wong
PA		5	4	4
PB	5		3	
PC	4	3	3	3
PD	3	4	4	
PE	3	3		
PF	3	3	3	

@ The students were asked to rate the performance of each program according to a 5 point scale: 1 - totally unacceptable, 2 - not reasonable, 3 - may be possible, 4- quite possible, 5 - good answer.

\# Though Fong also participated in the rating exercise, but because he refused to rate them according to the five point scale, it cannot be recorded here.

* The authors for the six programs are: pa - Brenda, pb - Evon, pc -Fong, pd - lin, pe - Sandra and pf - Wong.

Fong's program is the best in terms of its coherent structure and consistency with the Newtonian view of motion. However, it was not well received by the other students. The main reason was that because of its concern for logical sufficiency and general applicability to all situations, the users have to answer a whole host of apparently irrelevant questions. Furthermore, because of the symmetrical relationship between the agent and the affected encapsulated by the program, the program easily went into infinite loops while answering a query and this was extremely frustrating for the users. The following is an extract from a typical rating session for Fong's program:

Fong : basketball motion controlled by what?

Computer : basketball move-along-path what?

Fong : parabola.

Computer : what plan the path parabola?

Fong : man.

Computer : is it true that basketball possesses mass?

Fong : yes.

Computer : is it true that basketball nears man?

Fong : yes.

Computer : is it true that man possesses a huge mass?

Fong : no.

Computer : is it true that basketball is a charged body?

Fong : no.

Computer : is it true that man is a charged body?

Fong : no.

Computer : *is it true that basketball is a magnetized body?*

Fong : no.

Computer : *is it true that man is a magnetized body?*

Fong : no.

Computer : *is it true that basketball is in contact with man?*

Fong : no.

Computer : *is it true that basketball is in man?*

Fong : no.

Computer : *is it true that man collides with basketball?*

Fong : no.

Computer : *basketball nears what?*

Fong : man, earth, air.

Computer : *is it true that earth possesses a huge mass?*

Fong : yes.

Computer : *is it true that earth possesses mass?*

Fong : yes.

Computer : *is it true that earth nears man?*

Fong : yes.

Computer : *is it true that earth is a charged body?*

Fong : no.

Computer : *is it true that earth is a magnetized body?*

Fong : yes.

Computer : *is it true that earth is in contact with man?*

Fong : yes.

Computer : *is it true that man has a rough surface?*

Fong : yes.

Computer : *is it true that earth is in man?*

Fong : no.

Computer : *is it true that man collides with earth?*

Fong : no.

Computer : *earth nears what?*

Fong : man, air, basketball.

Computer : *is it true that air possesses a huge mass?*

......

......

Computer : *earth is in what?*

Fong : air.

Computer : *is air a fluid media?*

Fong : yes.

......

Computer : *Answer is: basketball motion controlled by man.*

Fong : why?

Computer : *basketball motion controlled by man because*
basketball move along path parabola and
man plan the path parabola and
man has the power to influence motion of basketball.

man has the power to influence motion of basketball because
man indirectly influence the motion of basketball.

man indirectly influence the motion of basketball because
earth directly influence the motion of basketball and
man indirectly influence the motion of earth.

man indirectly influence the motion of earth because
air directly influence the motion of earth and
man directly influences the motion of air.

air directly influence the motion of earth because
earth is affected by viscosity of air.

man directly influences the motion of air because
air is affected frictionally by man.

Fong : Any more answers?

Request for more answers to the question triggered off another burst of similar questions, and the program went into an infinite loop of looking for 'indirect influences'.

Fong himself was aware of the shortcomings of his own program when put to the test for pragmatic everyday situations. However, he was not convinced that Brenda's was the best. In fact, he was the only one who cared more about how the resolutions were made, that is how the computer reasoned in getting the answers, instead of what the actual answers were. When he first met Brenda's program in the rating session, he was very surprised at the efficiency of the program and was not satisfied with the reasoning given by the computer. His immediate reaction was to reject the way the program works as it was so different from his own. He then asked for a listing of all the programs to scrutinize at his own leisure. An interview with Fong was conducted after he had finished reading the programs several days later. The following is an extract from the interview:

> Fong was commenting on that part of Evon's program dealing with agencies of control for motion. (* PA refers to Brenda's program, PB refers to Evon's program and PC is Fong's program.)
>
> Fong: The rules on 'motion controlled by' have great potential of development.
>
> I: Why do you think that the rules on 'motion controlled by' has potential?
>
> Fong: I don't know... I had never thought of such a categorization. It is more useful, say. ... It has practical usage, it helps to anticipate ...
>
> I: You find the categorization into external and internal is good?
>
> Fong: Yes.
>
> I: So, you think its potential lies on the distinction between external and internal.
>
> Fong: Yes. It can distinguish the causes of motion, whether the motion is determined by internal factors or external factors. It has practical use. Besides, if this part is general enough, it can be linked up with the other parts.
>
> I: In Physics, we have not talked about 'force supplied by', we only talk about 'force', and there is nothing concerning 'supplied by'. Isn't that right?
>
> Fong: I don't know.
>
> I: In Physics there are internal forces and external forces. Say, the bird flies - Is it caused by an internal force?
>
> Fong: His [referring to author of PB] concept of internal force is different from what we are talking about. It talks about those forces that are generated out of its own wishes.
>
> I: You mean it is out of its own wishes?
>
> Fong: Not wish, but his 'internal force' is different from that used in Physics because ... they are very different.
>
> I: What are the differences?
>
> Fong: I think internal forces refer to those forces that are generated when the object is in motion, i.e., force supplied by object in motion. If the force comes from other object, it will be external force.
>
> I: Is this concept different from that in Physics?
>
> Fong: In Physics, internal forces come from the interaction of the particles inside the system.

I: If we follow the definition in Physics, say the statement 'force supplied is internal', what does it mean?

Fong: I cannot tell much from this. If there is external force, I can only predict what would happen to the centre of mass. What happens to the object, I do not know. For internal forces, there is no change in the centre of mass, but other things in the surroundings move. I cannot predict what will happen. It depends on the system.

....

Fong then went on to discuss Brenda's program. He was not entirely happy with the rule 'object motion controlled by friction if object move by using wheels'.

I: What is wrong with the rule about 'wheel' and 'friction'?

Fong: Here it says if it uses the wheel to generate motion, it must be affected by friction. At first, it seems meaningless but later, after more careful thinking, I found that he must have considered the case clearly before he wrote these things. His thinking was helped by practical cases. The program is not general. So, when there are some general questions, it doesn't work. The program becomes meaningless, say, I test it with an electron, I asked about the motion of an electron. It gave 'no more answers'. However, I still found that it is intelligent in some parts of the program.

I: Where are they? Are they inside specific cases?

Fong: Yes, like when dealing with the cases car, basketball, ... , it asks a lot of questions and they are meaningful too. The answers are exact and correct but those in between inferences are jumping too fast. So, I rank it as the 3rd in the potential to develop.

....

Fong: At first I thought that my program also has greatest developmental potentiality, but later, I find that in PB's program, when it is developed, it has a lot of practical uses, especially this part. If PB is properly developed on the part 'motion controlled by', it will have practical uses.

I: Can you give examples? I cannot think of any practical use.

Fong: Like, to determine whether a particle can control itself or not.

I: You mean the division in force supply into internal or external is very crucial?

Fong: Yes. It is very important. The linkage of 'affected by' and 'controlled by' in my program is practically useless. Also the arguments are not accurate. But they can be linked in my opinion.

I: You mean the linkage can be made in his program?

Fong: No. I mean adding my part into his program to make the linkage. So I think it has the greatest developmental potentiality. I ranked it as the 1st.

I: How about yours when compared to this one, PC versus PB?

Fong: PB is more accurate and practical. The arguments are acceptable. But as for develop-ment, I find that they are not flexible actually. If they can allow more areas for themselves, they can change or add things easily afterwards.

I: So, they are not easy to develop?

Fong: Not easy to be developed but by using his concept, it can be developed still.

I: How about reasoning? Not on the whole structure. Are the statements logical and acceptable?

Fong: PA is the most reasonable. Mine is not reasonable at all.

I: The answers are not reasonable in your program. But I am interested in the statement.

Fong: He [referring to author of PA] has simplified the statements. His performances are good in the case of 'car' and 'basketball'.

I: You think his statements are reasonable.

Fong: Yes. Some parts in PA are better than this one PB. Because in PB, some parts on "controlled by" are not correct. His way of thinking is good but still there are some mistakes.

I: So, in terms of abstract reasoning, PB is the best but in testing results, PA is the best. How about yours when compared with PB on development potential?

Fong: It is hard to say. In the short term, PB is better, I think the result will be the best when three are combined. You have to find somebody who think as accurate as PA, and flexible as PC, the ideas have to be as new as PB, so combining the three will be the best.

I: You think PB is inventive?

Fong: Yes, perhaps it is because I have never thought of these things before.

The clarity of thought expressed by Fong in this interview is remarkable and his comments on the value of the three different programs written by him, Evon and Brenda provide important insight for science educators.

It is evident from the interview that Fong can distinguish clearly the difference in meaning for the terms internal and external forces as used scientifically in the Newtonian sense from Evon's usage of these two terms. However, he found Evon's intended usage of these two terms to stand for sources of power for motion to be of much greater pragmatic value than the Newtonian connotation he used them for. He also found the ability of Brenda's program to deal with specific cases very impressive. In a way, Fong's criticism of his own program is applicable to the way school science knowledge is presented to students.

Scientific knowledge as presented (if interpreted correctly) represents an entirely different system of looking at the world which does not address important issues and apparent distinctions in daily life. Attempts to link up school knowledge with everyday concerns as in Evon's case bring about misconceptions. Further, the scientific laws and principles are usually so general that they fail to cope reasonably with specific daily situations.

As Fong pointed out, there is value in general rules in that they can go beyond the specifics and handle wider ranges of physical phenomena. In the long run, it is still programs like his and Evon's that has long term potentials for development. However, is Fong's proposal to link up the three programs workable? Is it possible to make science learning more meaningful and motivating by showing how scientific principles relate to pragmatic concerns and distinctions, and how they can be applied to deal with actual specific (and "messy") situations of everyday life?

Chapter 9 CONCLUSIONS

The two tasks in this research have yielded rich data which throw light on a variety of issues, theoretical as well as methodological. This final chapter is an attempt to summarize the findings for each task before presenting a synthesis of what we can conclude in terms of our understanding of commonsense reasoning about motion as well as the techniques used for elicitation in this research.

9.1 The Classification Task

A classification task was chosen as one of the methods of elicitation because as a cognitive activity, categorization is somewhere between perception and thinking (Neisser, 1987). It was anticipated that the classification schemes used would be affected both by features of the entities classified, in this case motion events, as well as the beliefs held by the subject about the domain.

9.1.1 Types of classification schemes used

An analysis of the contents of the classification schemes used by the students shows that they can broadly be classified into two types of schemes, descriptive ones and causal ones. The former accounts for two thirds of all classification schemes produced and these attempt to answer the *'what'* and *'how'* questions about motion, while the later are those that are concerned with the *'why'* questions. The descriptive schemes are thus concerned with issues roughly within the domain of kinematics and the causal schemes broadly dynamical ones.

There are three kinds of descriptive schemes: those classifying the motions according to pragmatic concerns and non-physical aspects of the motion, for example the intention or feelings of the moving object (D1 schemes); those classifying the motions according to holistic descriptions of the physical features of the motion, for example the type of motion (i.e. falling, sliding, etc.) or type of trajectory (D2 schemes); and those classifying the motion events according to a single physical parameter of the motion, for example speed or type of object involved.

The causal schemes can also be further subdivided into three groups: those classifying motion by naming the specific causal agent involved, whether concrete or abstract (A1 schemes); those classifying according to specific characteristics of the causal agents, for example the nature of the agent and whether the agent is moving (A2 schemes); and those classifying according to the mechanisms involved in production of the motion events (A3 schemes).

9.1.2 Commonsense reasoning about motion is predominantly concerned with pragmatic issues of the immediate situation

A close inspection of the classification protocols reveals that parameters used in the D1 schemes concern pragmatic issues and perspectives which permeate through all the other schemes, descriptive as well as causal. Issues like the intention or feeling of the moving object, whether the moving object has the ability to control the motion, or the motivation for the motion are never dealt with nor mentioned in mechanics lessons. They bear no definite relationship with the actual physical attributes of motion, and yet they are very real and important problems in everyday encounters with them. Parameters used in the D1 schemes thus define the context for the motions and the ultimate goals for the processing, that is, the purpose for reasoning about the motion events in commonsense contexts. The D2 schemes, on the other hand, tries to look for prototypical motion labels for each of the events. The motivation for such classifications, as revealed by the classification protocols, seems to be pragmatic: The ability to assign a label to a motion would allow a person to anticipate subsequent developments for that motion, and subsequently to arrive at an optimal strategy for dealing with the situation. The D3 schemes are also related to similar concerns. The animicity of the moving object and speed are the most popular schemes under this category. These two parameters, especially the former, determine largely how the motion event is handled.

All of the subjects, except three, produced causal schemes and these account for one third of all the schemes produced. This can be taken as an indication of the strong cognitive desire to look for causal explanations, to find out why a motion is developing the way it is, when reasoning about motion in commonsense everyday contexts. However, the level of explanation that is sought for is generally very low, mainly at the level of explaining away (Schank, 1986) the problem. A1 schemes offer explanations to the extent of naming the causal agents: falling motions are caused by falling forces, tripping motions by obstructions, etc. A2 schemes classify causal agents according to their characteristics, and 70% of these classified causal agents into animate and inanimate ones. Thus these two types of causal schemes do not contribute to deeper understanding of the causal agency for the particular situations than the D2 or D3 schemes. This situation is reasonable and expected if Schank's hypothesis that one would not try hard to understand deeply unless there is a strong need for it is valid. Such explanations are good enough for the pragmatic concerns of classifying motion events into prototypes so as to work out the best way to handle them in commonsense contexts.

Schemes in the A3 category offer deeper levels of causal analysis, classifying motion according to the mechanisms responsible for bringing about the motions. Two types of explanations were offered in these schemes: analogical models and aggregate effect of forces model. The former identifies certain physical components as playing the crucial causal role as in some common mechanical devices, for example springs, levers, etc., and the whole system is assumed to behave in an analogous fashion. The latter type of explanation is based on the belief that motions occur in the direction of the aggregate effect of the forces acting, and it is frequently the case that forces acting in opposite directions would work in contention, and the 'winning' one would take effect while the 'losing' one would remain dormant, producing no effect on the motion. It is believed that these schemes provide linkage points to the study of mechanics.

A comparison of the classification behaviour of the two age groups shows that there is no fundamental difference between them except that the older group is able to produce significantly more D2 schemes, possibly because their background knowledge in Physics enables them to have more ways of describing the motions and giving it a holistic label. The same fundamental concern for pragmatic issues is prevalent in both groups, indicating that instructions in mechanics have not fundamentally changed the basic outlook in commonsense reasoning about motion.

9.1.3 Commonsense knowledge about motion stored in large complex encapsulated units of memory structure based on prototypes

Another important observation from the results is that the reasoning about motion as exhibited in the classification protocols is mainly in the form of pattern matching and not logical inferencing according to certain principles. The most fundamental processing seems to be to look for prototypes with which to label these motions. Knowing the type of motion involved, together with a knowledge of one or two other important parameters like the animicity of the moving object, the speed, or the animicity of the causal agents, seems to be sufficient to act as indices for determining the value for all other parameters. This supports the idea that knowledge about motion events for the purpose of commonsense reasoning is stored in units that are larger and more structured than semantic nets. Such knowledge structures are likely to be at the level of complexity similar to Schank's (1975) scripts or Minsky's (1975) frames. The advantage of using such large and complex structures (or scripts in the general sense) for memory storage is that the processing time can be much reduced since each unit would hold default values for most of the motion parameters (or slots).

9.2 Conceptions of Motion as Expressed by the Expert Systems Developed

The students participating in the programming task had no difficulty in writing PROLOG rules about motion, despite the fact that most of them had had no previous programming experience. However, their interpretation of the task at hand was much less formal than may be expected from the logical syntax and the nature of the programming. Very often, the students were thinking of very specific situations or 'prototypes' when they were writing down the PROLOG rules, so that though the rules took the form of generally applicable statements, it was usually not the students' conscious aspiration to do so.

9.2.1 Important aspects of motion as revealed by a rule level analysis

At the individual rule level, the kinds of concerns expressed in the rules are very similar to those found in the classification task. About half of the rules were about aspects and attributes of motion and another one third were about agency of cause or control for the motion. The percentage content distribution of the rules for the two age groups of subjects are very similar, with the older group showing more interest in issues of agency and the younger group relatively more interest in describing motion.

Of the rules describing attributes of motion, the two most popular subgroups are of the format:

motion type if
observable physical conditions

and

motion attribute (abstract) if
observable physical conditions.

It can be seen that the former type of rules are essentially empirical correlations trying to determine a holistic label for a motion given a small set of observable features of the motion, and as such perform functions very similar to the D2 classification schemes. The latter type of rules try to make deductions about some abstract attributes of the moving object, like force or energy. These rules provide important links to the next largest group of rules, those about causal agency for the motions concerned.

The rules about causes of motion try to predict the causal agents involved when certain physical attributes of the motion are given, or vice versa. Both age groups expressed autonomous and non-autonomous motion as one major distinction in relation to agency, though the younger group tended to do it much more explicitly in the form of "*cause agent <- nature of object*" whereas the older group did it more subtly in terms of "*internal*" and "*external*" forces as the agents for the two kinds of motion. There is more diversity in the entities the older students used as labels for the agents: physical objects, forces or circumstantial factors such as the actual physical situation for the motion appeared in their programs as causal agents. In general, there was a tendency for the older group to label force as the immediate cause of motion while the younger group confined causal agents to physical objects, and force is seen as the necessary ingredient for motion to take place.

There is little consensus in the understanding of the term "control". It may be used as a synonym for cause, the dominating factor amongst a number of influences, a causal agent that can exercise intentions and will-power, or a possible state of a moving object (that is, the motion being under control or not).

9.2.2 Patterns of agency

The notion of causality is a consistently important theme coming up in both the classification task and the programming task. Whilst it is not easy to map out a clear and relatively complete picture of a subject's conception of agency from the classification protocols, it is possible to construct a more systematic picture of a student's causal conception from the expert system written. This is mainly because the programming task is a creative task unconstrained by any given motion event, problem situation or concept, as is often the case in other techniques of elicitation.

An analysis of the programs in terms of the "story" or underlying theme of the programs with reference to the causal conceptions they tell shows that they differ from each other along two important dimensions: mediation and reciprocity. Three out of the twelve programs only referred to concrete physical entities as causal agents, with no explicit mention of the need for any mediating entities for the causal action to take place. For these programs, the agent and the affected necessarily play distinctly different roles, with the former being the active party and the latter the passive one. Of those hypothesizing a mediating entity, normally force, as playing a crucial role in the causal action, most of them still highlighted the "active" role of the agent as opposed to the affected. Only two of the programs detached the attribution of agency from an entity to an interacting event. The mediating agent for the causal action was no longer derived from the physical agent directly, but from an interaction between the agent and the affected. Such a move highlights the fact that both parties are affected during such an interaction, making the causal action much more symmetrical.

It is hypothesized here that some programs did not make any reference to mediation not because the authors did not hold ideas of mediation, but that they do not see any pragmatic value in writing these down. A rating exercise carried out by some of the older students on the programs written by others in the same group indicates that a well thought out program written according to a model of unmediated agency can be the most efficient. Thus whether mediation is mentioned is related to what the author considers as important knowledge for commonsense operations. Those who referred to it did so mainly to satisfy their innate desire to express generalizations and abstractions.

The ability to see a reciprocal reaction in the causal interaction is a crucial step towards a more scientific understanding and conceptualization of motion. It is not clear how this change was accomplished in the two cases found. However, the teaching of classical mechanics might be promoted if the notion of causation can be discussed and linked explicitly to an interacting event rather than a physical object.

9.3 Epistemic Planes of Operation - a Matter of Intellectual Aspiration

It is the contention here that the programming task, as an unstructured, unconstrained, creative task, can be viewed as an explanation task where the author of a program tries to explain how the world of motion works in general by answering questions s/he sets up herself/himself. Under such circumstances, there is no fixed criterion as to what counts as a good explanation and it may vary from the very specific to the very general. The choice of questions and the level at which they are answered reflect the kind of knowledge the student figures as most important for understanding the world, that is, the sort of *epistemic questions* s/he is concerned with.

At the individual rule level, three kinds of rules can be distinguished along the dimension of generality. On the most specific end, there are the *empirical observations* which describe specific situations or concrete entities. These essentially capture prominent features of important prototypes of motion, and as such are descriptions at the level of encapsulated scripts for motion. Next are the *specific conjectures* which hypothesize the role of abstract entities in specific situations. Though these rules are also concerned with specific phenomena, they represent the first move towards generalization. This is because in the search for generalities applicable to all types of motion, invariances cannot be sought from concrete physical attributes but must necessitate the invention of an abstract entity, in essence the mediating entity for the causal action (there is simply no single kinematic system applicable to all motions and invariances only exist at the dynamical level). The specific conjectures also play the crucial role of providing possible linkage points between specific situations and further generalizations. Last came the *general hypotheses* which postulate general relationships between entities, abstract or concrete, that would be applicable to motions in general.

To look at each program as an integrated whole, inspecting how the different kinds of rules relate to and orchestrate with each other provides a deeper insight into the sort of epistemic questions the program is attempting to answer. There are programs that contain only empirical observations and specific conjectures. For these, the program taken as a whole does not go beyond the specifics of particular prototypes of motion and may justifiably be said to operate within a *'specific' epistemic frame*. It is hypothesized here that these programs do not contain any general hypotheses not because the authors do not entertain analyses of a more general nature, but that they do not see much utility in including them. This hypothesis is especially reasonable if the problems the authors are concerned with are the pragmatic ones of seeking appropriate ways of handling the situation efficiently. The programs thus reflect where the authors' priorities and values lie.

Most of the programs contain some rules in the general hypothesis category. However, in a large proportion of these programs, there are often no explicit relationships amongst the different general hypotheses, nor links across the general hypotheses, the specific conjectures and the empirical observations. This isolation of the general hypotheses from the other two categories of rules means that the former would have no utility at all in terms of contributing to reasoning about specific situations. Thus such programs though aspiring to operate at a *'general' epistemic frame*, in fact face the dilemma of not being able to cope with specific situations except by operating within the specific epistemic frame alone. This dilemma may account for the fact that one can find both general hypothesis and empirical observations within the same program written for the same kind of goal statements: this indicates an oscillation between the two epistemic frames.

Two of the programs exhibited much higher levels of structuring as to enable them to operate as one integrated system. Here, the actual number of general hypotheses may be very few, but they are hierarchically related to each other as well as linking up with the empirical observations via specific conjectures. Each of these two well structured programs in fact present a prominent core hypothesis, a theoretical framework for the interpretation and processing of motion events, and as such operates at the 'structural' epistemic frame. This structuredness is independent of whether the core hypothesis is consistent with the Newtonian view or not. Rather, it relates to how committed the person is to the value of a generalizable framework as opposed to empirical piecemeal strategies when the latter probably perform more efficiently and adequately for commonsense situations. This commitment possibly corresponds to what Polanyi (1958) referred to as "intellectual passions".

9.3.1 Epistemic frameworks and learning

One interesting observation here is that whilst the younger group produced roughly the same proportion of rules in all three categories, more than half of the older group's production of rules is in the category of specific conjectures, and proportionally less for the other two categories. A possible explanation for the F.6 group's greater tendency to produce specific conjectures is that they were primed to do so through the formal physics instructions they had received: they had been taught to analyze the nature, direction and magnitude of forces acting on different objects in specific settings.

If the specific conjectures do play a crucial role in enhancing the utility of generalizations as suggested earlier, then instructions in mechanics should help a person to move towards operating in the structural epistemic frame. Furthermore, learning in mechanics (or possibly other areas of science) may be made more meaningful if knowledge about the particular domain can be more clearly delineated into these three categories according to their generality and the issue of links across them be addressed more explicitly in teaching.

Another remarkable feature of the epistemic frames as elicited from the programs is that they correspond closely to the three mechanisms for equilibration , intra, inter and trans, as proposed by the Piagetian school (Piaget, 1975). The intra, or within scheme, mechanism takes place when a number of isolated facts are known, identified and analyzed independently of each other and is thus similar to the specific epistemic frame of operation. The inter, or between scheme, mechanism takes place when specific facts are found to be connected by transformations that leave something invariant, and is thus very similar to operations within the general epistemic frame. The trans, or totality, mechanism takes place when a structure is found that explains the transformations and the invariance, and each particular case. This last mechanism corresponds to operations within the structural epistemic frame. This close correspondence is an unanticipated and pleasing outcome which gives one some confidence that the method has elicited a real and in some sense a profound level of students' thought. Furthermore, this suggests that the three epistemic levels identified correspond to a progression that necessarily accompanies different phases of development in a progression towards the development of a scientific understanding.

9.4 In the Face of Conflicts - Program Testing and Modification

One important advantage of using expert system development as a task for elicitation is that the elicited knowledge is explorable and modifiable. The testing and modification episodes give rise to an explicit trace of the thought processes going on when a person tries to apply his/her externalized knowledge and meets conflicting situations. These two processes not only allow the researcher to probe deeper into the students' understanding even after the initial elicitation in the form of program development, but also provide important insight into a variety of issues: the kinds of possible cognitive conflict that may arise when one's beliefs are faced with challenges from actual everyday observations, the cognitive strategies that may be employed to cope with such conflicts, and the possible cognitive outcomes for different conflict resolution strategies.

9.4.1 Criterion for satisfaction

An inspection of the program modification behaviour of the students reveals that the criterion for satisfactory performance set up for the programs varied greatly across individuals. Most of the students took a pragmatic and utilitarian stance in assessing the performance of programs: they are concerned only with the acceptability of the answer itself so that the reasoning is of no consequence as long as it can give the right answer for the case under consideration; there is no expectation that rules should be infallible for all cases and answers are judged to be acceptable within a tolerance limit as long as they do not differ significantly from commonsense expectations. However, a few of the students went beyond the pragmatic and employed more "aesthetic" standards. These students tended to be the same group who produced programs at the structural epistemic level. Their programs built up a systematic and general explanation for motion, and they not only demanded that these produce correct answers, but also that the reasoning used is acceptable. Inherent in such an "aesthetic" position is the belief that there are some hidden principles, some grand laws of nature, that govern the world of motion and give a systematicity to the myriad manifestations of different motion characteristics.

9.4.2 Program structure and conflict resolution strategies

As the programs were tested using specific instances of motion, the first rules that would be fired would necessarily be the empirical observations or specific conjectures. There is little opportunity for the general hypotheses to be challenged except for programs written at the structural epistemic level. Most of the modifications that actually took place were made to the empirical observations and specific conjectures and were thus concerned with rules written for specific stereotypical motion events.

There were five types of modification strategy as exhibited by the program modification behaviour of the students.

First, there was the *maintenance strategy* which tried to cope with specific failures by adding, deleting or modifying conditions from the premises of rules. This effectively retains the same basic knowledge structure while making minor adjustments by relaxing or tightening the conditions so as to cope with the specific case in hand. There is no evidence that there were strong commitments to these changes and it is expected that this strategy only produced *episodic accommodations*.

Sometimes, a program did not produce expected answers not because of the failure of certain rules, but the absence of suitable rules to handle the specific cases in hand. Under such circumstances, students normally responded by adding more rules that adopted the fundamental basic structure as the existing ones in order to cater for these extra cases. This strategy is thus one of *extension* of the existing knowledge structure to account for more specific cases. This modification can be interpreted as *accretion* (Rumelhart & Norman, 1978) in the context of learning.
In a few of the cases, the students could not find acceptable alternatives to the failed rules and had to resort to deleting the rules altogether. This was actually a case of *evasion* on the student's side to face and resolve the conflict. All the evasion cases involved general hypotheses, indicating that failure of rules in this category in fact posed the most serious conflicts to the students.

A few of the modifications brought about more significant changes in the knowledge structure. These came about when a student decided to *reformulate* the statement of a rule so as to change the level of generality of the rules (to make it more specific or general), or to re-define the limits of applicability. Here, the students were not modifying the rules because of failed rules for specific cases. Rather, they were concerned with the general performance of the rules, how well the particular rules would work for a variety of situations. The rules were not wrong, but need refining. This category of modifications is similar to *schema tuning* (Rumelhart & Norman, 1978) where the basic relational structure of the schema remains unchanged and only the constant and variable terms referred by the schema are modified. A potentially important observation here is that many of the changes in this category were not directly brought about by failures, but came about as a result of the students' unprompted reflections on what they have read or heard about the topic in scientific contexts.

Only one student made syntactic modifications to the program in order to streamline the relation names to enable hierarchically related rules to be fired as such. This means that the student took the programming task as one of *theory building*, presenting a unified view of looking at motion events. The motivation for so doing is unrelated to the immediacy of coping with particular situations, but follows from an aesthetic desire and a respect for well structured knowledge. Such modifications also imply that this student had the remarkable ability to pick up the features and possibilities of the programming language which were not introduced to the students.

9.4.3 Cognitive conflict and learning

Results of the present research show that cognitive conflict rarely leads directly to learning in a significant sense. Most of the program failures led to episodic accommodation only and as such does not qualify to be called learning. A quarter of the rule modifications involved application of existing knowledge structures to new cases. These constitute the lowest form of learning - accretion. The highest form of learning observed during the experimental period was schema tuning. Here the knowledge structures were modified slightly to improve their general applicability. There was no observation of restructuring taking place.

A fair hypothesis from the above observations is that there is a general tendency to handle conflicts in ways that make the minimum possible changes on the existing knowledge structure. Conflicts lead to more serious challenges if the knowledge schemas are better structured while less structured ones can cope with failures by patching: piecemeal modifications for specific situations. Failures for programs at the structural level posed the greatest difficulties to the authors because peripheral modifications very often cannot solve all the problems because of the inter-relatedness of the whole program, and such failures are the only ones that have the possibility of leading to restructuring. As the structuredness of a person's knowledge in a particular domain relates to his/her epistemic values, and whether he/she believes in the existence of generalization descriptions, theories for the domain, these factors also affect the possibility of significant learning taking place.

Another important observation from the present research is that there is a general tendency to incorporate school science knowledge into existing knowledge structures. Evidence of this occurring can be found in the program of every student in the older group. However, most of the incorporation was done at the syntactic level - incorporating technical terms into the existing knowledge structure. School science instruction may possibly lead to serious cognitive conflict if the intuitive conceptions of the learner form a well structured schema. However, there is evidence that formal instruction in Physics does play a critical role in promoting cognitive changes: It was observed that the schema tuning incidents were mostly stimulated by information gathered by students during lessons or while watching TV programs. The most serious cognitive conflict came about in the case of Evon when she tried to incorporate two incompatible theoretical frameworks: the Newtonian framework of internal and external forces and her intuitive framework of differing sources of power for autonomous and non-autonomous motion.

9.5 Expert System Development as a Knowledge Elicitation Technique

This final chapter cannot evade the responsibility of discussing issues related to the novel technique used in this research - expert system development as a knowledge elicitation technique. The rationale for developing such a method was that it provides a non-threatening environment for elicitation, the externalized knowledge is accessible to both subject and researcher alike and is explorable and modifiable. Such an elicitation environment thus supports reflection, exploration of consequences of the knowledge and provides an objective record of all the transactions, including records of how conflicts are resolved.

Results from the present research provide encouraging evidence that such a technique is valuable and has potential for further development. Analysis at rule level revealed a prominent concern with holistic descriptions for motion and causal attributions that are consistent with results of the classification task. Many of the elicited ideas are also consistent with what has been reported by others working in this area. For example,the ideas of differentiating between autonomous and non-autonomous motion, and that animate and inamimate causal agents constitute a major distinction in conceptualizing motion are similar to those reported by Bliss, Ogborn & Whitelock (1989); the idea that the causal agent acts as the source of force for the affected is also suggested by Andersson (1985). These give one confidence that the elicited knowledge is not just an artifact of the programming task but that it does reveal important aspects of students' commonsense conceptions about motion.

In addition to yielding elicitations consistent with other methods, it has provided further insight into aspects of students' commonsense conceptions about motion as well as more general aspects of cognition which have not been exposed by other methods reported in this area.

First of all, the fact that this task allows students to create over an extended period of time a representation of their own conceptions without prior conceptual or situational constraint, and that such representation is meant to operate as a system for problem solving allows one to probe much deeper into the structural aspects of the knowledge schema. The different agency gestalts elicited from the students have revealed much finer distinctions in the students' notions of causality than has been reported (e.g. Andersson, 1986). Such distinctions may indicate a possible progression from the intuitive conceptions to a more scientific model.

Another special feature of the present method is that it demands that the subjects take up a rational and reflective attitude in their attempts to build an explicit and generally applicable system about motion. This contrasts greatly with the spontaneity with which responses are elicited in most other methods of elicitation. The completion of such a task requires the subjects to define their own explanation goals, that is, what knowledge they consider to be important for "knowing" about the domain under discussion. This method thus allows the researcher to probe into a completely different aspect of a person's knowledge structure: the epistemic plane of operation represented by the elicited knowledge, including the level of generality the knowledge structure is working at and how well it operates as an orchestrated system. This aspect of a person's organization of commonsense knowledge is not accessible by other means of knowledge elicitation and thus has not been reported. However, this aspect does represent a real and important dimension of knowledge if one really believes in knowledge being organized.

The explorability and modifiability of the programs allow one to study the kinds of conflict that may arise when a person's conceptions about a particular domain are tested against physical observations and what strategies s/he may use to resolve them. The program modification behaviour of students thus provided a chance for us to look into a much discussed issue of promoting learning through the setting up of conflicts (Hewson, 1981, 1984). Analysis of results from the present research suggests that very often conflict resolution does not lead to learning, and in the cases when it does, it is still mainly low level learning as accretion. Even learning as schema tuning rarely happens. The results further suggest that generally strategies leading to minimum structural changes would be adopted for conflict resolution, and only when the existing knowledge structure is well organized (at the structural epistemic level) would there be a possibility for the seriousness of the conflict to escalate to the extent of challenging the validity of the core hypothesis, leading to a possibility of restructuring. Such observations with regard to learning are consistent with the theoretical predictions of learning theories as proposed by Rumelhart and Norman (1978) and to some extent that of Piaget (1975). Such an outcome is very pleasing and does give one some confidence that the method has provided some important insights into learning.

9.6 Beyond Elicitation and Understanding

A better understanding of students' commonsense reasoning about motion should have pedagogical implications. At the content level, results from the present research suggest that relating the Newtonian framework more explicitly with the pragmatic concerns of commonsense reasoning might be a way to promote better understanding of the Newtonian framework, in particular, to address the issue of agency and trying to highlight the role of interacting events in the causal action as responsible for the production of the interacting forces. This might lead to a gradual diminution of the role differences attached to the agent and the affected in motion events. Such an effect might also bring about less differences being attached to animate and inanimate objects, or at least clarify the essential differences between them.

At a broader level, an adoption of the Newtonian frame implies a conceptual revolution. The superiority of the Newtonian framework cannot be appreciated within the context of the pragmatic need for handling specific stereotypical motion which is the predominant case in commonsense reasoning about motion. Unless a student can detach himself/herself from the pragmatic immediacy of everyday motions and cultivate an aesthetic, intellectual perspective in looking at the world, real learning cannot take place. A person needs to operate at the level of the structural epistemic frame before the adoption of a Newtonian framework can take place.

BIBLIOGRAPHY

Andersson, B. (1986). The Experiential Gestalt of Causation: a Common Core to Pupils' Preconceptions in Science. *E.J.S.E.*, 8(2), 155-171.

Ausubel, D. P. (1968). *Educational Psychology: A Cognitive View*, New York: Holt, Rinehart & Winston.

Bell, B., Brook, A., & Driver, R. (1985). An Approach to the Documentation of Alternative Conceptions in School Students' Written Responses. *Briti sh Educational Research Journal*, 11(3), 201-213.

Berger, P. & Luckmann, T. (1966). *The Social Construction of Reality*. New York: Viking Penguin Inc.

Black, M. (1962). Models and Archetypes. In M. Black, *Models and Metaphors*. N. J. : Cornell University Press.

Bliss, J. (1986). *The Relevance of Piaget to Research into Children's Conceptions*. London: Centre for Studies, KQC.

Bliss, J., Monk, M., & Ogborn, J. (1983). *Qualitative Data Analysis for Educational Research: A Guide to Users of Systematic Networks*. Croom Helm, London.

Bliss, J., Morrison, I. & Ogborn, J. (1988). A Longitudinal Study of Dynamics Concepts. *International Journal of Science Education*, 10(1), 99-110.

Bliss, J., Ogborn, J. & Whitelock, D. (1989). Secondary School Pupils' Commonsense Theories of Motion. *International Journal of Science Education*, 11(3), 261-272.

Briggs, J. (1984). *Micro-PROLOG Rules!* PEG-Exeter School of Education, University of Exeter.

Buchanan, B.G. & Shortliffe, E.H. (1984). *Rule-Based Expert Programs: the MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison-Wesley.

Caramazza, A., McCloskey, M., & Green, B. (1981). Naive Beliefs in 'Sophisticated' Subjects: Misconceptions about Traject ories of Objects. *Cognition*, 9, 117. c/ r/

Cercone, N. (1975). *Representing Natural Language in Extended Semantic Networks*, Technical Report TR75-11, Department of Computing Science, Edmonton: University of Alberta.

Champagne, A.B., Gunstone, R.F. & Klopfer L.E. (1985). Instructional Consequences of Students' Knowledge about Physical Phenomena. In West, L. & Pines A. (Eds.), *Cognitive Structure and Conceptual Change*. Orlando, Florida: Academic Press Inc.

Chomsky, N. (1965). *Aspects of the theory of Syntax*, Cambridge,Mass: MIT Press.

Clark, K.L. & McCabe, F.G. (1984). *Micro-PROLOG: Programming in Logic*. New Jersey: Prentice-Hall International, Inc.

Claxton, G.L. (1985). Teaching and Acquiring Scientific Knowledge. In Keen T. & Pope M. (Eds), *Kelly in the Classroom: Practical Applications of Personal Construct Psychology.* Montreal: Cybersystems.

Clement, J. (1982). Students' Preconceptions in Introductory Mechanics. *American Journal of Physics*, 50(1), 66-71.

Collins, A. M., & Quillian, M. R. (1970). Does Categoric Size affect Categrization Time? *Journal of Verbal Learning and Verbal Behaviour*, 9, 432-438.

Confrey, J. (1987). `Misconceptions' Across Subject Matters: Science, Mathematics and Programming. In Novak J.D. (Ed.). *Proceedings of the 2nd International Seminar on `Misconceptions and Educational Strategies in Science and Mathematics'*, July 26-29, Ithaca.

Cosgrove, M. & Osborne, R. (1985). Lesson Frameworks for Changing Children's Ideas. In Osborne, R. & Freyberg, P. (Eds), *Learning in Science.* Auckland: Heinemann.

Dean, J. (1986). Prolog: a Language for Cognition. *Proceedings of the First PEG Conference.* University of Exeter.

Denney, D.R. & Moulton, P.A. (1976). Conceptual Preferences among Preschool children. *Developmental Psychology*, 12, 509-513.

DiSessa, A. A. (1982). Unlearning Aristotelian Physics: A Study of Knowledge-based Learning. *Cognitive Science*, 6, 37-75.

DiSessa, A.A. (1983). *Phenomenology and the Evolution of Intuition.* In Stevens, A.L. & Gentner, D. (Eds), *Mental Models.* London: Lawrence Erlbaum Associates.

DiSessa, A. A. (1984). *Learning about Knowing*,Technical Report, MIT Laboratory for Computer Science.

DiSessa, A.A. (1985). *Final Report on Intuition as Knowledge.* Report prepared for the Spencer Foundation, MIT Laboratory for Computer Science.

Doran, R.L. (1972). Misconceptions of selected science concepts held by elementary school students. *Journal of Research in Science Teaching*, 9(2), 127-137.

Driver, R. (1973). *The Representation of Conceptual Framework in Young Adolescent Science Students.* (Unpublished PhD Thesis). Illinois: University of Illinois.

Driver, R. & Easley, J. (1978). Pupils and Paradigms: a Review of Literature Related to Concept Development in Adolescent Science Students. *Studies in Science Education*, 5, 61-84.

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. (1983). *The pupil as Scientist*, Milton Keynes: Open University Press.

Driver, R., Guesne, E., & Tiberghien, A. (Eds) (1985). *Children's Ideas in Science*, Milton Keynes: Open University Press.

Duhem, P. (1954). *The Aim and Structure of Physical Theory*. Trans. by Wiener, P.P. Princeton: Princeton University Press.

Duit, R. (1987). Research on Students' Alternative Frameworks in Science -Topics, Theoretical Frameworks, Consequences for Science Teaching. In Novak J.D. (Ed.). *Proceedings of the 2nd International Seminar on 'Misconceptions and Educational Strate-gies in Science and Mathematics'*, July 26-29, Ithaca.

Duncan, I.M. and Johnstone, A.H. (1973). The mole concept. *Education in Chemistry*, 10, 213-214.

Einstein, A., & Infeld, (1938). *The Evolution of Physics*. Cambridge: Cambridge University Press.

Engel, M.E.T. (1982). *The Development of Understanding of Selected Aspects of Pressure, Heat and Evolution in Pupils Aged between 12 and 16 years*. Unpublished Ph.D. Thesis. Leeds: University of Leeds.

Feigenbaum, E.A., Buchanan, B.G. & Leiderberg, J. (1971). On Generality and Problem Solving: a Case Study using the DENDRAL Program. *Machine Intelligence*, 6, 1971.

Feyerabend, P. (1978). *Science in a Free Society*. London: New Left Books.

Fillmore, C.J. (1968). The Case for Case. In E. Bach, & R.T. Harms (Eds.), *Universals in Linguistic Theory*. N. Y.: Holt, Rinehart & Winston.

Fivush R. (1987). Scripts and categories: interrelationships in development. In Neisser U. (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 234-254). Cambridge: C.U.P.

Fransella, F. & Bannister, D. (1977). *A Manual for Repertory Grid Technique*. Academic Press.

Furth, H.G. (1981). *Piaget and Knowledge,* 2nd edition. Chicago: The University of Chicago Press.

Gagne, R.M., & White, R.T. (1978). Memory Structures and Learning Outcomes. *Review of Educational Research*, 48, 187-272.

Garcia, R. (1983). Psychogenesis and the History of Science. In *Cahier de la Fondation Archives Jean Piaget No. 4*. Geneva: Fondation Archives Jean Piaget Publication.

Garcia, R. (1987). Sociology of Science and Sociogenesis of Knowledge. In Inhelder, B., de Caprona, D. & Cornu-Wells, A. (Eds.), *Piaget Today*. Hillsdale: Lawrence Erlbaum Associates.

Gentner, D., & Gentner, D.R. (1983). Flowing Waters or Teening Crowds: Mental Models of Electricity. In D. Gentner & A.L. Stevens (Eds.), *Mental Models*. Lawrence Erlbaum Associates, London.

Gibbs, G. (1982). A Review of the Research of Ference Marton and the Gotegorg Group: A Phenomenological Research Perspective on Learning. In *Higher Education*, 11, 123-145.

Gibbson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gilbert, J.K., Watts, D.M. & Osborne, R.J. (1981). *Eliciting Students' Views Using an Interview-About-Instances Technique*. Paper presented at AERA conference, Los Angeles, April.

Gilbert, J.K. & Pope, M.L. (1982). *School Children Discussing Energy*. (Mimeograph). University of Surrey, Institute of Educational Development.

Gilbert, J.K., Osborne, R., & Fensham, R. (1982). Children's Science and Its Consequences for Teaching. *Science Education*, 66(4), 623-633.

Gilbert, J.K., & Watts, D.M. (1983). Concepts, Misconceptions and Alternative Conceptions: Changing Perspectives in Science Education. *Studies in Science Education*, 10, 61-98.

Gilbert, J.K., & Zylbersztajn, A. (1985). A Conceptual Framework for Science Education: The Case Study of Force and Movement. *E.J.S.E.*, 7(2), 107-120.

Gilbert, J.K., Watts, D.M. & Osborne, R.J. (1985). Eliciting Student Views Using an Interview-About-Instances Technique. In West, L. & Pines, A. (Eds), *Cognitive Structure and Conceptual Change*. Orlando, Florida: Academic Press.

Giordan, A. (1987). Bibliographie concernant les recherches sur les conceptions des apprenants en biologie. *Annales De Didactiques Des Sciences*, no. 2, 1987.

Guidoni, P. (1985). On Natural Thinking. *E.J.S.E.*, 7(2), 133-140.

Gunstone, R. & White, R. (1981). Understanding of Gravity. *Science Education*, 65(3), 291-299.

Gunstone, R., & Watts, M. (1985). Force and Motion. In R. Driver et al (Eds.), *Children's Ideas in Science*. Milton Keynes: Open University Press.

Hamlyn, D. W. (1978). *Experience and the Growth of Understanding*. London: Routldge & Kegan Paul.

Hammond, P. & Sergot, M. (1984). *APES: Augmented Prolog for Expert Systems, Reference Manual*. Surrey: Logic Based Systems Ltd.

Hayes, P. (1979). The Naive Physics Manifesto. In D. Mitchie (Ed.), *Expert Systems in the Microelectronic Age*. Edinburgh University Pres, Edinburgh.

Heidegger, M. (1962). *Being and Time*. Translated by Macquarrie, J. & Robinson, E., 1962. New York: Harper & Row.

Helm, H. (1980). Misconceptions in Physics Amongst South African Students. *Physics Education*, 15, 92-105.

Hewson, P. (1981). A Conceptual Change Approach to Learning Science. *E.J.S.E.*, 3(4), 383-396.

Hewson, P.W. (1980). *A Case Study of the Effect of Metaphysical Commitments on the Learning of a Complex Scientific Theory*. Paper presented at Australian Educational Research Association, April.

Hewson, P.W. (1984). The Role of Conceptual Conflict in Conceptual Change and the Design of Science Instruction. *Instructional Science*, 13, 1-13.

Husserl, E. (1913), *Logical Investigations*.

Jagger, J.M. (1985). A Review of Research into the Learning of Mechanics. In Orton, A. (Ed.), *Studies in Mechanics Learning*. Leeds: The University of Leeds, Centre for Studies in Science and Mathematics Education.

Johnson, N. (1983). *Elicitation and Representation of Children's Arithmetic Knowledge*. Unpublished Ph.D. Thesis, University of London.

Keil, F.C. & Batterman, N. (1984). A Characteristic-to-Defining shift in the Development of Word Meaning. *Journal of Verbal Learning and Verbal Behaviour*, 23, 221-236.

Keil, F.C. (1986). The acquisition of natural kind and artifact terms. In W. Demopolous & A. Marras (Eds.), *Language learning and concept acquisition*. Norwood, NJ: Ablex.

Kelly, G. A. (1955). *The Psychology of Personal Constructs Vol 1 & 2*. New York: W. W. Norton.

Kuhn, T.S. (1970). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lakatos, I. (1970). Falsification and the Methodology of Science Research Programs. In Lakatos & Musgrave (Eds), *Criticism and the Growth of Knowledge*. Cambridge: The University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.

Langford, J.M. & Zollman, D. (1982). *Conceptions of Dynamics Held by Elementary and High School Students*. Paper presented at the annual meeting of the American Association of Physics Teachers, San Francisco, January.

Larkin, J.H. (1983). The Role of Problem Representation in Physics. In D. Gentner & A.L. Stevens (Eds.), *Mental Models*. London: Erlbaum.

Larkin, J.H., & Rainard, B. (1984). A Research Methodology for Studying How People Think. *Journal of Research in Science Teaching*, 21(3), 235-254.

Law, N. & Ki, W.W. (1987). *AI Programming Environment as a Knowledge Elicitation and Cognitive Modelling Tool*. Paper presented at the Third International Conference on Artificial Intelligence and Education, University of Pittsburgh, May 7-11.

Law, N., Ogborn, J. & Whitelock, D. (1988). Knowing What the Student Knows: a Use of APES in Science Education. In Nicol, J., Briggs, J. & Dean, J. (Eds), *PROLOG, Children and Students*. London: Kogan Page.

Law, N. (1988). Knowledge Structures - Where Can We Find Them? *Educational Research Journal* (Hong Kong), 3, 23-29.

Lawson, R.A. (1984). *Student Understanding of Single Particle Dynamics*. Ph.D. Dissertation, University of Washington.

Lucariello J. & Nelson K. (1985). Slot-Filler Categories as Memory Organizers for Young Children. *Developmental Psychology*, 21(2), 272-282.

Magee, B. (1987). *The Great Philosophers: An Introduction to Western Philosophy.* London: BBC Books.

Mandler, J.M. (1979). Categorical and schematic organization in memory. In Puff C.R. (Ed.), *Memory Organization and Structure* (pp. 259-299). New York: Academic Press.

Marton, F. (1981). Phenomenography - Describing Conceptions of the World Around Us. *Instructional Science*, 10, 177-200.

McCabe, F.G. et al. (1985). *Micro-PROLOG Professional Programmer's Reference Manual.* London: Logic Programming Associates Ltd.

McClelland, J.A.G. (1985). Misconceptions in Mechanics and How to Avoid them. *Physics Education*, 20, 159-162.

McCloskey, M. (1983). Naive Theories of Motion. In D. Gentner & A.L. Stevens (Eds.), *Mental Models.* Lawrence Erlbaum Associates, London.

McCloskey, M., Carmazza, A. & Green, B. (1980). Curvilinear motion in the absence of external forces: naive beliefs about motion of objects. *Science*, 210, 1139-1141.

McDermott, L.C. (1984). An Overview of Research on Conceptual Understanding in Mechanics. *Physics Today*, July, 22-32.

Minsky, M. (1975). A Framework for Representing Knowledge. In Winston, P.H. (Ed.), *The Psychology of Computer Vision.* McGraw-Hill.

Minstrell, J. (1982). Explaining the 'at rest' Condition of an Object. *Physics Teacher*, January.

Nagel, E. (1961). *The Structure of Science.* New York: Harcourt, Brace.

Neisser, U. (1987). From Direct Perception to Conceptual Structure. In Neisser, U. (Ed.), *Concepts and Conceptual Development.* New York: Cambridge University Press.

Nelson K (1983). The derivation of concepts and categories from event representation. In E. Scholnick (Ed.), *New trends in conceptual representation: Challenges to Piaget's theory* (pp. 129-150). Hillsdale, NJ: Erlbaum.

Newell, A. & Simon, H.A. (1972). *Human Problem Solving.* Prentice-Hall.

Newton, I. (1687). *Philosophie Naturalis Principia Mathematica.*

Norman, D. A. (1978). Notes towards a Theory of Complex Learning. In Lesgold et al (Eds.), *Cognitive Psychology and Instruction.* New York: Plenum Press.

Novak, J. (1977). *A Theory of Education.* Ithaca: Cornell University Press.

Novak, J.D., Gowin, D.B. & Johansen G.T. (1983). The Use of Concept Mapping and Knowledge Vee Mapping with Junior High School Science Students. *Science Education*, 67(5), 625-645.

Nussbaum, J., & Novick, S. (1981). Brainstorming in the Classroom to Invent a Model: A Case Study. *School Science Review*, 62, 771-778.

Nussbaum, J. (1985). The Particulate Nature of Matter in the Gaseous Phase. In Driver, R. Guesne, E. & Tiberghien, A. (Eds), *Children's Ideas in Science.* Milton Keynes: Open University Press.

Ogborn, J. M. (1985). Understanding Students' Understanding: An Example from Dynamics. *European Journal of Science Education*, 7(2), 141-150.

Ogborn, J. (1987). Newton and Common Sense. *Physics Education*, 22, 321.

Ogborn, J. (1989). *Primitive Structures of Commonsense Reasoning and the Understand-ing of Science.* in Le Transfert des Connaissances en Sceinces et Techniques, Actes du Colloque Athena 89, Universite Montpellier II, 4-7 Sept.

Oppenheimer, R. (1954). *Science and the Common Understanding.* New York: Simon and Schuster.

Osborne, R.J. (1980). *Force.* Learning in Science Project. Working Paper No.16. Hamilton (NZ): University of Waitkato.

Osborne, R. & Gilbert, J. (1980). A method for investigation of concept understanding in science. *European Journal of Science Education*, 2(3), 311-321.

Osborne, R., & Freyberg, P. (1985). *Learning in Science: The Implications of Children's Science.* Auckland: Heinemann.

Petrie, H. G. (1979). *Metaphor and Thought.* Cambridge: The Cambridge University Press.

Pfundt, H. & Duit, R. (1987). *Bibliography: Students' Alternative Frameworks and Science Education* (2nd ed.)

Piaget, J. (1954). *The Construction of Reality in the Child.* New York: Basic Books Inc.

Piaget, J. (1970a). *The Child's Conception of Movement and Speed.* New York: Ballantyne Books.

Piaget, J. (1970b). *The Child's Conception of Time.* New York: Ballantyne Books.

Piaget, J. (1970c). *Genetic Epistemology.* N. Y. : Columbia University Press.

Piaget, J. (1974). *Understanding Causality.* Toronto: G.J. McLeod Ltd.

Piaget, J. (1975). *L'equilibration des structures cognitives.* Paris: Presses Universitaires de France.

Piattelli-Palmarini, M. (1979). *Language and Learning-The debate between Jean Piaget and Noam Chomsky.* London: Routledge & Kegan Paul.

Pines, A.L. (1985). Towards a Taxonomy of Conceptual Relations and the Implications for the Evaluation of Cognitive Structures. In West, L.H.T. & Pines, A.L. (Eds), *Cognitive Structure and Conceptual Change.* Orlando, Florida: Academic Press.

Polanyi, M. (1958). *Personal Knowledge, Towards a Post-Critical Philosophy.* Chicago: University of Chicago Press.

Popper, K.R. (1972). *Objective Knowledge.* Oxford: Oxford University Press.

Quillian, M.R. (1967). Word Concepts: A Theory and Simulation of some Basic Semantic Capabilities. *Behavioral Science,* 12(5), 410-430.

Quillian, M.R. (1968). Semantic Memory. In M. Minsky (Ed.), *Semantic Information Processing.* Cambridge, Mass: MIT Press.

Rich, E. (1983). *Artificial Intelligence.* Singapore: McGraw-Hill Inc.

Rosch, E. (1978). Principles of Categorization. In E. Rosch and B.B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27-48). Hillsdale, N.J.: Erlbaum.

Rousseau, J.J. (1760). *Emile.*

Rowell, J.A. & Dawson, C.J. (1983). Laboratory counter examples and the growth of understanding in science. *European Journal of Science Education,* 5(2), 203-216.

Rumelhart, D.E. & Ortony, A. (1977). The Representation of Knowledge in Memory. In Anderson, R.C., Spiro, R.J. & Montague, W.E. (Eds), *Schooling and the Acquisition of Knowledge.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Rumelhart, D. E., & Norman, D. A. (1978). Accretion, Tuning and Restructuring: Three Modes of Learning. In J. W. Cotton & R. Klatzky (Eds.), *Semantic Factors in Cognition.* New Jersey: Lawrence & Erlbaum Associates.

Rumelhart, D.E., & Norman, D.A. (1985). Representation of Knowledge. In A.M. Aitkenhead, & J. M. Slack (Eds.), *Issues in Cognitive Modelling.* London: Erlbaum.

Saltiel, E., & Viennot, L. (1985). What do we learn from similarities between Historical Ideas and the Spontaneous Reasoning of Students? In P. L. Lijnse (Ed.), *Conference on Physics Education.* Utrecht.

Schank, R.C. (1975). *Conceptual Information Processing.* New York: North Holland.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, Plans , Goals and Understanding.* Hillsdale, New Jersey: Erlbaum.

Schank, R.C. (1986). *Explanation Patterns: Understanding Mechanically and Creatively.* N.J.: Lawrence Erlbaum Associates.

Schutz, A., & Luckman, T. (1973). *Structure of the Life World.* London: Heinemann.

Shavelson, R.J. (1971). *Some Aspects of the Relationship between Content Structure and Cognitive Structure in Physics Instruction.* Doctoral Dissertation, University Microfilms No. 71-19, 759, Stanford University.

Shavelson, R.J. (1974). *Methods for Examining Representations of a Subject Matter Structure in a Student's Memory.* Journal of Research in Science Teaching, 11(3), 231-249.

Sjoberg, S. and Lie, S. (1981). *Ideas about force and movement among Norwegian pupils and students*, (mimeograph). Oslo: University of Oslo, Centre for School Science.

Solomon, J., Black P., Oldham, V., & Stuart, H. (1985). The Pupils' View of Electricity. *European Journal of Science Education*, 7(3), 281-294.

Stewart, J. (1979). Content and Cognitive Structure: Critique of Assessment and Representation Techniques Used by Science Education Researchers. *Science Education*, 63(3), 395-405.

Stewart, J. (1980). Techniques for Assessing and Representing Information in Cognitive Structure. *Science Education*, 64(2), 223-235.

Stillings, N.A., Feinstein, M.H. & others (1987). *Cognitive Science: An Introduction.* Cambridge, Massachusetts: The MIT Press.

Sutton, C.R. (1980). The Learner's Prior Knowledge: A Critical Review of Techniques of Probing its Organization. *E.J.S.E.*, 2, 107-120.

Taylor, J., & Du Bonlay, B. (1985). *Why Novices May find Programming in Prolog Hand.* Cognitive Studies Programme, University of Sussex.

Trowbridge, D., & McDermott, L.C. (1980). Investigation of Student Understanding of the Concept of Velocity in one dimension. *American Journal of Physics*, 48, 1020.

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.

Viennot, L. (1977). *Le Raisonnement Spontane en Dynamique elementaire.* Paris: The University of Paris VII.

Viennot, L. (1979). Spontaneous Reasoning in Elementary Dynamics. *European Journal of Science Education*, 1(2), 205-221.

Viennot, L. (1985). Analysing students' Reasoning in Science: a Pragmatic View of Theoretical Problems. *European Journal of Science Education*, &(2), 151-162.

Wartofsky, M.W. (1968). *Conceptual Foundations of Scientific Thought.* New York: Macmillan.

Watts, D.M., & Zylbersgtajn, A. (1981). A Survey of Some Children's Ideas about Force. *Physics Education,* 16(6), 360-365.

West, L.H.T. & Pines, A.L. (1985). *Cognitive Structure and Conceptual Change.* Orlando, Florida: Academic Press.

Whitaker, R.J. (1983). Aristotle is not Dead: Student Understanding of Trajectory Motion. *American Journal of Physics*, 51, 352.

White, B.Y. (1984). Designing Computer Games to Help Physics Students Understand Newton's Laws of Motion. *Cognition and Instruction*, 1, 69.

White, R.T. (1985). Interview Protocols and Dimensions of Cognitive Structure. In West, L.H.T. & Pines, A.L. (Eds), *Cognitive Structure and Conceptual Change.* Orlando, Florida: Academic Press.

Winograd, T. (1975). Frame Representations and the declarative Procedural Controversy. In D. Bobrow & A.Collins (Eds.), *Representation and Understanding*. New York: Academic Press.

Winston, P.H. (1984). *Artificial Intelligence*, 2nd edition. Reading, Massachusetts: Addison-Wesley.

Appendix 1

A.I. PROGRAMMING ENVIRONMENT AS A KNOWLEDGE ELICITATION & COGNITIVE MODELLING TOOL

Paper presented at the Third International Conference on Artificial Intelligence and Education, University of Pittsburgh, May 7-11, 1987.

Nancy Law & Wing Wah Ki
Dept. of Professional Studies in Education
University of Hong Kong

This study attempts to elicit students' commonsense ideas about motion and how they structure their learnt physics concepts in relation to their commonsense ideas through Prolog programming. Prolog is offered as an expressive tool to students to externalize and formalize their intuitive ideas about motion. The main aim of this investigation is to provide deeper insight into the problem of students' misconceptions/alternative frameworks in mechanics (see Law N. 1986).

The important effect of pre-instructional ideas on children's learning is now widely recognised and a lot of effort has gone into investigating what these ideas are like in various domain areas in science in the past few years. In the domain of mechanics, we now have accumulated a very comprehensive catalog of phenomenological descriptions of various aspects of children's reasoning about forces and motion, without much understanding of what lies underneath such phenomenological manifestations of children's thinking. There is a need for us to really take children's ideas seriously, in their own right and not as a distortion of the scientific view. We need to find out a child's basic cognitive framework and mechanisms of thought when thinking about a particular domain area before we can begin to find a better way of teaching.

In this study, the student plays the role of the teacher and the computer that of a learner. The idea is that in the course of teaching the computer, one is obliged to externalize one's ideas. And once these ideas and intuition have been externalized in the form of a computer program, they become accessible to verification, reflection, and modification. In this way the computer serves as a 'mudpie' for the learner to develop his/her own understanding. Furthermore, the resulting programs and the interaction process of the student with his/her own externalized ideas provide rich material for an understanding of the mental models and processes of the student.

One important pre-requisite for such a scheme of work to be viable is the availability of a programming environment that is amenable to use and effective in representing human knowledge.

Prolog, one of the A.I. programming languages that is gaining more and more attention, is chosen for this research for several reasons:

(1) It provides a declarative programming environment and the user only needs to declare the relevant facts and inference rules. There is an in-built inference mechanism so that the knowledge base can be put to work without extra programming effort.

(2) It excels conventional languages in its ability to handle symbolic and non-numerical relationships.

(3) The PROLOG programming environment is highly interactive. The man-machine interaction can be a continuous and inter-leaving sequence of knowledge input and queries. When used with a front-end extension that can provide a query-the-user facility, the knowledge base can increase incrementally even when the user is using the program to answer queries (Law N., Ogborn J. & Whitelock D., 1986).

Methodology

The main task for the participating students was to build an expert system of their own understanding of some aspects of motion in micro-Prolog (McCabe F.G. et.al., 1985) using the front-end APES (Hammond P. & Sergot M, 1984). Each student participate in the design and actual construction of a computational model of their own thinking, assess the model built (which is the resulting expert system) and then modify it as they think fit. A representation of the student's knowledge in the form of an expert system has the advantages of being explicit, explorable, and capable of offering explanations for deduction paths (Law N, Ogborn J. & Whitelock D., 1986). This involvement of the students in the programming process forces the students to externalize and formalize their thoughts, a lot of which may have never been raised to a conscious level. It is thus hoped that this method offers an effective tool for probing deeper into students' knowledge structures. Furthermore, as an externalized entity, the resulting program can be explored, and so offers a good opportunity to find out the possible outcomes of confronting students with explicit representations of their own knowledge.

So far, eight sixth form Physics students have participated in this study. Each student is treated as a deep case study on its own, involving eight programming sessions, each of one and a half hour duration. The researcher is with the student throughout the programming sessions for data collections purposes and as a programming adviser to the student. The data collected include the developed programs, as well as screen dumps and diary notes taken during the programming sessions. As the aim of the research is not to teach programming but to elicit the students' knowledge structures, efforts were made to reduce the students' efforts spent on mastering the Prolog language. Only a restricted subset of the language was introduced, and they were only required to write programs in the infix format.

From the preliminary results of this study, several interesting observations can be made.

Some Observations on the cognitive processes of the students

a. Mental models anchoring on one stereotypical example -

   Although the task set to the students was to write up rules about some aspects of motion such that the rules can be applied in general to all situations, most of the students start with one specific instance in mind and write up a rule that they think can be satisfied by that particular instance.

   For example, a student made up the following rule about how to decide that something is experiencing gravity:

```
_object experience gravity if
     _object travelling parabolically
```

   Evidently the student was only thinking about the cases when an object is thrown upwards at an angle, moving in a parabola befor hitting the ground. But this is only one set of instances when the motion is affected by gravity - what about the cases when the object is thrown vertically upwards or just moving downwards vertically?

b. General statements used for a *restricted* and *fuzzy* domain of reference

   This peculiarity is connected to the imprecise nature of our everyday language. Our language would be very dull indeed if all its usage is clearly defined without any ambiguity at all. Yet when the students are required to formalize their thoughts into generally applicable rules as they are required in this programming exercise, they operate in the same mode of imprecision. As mentioned above, the students normally think in terms of a restricted domain of instances. At the same time, the boundaries of the domain is not clear even to themselves.

   Again using the same example of the student writing the rule quoted above, her use of the word 'experience' actually meant 'during the course of motion of the object, it was being affected by'. When the research asked her to apply the rule to the case of a tape recorder placed on the table, the result was of course that the tape recorder does not experience gravity. Because she had her own particular definition of experience at that time, she did not think that the response from the computer had any problem until I asked her specifically whether for the tape recorder on the table, it does experience any gravity or not. If she had a clear focus for her statement, she could have pointed out that the word 'experience' is not used to encompass all the meanings that this word would carry in everyday language. Instead, after some thinking, she agreed to make an amendment by adding the following rule:

```
_object experience gravity if
     _object has weight and
     _object is still
```

Again this additional rule reflects that she is considering only the cases of still objects like the tape-recorders. When she tested the rules with the case of the kite which is moving and yet not parabolically, the response from the computer was that the kite does not experience gravity which is counter to the new meaning of the word 'experience'. At that point, she very relunctantly edited the last rule above to :

```
_object experience gravity if
     _object has weight
```

c. Patching up of the model when it doesnot fit expectations rather than reconsidering the structure of the whole model -

It is inevitable that hypotheses set up may be disproved when put to the test. There are different courses of action one may take when faced with such failures. One prominent feature that came up during the course of this investigation is that most of the students tend to stick on to the basic structure of their first hypothesis and try to make adjustments to it to fit the specific cases rather than re-assessing the validity of the structure of the whole hypothesis. The following is one typical example from the case studies:

A student was writing up rules to decide what causes the motion of an object. Her core hypothesis was that there is a '*source of supply of force*' for every action, and such sources can be divided into *internal* and *external* ones. She believed that the nature of such 'sources of supply of force' is a deciding factor in determining the cause(s) and source of control for a motion. A bird flies and a motor car runs because of an internal supply of force. A basketball starts moving because of an external supply of force from the hand. Yet there are some forces, like the weight of an object, which she was not so sure about the categorization.

She wrote the following rules for determining the cause of motion:

```
_object motion-caused-by itself if
     _object force-supplied-by _object
_object motion-caused-by machine if
     _object force-supplied-by machine
_object1 motion-caused-by _object2 if
     _object1 force-supplied-by _object2
```

```
_object motion-caused-by gravity if
     not(_object under-the-influence-of other-external-force)
```

She then set out to test the program with the case of an apple (thinking about the famous 'Newton's apple). To her dismay, the computer did not came out with any answer at all. The reason was that in one of her earlier sessions, she included weight as an external (supply of) force, together with forces like friction, air-current, etc. Faced with such a contradiction, her immediate response was to exclude weight from her list of external forces. After making this modification, the computer still did not give her the expected answer of gravity - According to one of her other rules, anything placed in air would be affected by air-current, and air-current is an external force. Now, excluding air-current from the list of external forces seems to be too much of a twisting of facts, so she modified her fourth rule above to :

```
_object motion-caused-by gravity if
     not (_object motioncausedby _something)
```

This alteration did succeed in giving her the expected answer for the apple because this set of rules on cause of motion does not involve anything concerning whether a force is an external or internal one.

Her next project was to write some general rules to determine what controls a particular motion. She began with the rules:

```
_object controlled-by itself if
     _object force-supplied-during-motion _something and
     _something force-type internal and
     not(_object under-the-influence-of other-external-force)
_object controlled-by _something if
     _object force-supplied-during-motion _something and
     _something force-type external and
     _object affected-by _something
```

These two lines clearly show that the student was still building her model up on her core hypothesis of internal and external supplies of force .

She began testing the rules with the case of a falling stone. This time, she considered weight to be an external force, and so the answer from the computer was acceptable - that the falling stone was controlled by an external force. Problem arose when she tested the same rules with the case of a bird which is flying. Instead of confirming that that 'bird motion-controlled-by itself'as she would have expected, the answer was negative because according to the data input the bird is affected-by weight which is an external force. Again, her response was to change weight from 'force-type external' to 'force-type internal'.

This tactic of patching up of the rules by modifying a fact is very frequently adopted, rather than changing the structure of the core hypothesis.

d. Change in attitude towards cognitive conflict -

The students participating in this study differ not only in their general ability, but also in their attitudes towards cognitive conflict. Some students are very reflective and sometimes, even before they finish typing in a rule, they would have started doubting or changing what they have typed in. But more often, the students tend to be very defensive, and does not like to see their rules fail. Prolog programming seem to provide an environment that promotes a change in attitude in the latter group of students, according to the case studies in this research.

A significant change in attitude was observed in the student who was the most defensive and least reflective. At first she felt very annoyed when a rule does not work out, and she disliked the examples suggested by the researcher for testing the program because she felt that they were chosen to fail her rules. An incident occurred during her fourth programming session that changed her entire attitude. Amongst her rules on control of motion was the follwing :

```
_object motion-controlled-by engine if
     _object has-ability-to-move continuously-without-
          getting-tired
```

She was rather pleased that the output from the program was reasonably acceptable after several testings. The she decided to test the program with the case of a falling leaf - and asked for the source of control for the leaf. To her surprise and dismay, the answer came out that the falling leaf was controlled by an engine. This outrageous result shattered her so much that she asked for the screen dump of that session to study at home and left very depressed. The next session she came back armed with a whole lot of amendments and additions. After this incident she became much more interested in the project. Moreover, she becomes much more adventurous in suggesting various novel situations to test her program, ones that she evidently had not thought about when she made up her rules initially.

Some Initial Evaluation on the usefulness of the Prolog Programming Environment

A preliminary evaluation of the results from the present study shows that Prolog is valuable as a knowledge elicitation tool. It allows us to uncover the core hypotheses and fundamental assumptions held by students. The expert systems that the students write tell us about the main considerations they would make concerning a specific aspect of the domain they are considering. But to find out which beliefs are held most strongly, which ideas are acting as the structural organizer of their knowledge, we have to observe the students' behaviour when their programs fail. Another kind of information we can elicit in such a process is the personal cognitive style of the students. We can find out which students are the more reflective ones, and also how each differs in their attitude towards cognitive conflict.

Knowing what the student believes in and finding about about his/her cognitive styles is of course a valuable information to a teacher. But can the Prolog programming environment offer anything towards promoting learning? In one of the case studies quoted above, the student did have a marked change in her attitude towards cognitive conflict which should be valuable for promoting structural changes in her knowledge. Yet this change is not evident in the other case studies. Probably there are other factors besides the Prolog programming environment that is producing such a change. Perhaps further research and analysis may provide more information on this aspect.

Another question we may like to ask about such a study is whether the cognitive conflicts provided by the Prolog programming environment encourage the students to change to a more scientific knowledge structure. Unfortunately, the answer from this study is negative. There are too many alternatives changes that the students can make, patching or otherwise, and the Prolog environment per se can tell nothing more than the internal logical consistency of the input statements. A set of more logically consistent statements does not necessarily possess greater physical validity. It seems that suitable intervention on the basis of the uncovered core hypotheses would be necessary. But what form should such intervention take and whether Prolog programming would be of value at this stage are questions that warrant further research and investigation.

Reference:

Hammond P. & Sergot M. (1984), APES: Augmented Prolog for Expert Systems, Reference Manual, Logic Based Systems Ltd., Surrey, England.

Law N. (1986), *Students' Creation and Interactions with Computational Representations of their own Knowledge Structure*, a paper presented at the Sixth Conference on Computers in Higher Education, April 15-17, 1986, Lancaster, England.

Law N., Ogborn J. & Whitelock D. (1986), *Knowing what the Student Knows: a Use of APES in Science Education* in Prolog Educational Group Newsletter, 1(1), Exeter, England.

McCabe F.G. et al. (1985), Micro-PROLOG Professional Programmer's Reference Manual, Logic Programming Associates Ltd., London, England.

Appendix 2

13. Knowing what the student knows: a use of APES in science education

Nancy Law, Jon Ogborn and Denise Whitelock
Dept of Science Education
University of London Institute of Education
Bedford Way
London WR1H 0AL
UK

A PROLOG program contains a description of a 'world', and running the program generates consequences which follow from the facts and rules in that description. Consider what might be told to a PROLOG program in the case of one thing supporting another: for example, a book resting on top of a table. Some simple rules about the general notion of support might be:

1. An object which is supported must rest on another object.
2. If things are piled on one another, the lower ones support those above them.
3. All supports require sufficient strength to sustain that which they support.

We can capture this 'real world' picture in a small PROLOG program shown below.

```
X rest-on Y if
          X on Y
X rest-on Y if
          Z on Y and
          X rest-on Z
X supports Y if
          Y rest-on X and
          X strong-for Y
X supported if
          Y supports X
```

However, how does the deductive capability of PROLOG manifest itself when employed in the expert system shell APES? The salient features of the APES shell are demonstrated if we engage in a dialogue with our program and ask whether Mary is supported. The program has no knowledge about a person named Mary but has been given a set of rules about support which should be applied in order to deduce an answer to our query. In fact, the system replies with two questions:

1. Is Mary on anything? and
2. Is Mary's support strong enough for her?

With this extra information which is then added to the database, the system can confirm that Mary is supported and will also explain how it reached that conclusion.

```
    which (X : Mary on X)?
          Answer is chair
          Answer is end
  Is it true that chair strong-for Mary? Yes
    = = = Yes, I can confirm Mary supported why
  To deduce Mary supported I used the rule
    X supported if Y supports X
  I can show chair supports Mary stop
  Execution terminated
```

If we add another rule to the database, which elaborates upon our general notion of support and suggests that the floor is the ultimate support since it appears to be very strong and cannot break, and then consider whether Dave who is sitting on a stool holding a pile of books is supported, we are again asked whether the stool is a strong enough support for Dave. Even when the reply is a negative one to this query, the program still confirms that Dave is supported by reasoning that the floor supports Dave, as is shown below.

```
    which (X : Dave on X)?
    Answer is stool
    Answer is end
  Is it true that stool strong-for Dave?
    which (X Y) : X on Y)?
    Answer is (Mary chair)
    Answer is (books Dave)
    Answer is (Dave stool)
    Answer is (stool floor)
    Answer is end
    = = = Yes, I can confirm Dave supported why
  To deduce Dave supported I used the rule
    X supported if Y supports X
  I can show floor supports Dave.
  Execution terminated.
```

This is an answer we would not have expected since we had previously stated that the stool could not support Dave, let alone Dave plus a pile of books! Therefore, our program demonstrated that a common-sense description of support was not accurate enough to convey the subtlety implied by this notion, and revealed flaws in our less-than-rigorous definition. However, the description can now be modified since the interaction with APES has provided clues as to where our descriptions were incomplete or invalid and will alsd quickly provide feedback about the appropriateness and reliability of any new representation which is constructed.

We have demonstrated PROLOG's query-the-user facility employed within APES, which enabled the database to expand and acquire more facts, which can in turn be manipulated by the system's inferencing rules. Explanation features are of particular interest to researchers in the field of education, for reasons discussed in the remainder of this chapter. The dialogue between computer and user can be recorded and noted; APES will also report what it has been told in a given session, as in the example below:

list dialogue
You told me that Mary on chair
You told me that books on Dave
You told me that Dave on stool
You told me that stool on floor
You told me that chair strong-for Mary
You told me that floor strong-for X if X rest-on floor
You denied stool strong-for Dave

Two important questions now need to be answered:

1. Why should science educators be concerned with representing such problems as support and movement in a non-scientific way?
2. How can the pertinent features of an expert system shell be of use in such a task?

There is a growing body of research interested in pupils' intuitive ideas about science and there have been a considerable number of investigations which support the view that pupils have their own conceptions about natural phenomena.

A variety of studies has been conducted in such areas as dynamics, heat, light, and many others (Gilbert and Watts, 1983). Although proof of existence of these prior beliefs is abundant, it is often difficult to fit a descriptive pattern to the results obtained, and the mismatch between pupils' understanding of science and formal science can persist even through to undergraduate level (McDermott, 1983). However these 'alternative conceptions' are so strongly taken for granted that they are often not made explicit, and need to be purposefully teased from individuals in situations where they feel it is both possible and reasonable to explain things which normally need no explanation.

Therefore, our current research interests lie in attempting to elicit and formalize in a computational model pupils' own thinking about dynamics which can capitalize upon the explanatory and query-the-user facilities of APES. The idea is to bring tacit knowledge into the open, knowledge which otherwise could dominate thinking without the person even being aware of the process. That is, if pupils' common-sense notions about dynamics can be brought out using this tool, in a way in which the consequences of holding incomplete knowledge or an inadequate rule system will cause the system to reach invalid answers or to ask unreasonable questions, then this combination could give rise to a new level of pupil awareness about their own ideas and could help to provoke a reconsideration of their current ways of thinking. The concept area of dynamics is particularly appropriate, since Newtonian dynamics does indeed require a fundamental shift in the basic concepts used to understand the work.

This work is at present at a very early stage, with more to show by way of ambitions than achievements. The example of 'support' is purely illustrative: other thinking we hope to capture would involve ideas of what makes things move, what makes them fall, what stops them moving or falling, and so on. The planned research has two essential dimensions: first, to explore the idea of formal representations of informal ideas and second, to explore the idea of using PROLOG as a mental 'scratch pad' used to externalize and so make available for scrutiny thoughts and theories about processes.

REFERENCES AND FURTHER READING

Gilbert, J K and Watts, D M (1983), 'Concepts, misconceptions and alternative conceptions: changing perspectives in science education'. In *Studies in Science Education*, Vol 10, pp.61-98.

McDermott, L C (1983), 'Critical review of 'research in the domain of mechanics'. In Delacote, G (Ed) *Research on Physics Education: Proceedings of the First International Workshop, La Londe les Maures*, Edition du CNRS, Paris.

Appendix 3

教育研究學報1988，第三卷，23—29頁
Educational Research Journal 1988, Vol. 3, pp. 23–29

Knowledge Structures — Where Can We Find Them?

Nancy LAW

University of Hong Kong

Over the past ten years, one of the major concerns amongst researchers in science education has been the problem of learners' pre-instructional ideas. It has been widely recognised that before students come to the science classroom, they already have developed some conception of the world around them and that such intuitive ideas very often persist despite instruction. Until very recently, most of the research has been focussed on where these ideas depart from the scientific view. Faced with the voluminous data from such research, mainly in the form of a catalog of misconceptions, many research workers in this field are recognizing the need to find a more comprehensive and deeper understanding of the nature of the students' alternative conceptions and different hypotheses are put forth for their interpretation. Some propose to look for a commonsense 'theory' of the content of 'alternative conceptions' which are 'more general than a set of rules about particular cases', capable of accounting uniformly for the various conceptions catalogued (Ogborn, 1985). Others, not expecting the students' conceptions to be consistent, structured theories, propose more pragmatic approaches, ranging from finding a set of correlational descriptions between situations and student responses (Viennot, 1985) to looking for a set of prototypes, not necessarily coherent, which the students would use to handle different situations (Guidoni, 1985).

There is no dispute that scientific knowledge is well structured, and so probably is the expert's scientific knowledge. What about commonsense, intuitive ideas? Are they structured as well? How coherent or consistent are they, and how persistent? Furthermore, how does the learning of school physics interact with their intuitive ideas? Are the two domains of knowledge compartmentalized as some have alleged (Claxton G., 1985), or are they integrated? If it is the latter, what are the different modes of integration? These are important questions for the understanding of the nature of alternative concepts, and for providing insight into the process of learning in situations where the student comes to the classroom with a lot of pre-conceptions. The present research tries to look for answers to these questions by eliciting students' intuitive ideas about motion through asking them to write expert systems about motion, and to observe their interactions with their own knowledge so expressed.

Method

The methodology employed in this research is essentially one of knowledge elicitation through programming. The basic idea is that here the student plays the role of the teacher and the computer that of a learner. In the course of teaching the computer, the student is obliged to externalize his/her ideas. Once these ideas and intuition have been externalized, they become in some sense accessible to reflection. The advantage of having such ideas externalized in the form of an expert system is that it 'freezes' the students' ideas at a certain instant, allowing the student to use it to explore the logical consequences of the externalized knowledge and at the same time supporting easy modification when this is deemed necessary.

An important prerequisite for such a scheme of work to be viable is the availability of a programming environment that is amenable to use and effective in representing commonsense knowledge. PROLOG is chosen for this research as it provides a declarative programming environment and a built-in logical inference mechanism so that the user only needs to declare the relevant facts and inference rules and the resulting knowledge base can then be put to work without extra programming effort. PROLOG also excels conventional languages in its ability to handle symbolic and non-numerical relationships and this is very important here since what needs to be elicited are qualitative intuitive ideas rather than mathematical relationships. In order to provide better programming support to the students, an expert system shell — APES (Hammond P. & Sergot M., 1984— is provided to the students on top of PROLOG. This allows the computer to explain how it arrived at a certain result based on the rules it has learnt. Further, it allows factual information required for getting inference results to be entered through the 'query-the-user' facility during the course of running the expert system developed (Law N, Ogborn J. & Whitelock D., 1988). It also records these inputs and thus the knowledge base increases even as the expert system is being used.

An example dialogue between the student and the computer goes like this:

Student: object's motion caused by itself if object has force supplied to it by itself.
Student: bird's motion caused by what?
Computer: bird has force supplied to it by what?
Student: itself.
Computer: Answer is: bird's motion caused by itself.
Student: Why?
Computer: Bird's motion caused by itself because bird has force supplied to it by itself.

It can be seen from the above example that a student can easily move between putting in new rules, editing them and executing the expert system at any stage during development very flexibly without having to explicitly move from one operating mode to another.

This study has been carried out with two groups of students. The first group was composed of sixth-form science students (17-yr-olds) who have just completed a two year O-Level course in Physics, while the second group was made up of form three students (14-yr-olds) who have had no formal instruction in Physics. Each student worked on the programming task for about seven sessions, each lasting approximately one and a half hours. The data collected include the developed programs, as well as screen dumps and diary notes taken during the programming sessions.

Students' Private Physics

One objection that may be raised against using PROLOG as the means of eliciting children's ideas is that it places an artificial constraint on the students to express their ideas in the form of logical statements, whereas one would expect intuitive ideas to be much less structured and possibly much less logical. A careful look at the programs reveals that the students use the PROLOG programming environment as the medium to express what they think general rules about motion should be like, and the syntax of PROLOG does not seem to have posed too much restriction on expressing their intuitive ideas. A typical rule written by a student is as follows:

```
_object motion-controlled-by itself if
    _object force-supplied-during-motion
    _name-of-force and
    _name-of-force force-type internal[1]
```

The students wrote their programs individually and they were discouraged from discussing their program with each other. From the wide difference in the overall content as well as structures of the programs written, there was strong evidence that there had not been much of a collaboration between the students in this task.

At a first level of analysis, one may leave the data collected during the process of program development aside, and just concentrate on the resulting final programs written by the students. What do these programs tell us about students' conceptualization of motion?

If we look at the contents of the programs, we find that many of the rules display recurrent themes that reveal certain very pertinent patterns of reasoning which when taken together form a 'mental gestalt' of naive mechanics. Though the details of the rules written differ from one student to the next, the cognitive primitives used and the framework of reasoning are very similar amongst the students in both age groups. Fig. 1 below tries to depict the framework of reasoning in the form of a network:

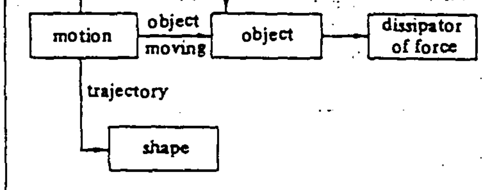


FIGURE 1. *Network showing the predominant framework of reasoning about motion.*

1 In order to keep the syntax of the program simple, students are taught to write each clause into three terms, and so hyphens are used to link up words to reduce the number of terms to that required while retaining readability. Words with underscores in front denotes a variable name.

This framework (gestalt) of reasoning goes as follows: For any motion there needs to be a cause. The cause of motion is normally an agent, and the agent supplies the force needed for the motion. Thus force is seen as an entity that can be possessed by an object and can be passed from one object to another. Furthermore, the force would be dissipated by the moving object during the course of motion. In some instances the agent of causation can be the moving object itself, e.g. a bird flying in the air, and in such cases, the source of force would be termed internal, otherwise the source of force would be labelled external.

Knowledge Structures & Intuitive Ideas

It would be an over-exaggerated claim to say that the structures of the students' programs are direct representations of the structures of the students' intuitive ideas. They nevertheless provide us with important information as to how they relate their various ideas and conceptions of motion together. There are four possible dimensions of looking at the degree of structuredness of the knowledge expressed in the programs:

i) whether there exists any prominent 'mental gestalt' as expressed in the program and their relative importance within the program,

ii) whether entities occupying the same syntactic role in similar relations share the same onto-logical status,

iii) consistency in the nature of the conditions used for deciding among the same group of goal statements,

iv) the inter-relatedness of the rules.

Note here that the point of interest is the structure of the expressed knowledge content per se and not the structure of the rules from the point of view of the PROLOG programming skill demonstrated by the students.

Existence of Central Mental Gestalts

It is interesting to note that in the twelve programs completed by the students, some clearly express strong adherence to some central cognitive 'gestalts' about motion similar to the ones depicted earlier. On the other hand, some of the students' programs only consist of ad hoc rules, revealing no causal relationships between the goal statements and the conditions, and the ideas revealed in one rule seldom recur again in other rules. The following extract (Fig. 2) taken from one of the students' programs gives a clear illustration of the ad hoc nature of the rules she has written.

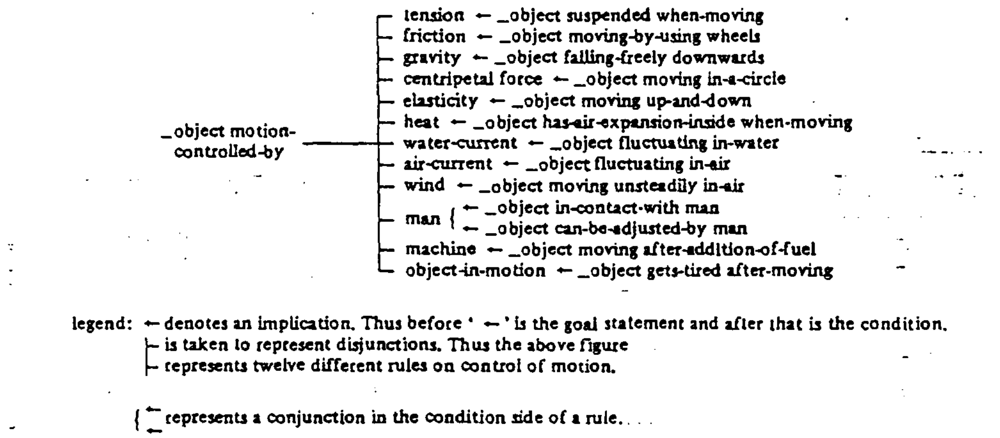


FIGURE 2. *A Portion of the Program written by Brenda represented in the form of a Network.*

Fig.3 is a network representation of the work of another student in the same group – the 17-yr-old group. While Branda's rules are all ad hoc in nature, Evon's program demonstrates clearly that she is using a central mental 'gestalt' similar to the one described earlier in Fig. 1 as the basis for constructing all her rules. In other words, Evon is displaying a strong persistent adherence to a framework of intuitive physics, while Brenda does not. This difference in the structure of the students' programs appeared in the work of both age groups, suggesting that this feature may not be particularly age dependent.

Ontological Status of Entities Participating in the Same Relation

Another important dimension for analysing the degree of structuredness, or perhaps coherence, of the programs is the variety in the ontological status of entities participating in the same relation in the programs. Evon's program is quite well

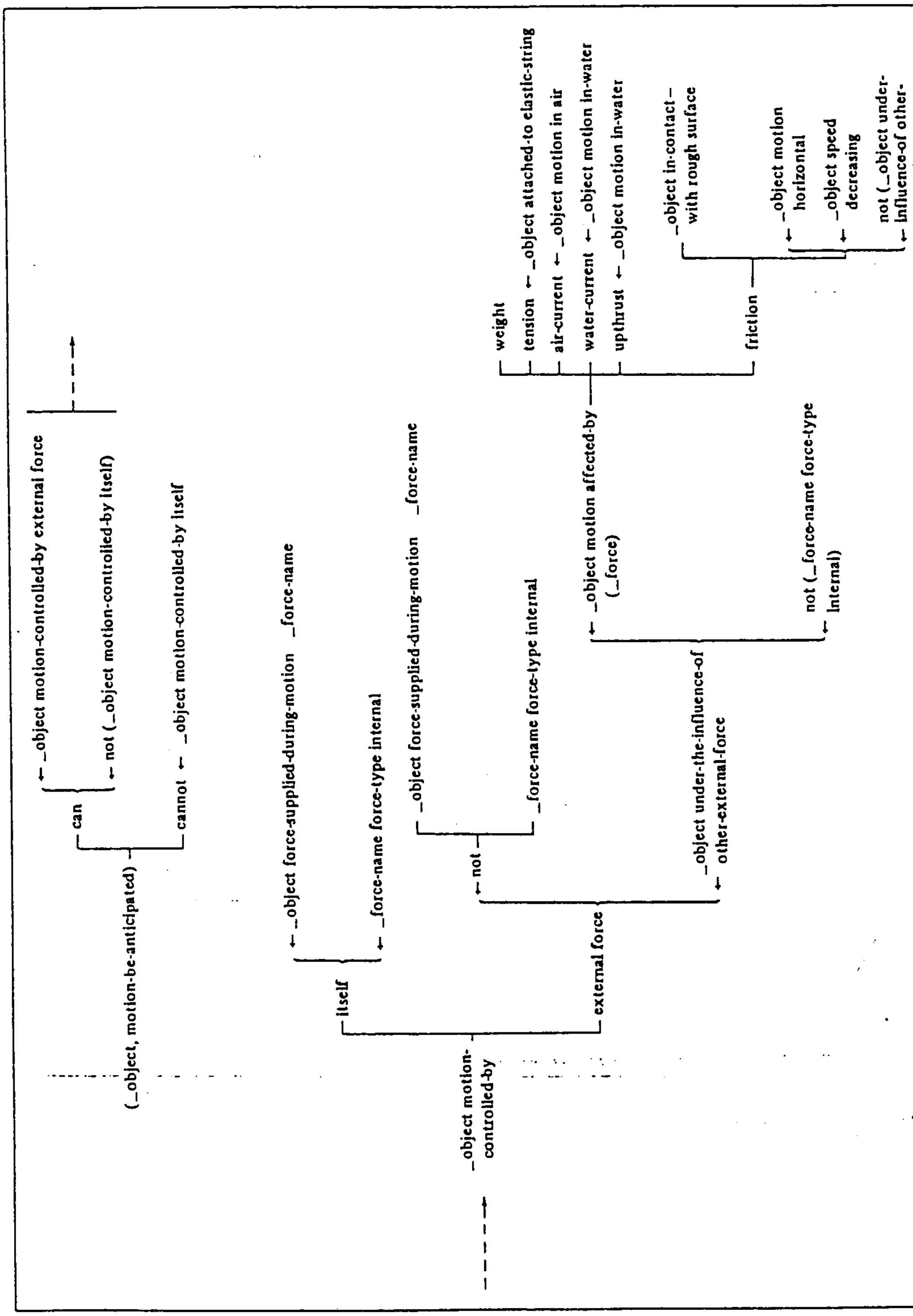


FIGURE 3. *Network representation of Evon's program.*

structured from this perspective — with the exception of the relation 'motion-controlled-by' where the entities for control of motion were 'itself' (which is an agent) and 'external-force', all the other relations were used for relating the same kind of entities. For example, in her program, the entities responsible for causing motion are all physical agents, and the entities that can affect motion are all different kinds of forces. On the other hand, Brenda's program is extremely unstructured (non-coherent) from this perspective — motion can be controlled by forces (like tension, friction, gravity), or material properties (like elasticity), or energy (like heat), or natural forces (like water current, wind) or physical agents (like man, machine or the object itself). This unstructuredness in the programs reflects a lack of clarity and rigour in the student's analysis of the situation. Differences along this dimension of structuredness appear in the work of both age groups — there are some very well structured programs in the work of the younger group as well as some very poorly structured programs in the work of the older group.

This kind of difference also gives us important insight as to the mode of integration of the student's learnt physics concepts with his/her intuitive understanding of motion. Both programs made reference to tension, a term which they picked up in the context of their school physics. Evon explicitly placed tension as a force that may affect an object's motion. Furthermore, this force would be labelled as an external force under her scheme of intuitive understanding. For Brenda, her program did not put any explicit label to classify the ontology of the term 'tension', and as a matter of fact it was just grouped together with a variety of different kinds of entities. It seems that she has only some vague impression that tension somehow affects the motion of an object, and that this happens in cases when an object is suspended. The learning of physics has not helped her to build up a theoretical framework for analysing motion, and her only gain seems to be the acquisition of a new term which is associated with a very specific situation.

In the work of the younger age group, many fewer jargons and technical terms like tension and upthrust were used. Learned school physics has not had an opportunity to add such items to their vocabulary.

Consistency in the Nature of Conditions Used for the Same Group of Goal Statements

For Evon's program, basically the same kind of conditions would be used for deciding among the same group of goal statements. For example, what controls motion would be decided by finding out what the nature of the source of supply of force is — if the source of supply is internal, then the control would be itself, and the control would be by external force when otherwise. To decide what forces affect an object's motion, the criteria she uses would be to find out the actual physical situation of the object — what medium it moves in and what it is in contact with. For Brenda, the conditions for deciding what controls a particular motion vary a great deal from case to case and there is no general principle to guide the decision. The conditions may be the physical situation of the moving object, or the trajectory of the object, or even an animistic criterion of whether the moving object would get tired after moving. This lack of consistency in the nature of the conditions used again pertains very much to whether the student has a clear theoretical framework for looking at motion, even if it were an intuitive one.

Inter-relatedness of the Rules

For the programs which display a definite adherence to a theoretical model for the interpretation of motion, the rules tend to fall into groups with inter-relationships linking the groups together. A careful look at Fig. 3 shows that the inter-relationships running between the groups of rules follow very closely that of the links between key concepts depicted in Fig. 1. On the other hand, for a program that does not demonstrate any adherence to a clear theoretical model, the rules still fall into groups, but there are no links between the rule groups. It is just a flat structure displaying no hierarchical relationship between the concepts involved and the rules are all just one level associations.

Knowledge Structures and Effects of Interacting with Own Knowledge

The PROLOG programming environment together with the expert system shell provided readily allows the students to query the knowledge bases they developed, and also provides facilities to explain how the inferences were arrived at when required. This allows students to explore the logical consequences of their externalized ideas. During the process of development and modification of their programs, the students were unavoidably confronted with unexpected results. The nature of the cognitive conflict and what the student may gain from the interactions depend very much on whether the programs reflect any core theoretical hypothesis and its degree of structuredness.

In the case when the programs demonstrate a strong theoretical hypothesis, the program normally fails because the fundamental assumptions/concepts are invalid. The normal strategy employed in such cases seems to be that of patching: he/she would try to adhere to the same basic theoretical framework and try to add in minor adjustments to get around the specific instances of failure, rather than to review the theoretical framework itself. The following episode taken from Evon's programming sessions is a typical example of one such interactions:

She had written the following rules for determining the cause of motion:

_object motion-caused-by itself if
  _object force-supplied-by _object

_object motion-caused-by machine if
  –object force-supplied-by machine

_object 1 motion-caused-by _object 2 if
  _object 1 force-supplied-by _object 2

_object motion-caused-by gravity if
  not (object under-the-influence-of other external-force)

She then set out to test the program with the case of a falling apple (thinking about the famous 'Newton's apple'). To her dismay, the computer did not come up with the expected answer. The reason was that in one of her earlier sessions, she included weight as an external (supply of) force, together with forces like friction, air current, etc. Faced with such a contradiction, her immediate response was to exclude weight from her list of external forces. As a matter of fact, in her scheme of intuitive physics, there was no clear orientation for classifying weight as internal or external forces since she could find justification for classifying it as either type of force. Weight arises because of attraction from the earth which is external to the object, so it makes sense to classify weight as external force. On the other hand, weight is related to the mass of the object which is an intrinsic property of the object itself, so it also makes sense to classify it as internal force. After changing weight to an internal force, the computer still did not give her the expected answer – gravity, because according to one of her other rules, anything placed in air would be affected by air-current, and air-current is an external force. Now, excluding air-current from the list of external forces seemed to be too much a twisting of facts. She just replaced the forth rule listed above by:

_object motion-caused-by gravity if
  not (_object motion-caused-by _something)

to avoid the problem with specific instances of internal or external forces. She did not review her underlying concepts about sources of supply of force, but then this was not the end of the problem for her.

She again met similar problems relating to internal and external supplies of force when she explored her rules on control of motion, and she could not find a satisfactory solution after patching in various parts of the program. At the end of the programming task, she was beginning to raise questions about the more central assumptions.

When the programs fail to give an expected answer, the kind of conflict faced by a student who do not hold a strong theoretical framework of intuitive physics, i.e. one whose program has an unstructured knowledge base, is much less serious and it was much easier to remedy by local patching. For them, the characteristics chosen to be put into the premises do not belong to the same ontological category. For example, whether something gets tired after moving describes the subjective feeling of an animate object whereas whether something needs fuel for motion is an objective fact. Yet both are used as defining characteristics for making inferences about the control of motion. When the program fails, the student's reaction was either that he/she has chosen a less appropriate characteristic, or that he/she has missed out some other important characteristics.

Looking at the interactions between the students and their own PROLOG programs, there is evidence that such explorations with one's own externalized knowledge do provide an opportunity for inconsistencies in one's ideas and concepts to show up. Further, when the knowledge base expressed is well structured, such interactions may eventually lead to fundamental restructuring of core concepts and hypothesis.

Conclusions

Results from the preliminary analysis indicate that it is possible to find out about a student's framework for conceptualization of motion through PROLOG programming. The elicited knowledge bases indicate that the intuitive ideas are not compartmentalized with respect to their learnt physics concepts, but rather become woven into their own intuitive framework for looking at the world of motion. The nature of the integration depends very much on how well structured their own knowledge bases about motion are. Intuitive ideas may or may not be well structured, and the degree of structuring seems to pertain to the cognitive style of the individual, and does not bear any relation to the age of the student nor is it related to the amount of Physics the student has learnt. For the 17-yr-old group, it is the higher ability students who show a much better structuring of ideas than the lower ability group; but no such trend can be found in the younger group. Most of the students found the performance of the program containing the ad hoc rules more acceptable than the better structured ones.

Acknowledgement

The author wishes to express her gratitude to Prof. Jon Ogborn who has provided invaluable support throughout the course of the present research. This research is supported by grants 336/053/0001 and 335/053/0007 from the University of Hong Kong.

References

Claxton G.L. (1985). Teaching and Acquiring Scientific Knowledge. In Keen T. & Pope M. (Eds.) *Kelly in the Classroom: Practical Applications of Personal Construct Psychology*. Montreal: Cybersystems.

Guidoni P. (1985). On Natural Thinking. In *European Journal of Science Education*, 7(2), 133-140.

Hammond P. & Sergot M. (1984). *APES: Augmented PROLOG for Expert Systems, Reference Manual.* Logic Rased Systems Ltd., Surrey, England.

Law N., Ogborn J. & Whitelock D. (1988). Knowing what the Knower Knows: a Use of APES in Science Education. In Nichol J, Briggs J. & Dean J. (Eds), *PROLOG, Children and Students*. London: Kogan Page.

Ogborn J. (1985). Understanding Students' Understandings: an Example from Dynamics. In *European Journal of Science Education*, 7(2), 141-150.

Viennot L. (1985). Analyzing Students' Reasoning in Science: a Pragmatic View of Theoretical Problems. In *European Journal of Science Education*, 7(2), 151-162.

知識結構—何處尋？

羅陸慧英
香港大學

（摘要）

很多研究顯示，學生於正式從課堂中學習科學知識之前已對其周圍環境作出認識，從而對一些事物產生不少直覺概念，而且這些概念非常牢固，不易爲課堂教學所改變。直至最近，這方面的研究大多集中於搜集學生直覺概念中的錯誤概念。然而近兩、三年來一些研究人員已提出需要對這現象作更全面、深入的研究。其中一個很重要的課題是如何認識這些直覺（錯誤）概念。究竟這些不同概念是否相容？是否形成一套「直覺理論」，或只是一些不大相關，有時可能是互相矛盾的個別經驗性聯繫？科學知識擁有一個被清楚公認的結構，但直覺概念又是否同樣擁有結構呢？它們是相容還是相矛盾？有多牢固？這些都是此研究所希望探索的問題。

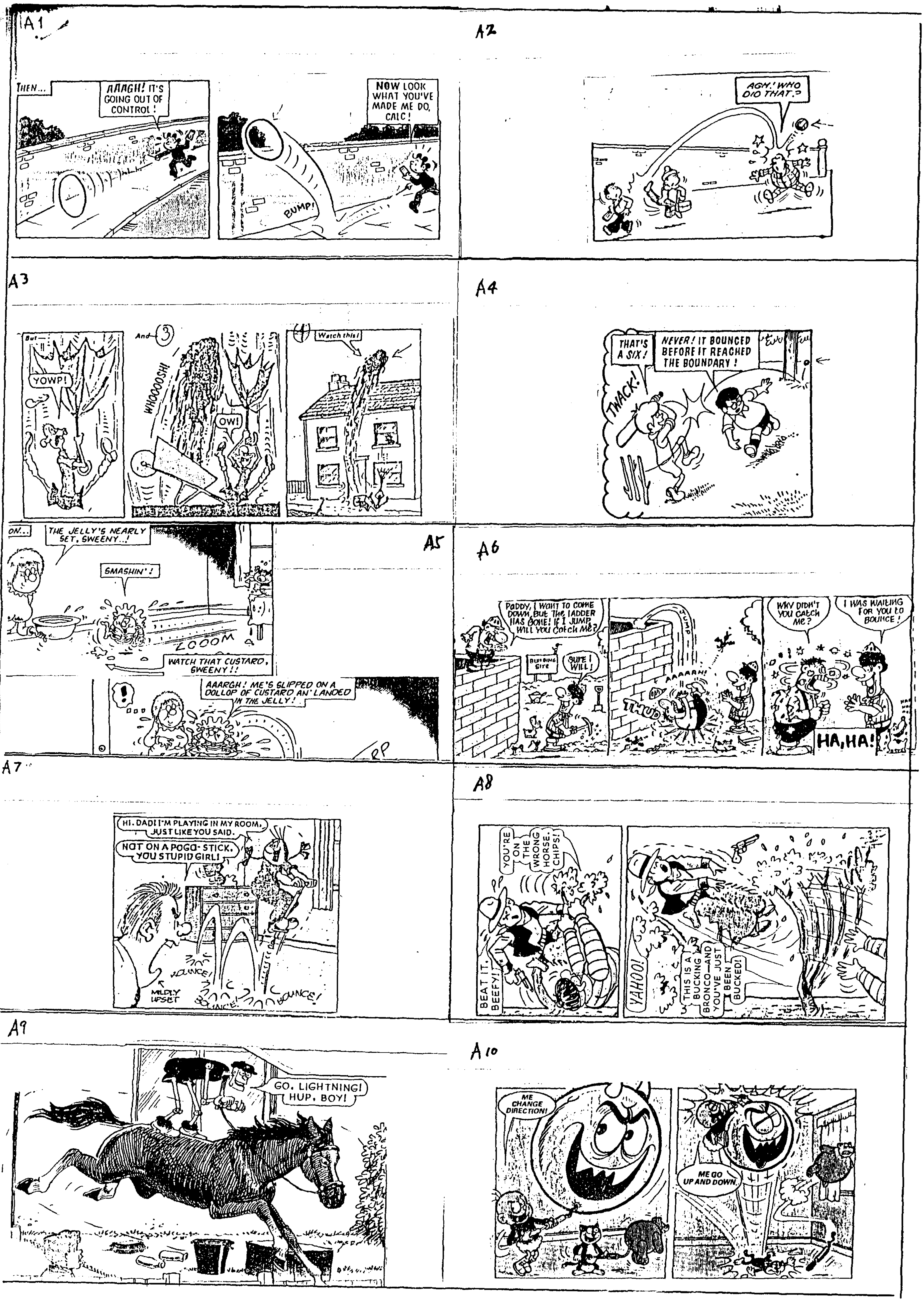
此研究採用方法主要是透過學生使用PROLOG電腦語言編寫專家系統，以得知其所持有關物體移動之觀點、概念，從而探索上列問題。

Author

Nancy LAW, Lecturer, Department of Professional Studies in Education, University of Hong Kong, Pokfulam Road, Hong Kong

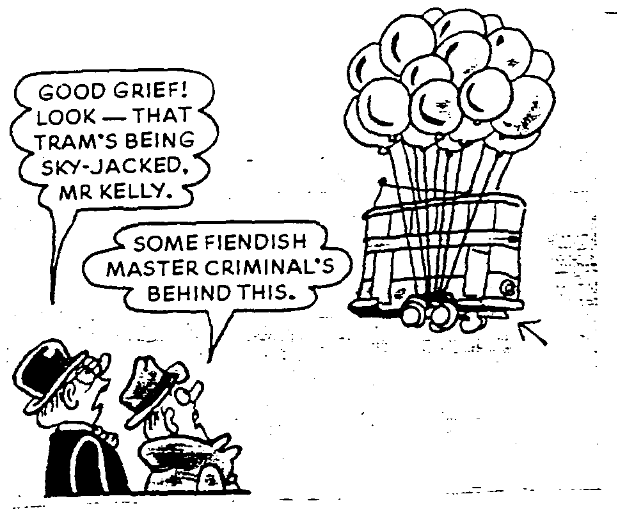
Picture Set A

Appendix 4. The three picture sets (sets A, B & C) used in the classification task.

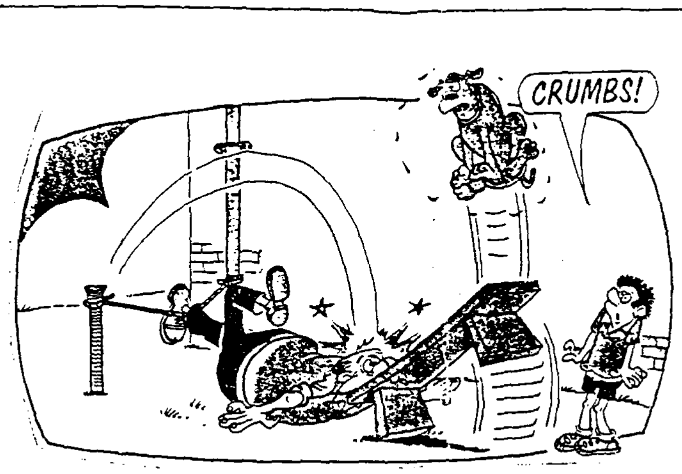


Picture Set B

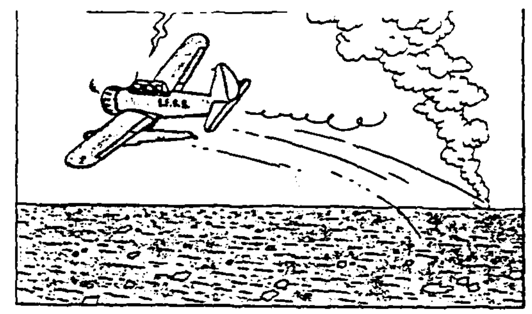
B7



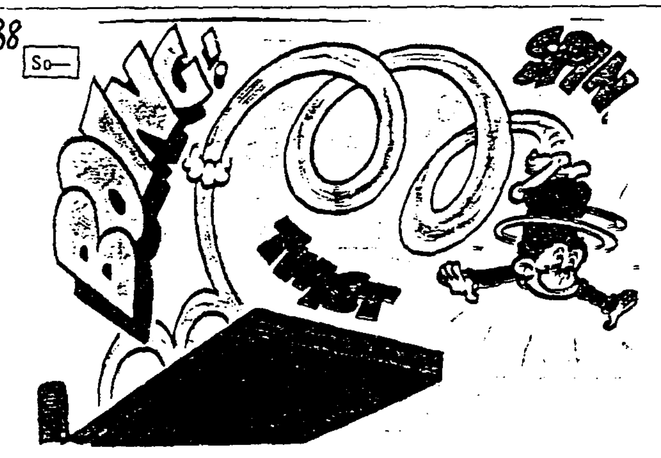
B4



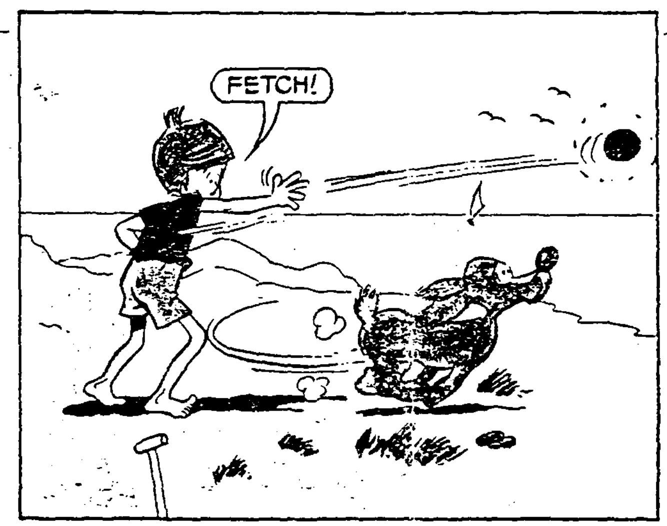
B1



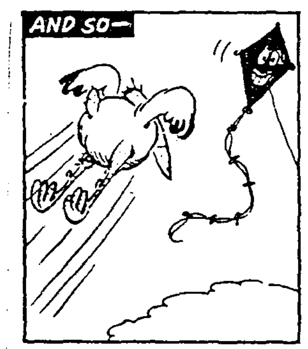
B8



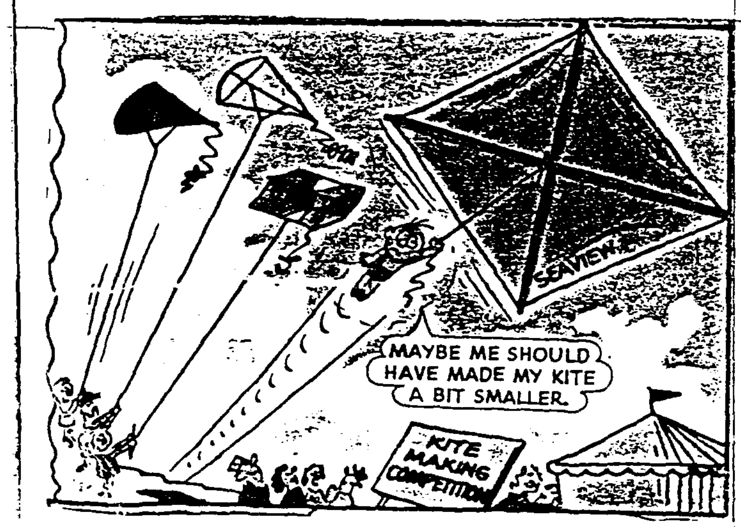
B5



B2

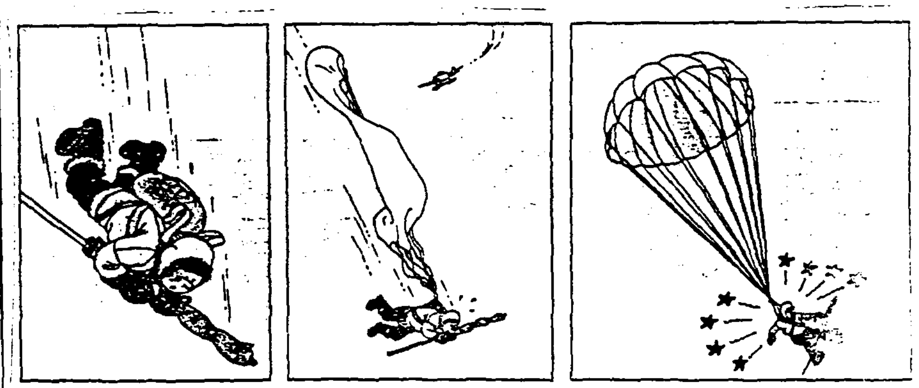


B9



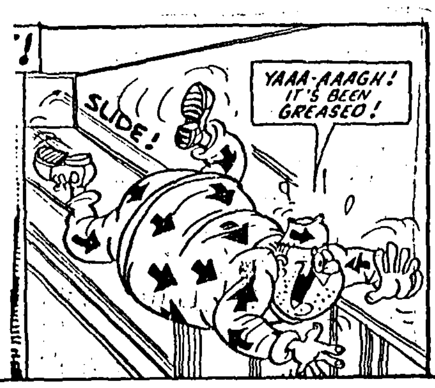
B6

B3

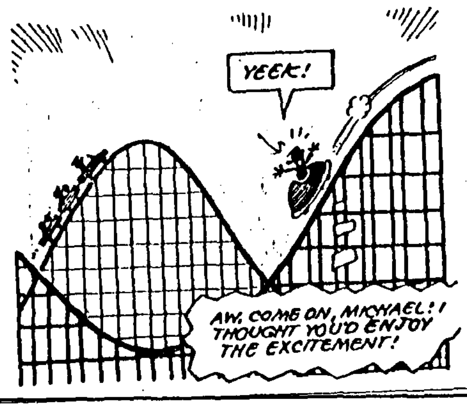


Picture Set C

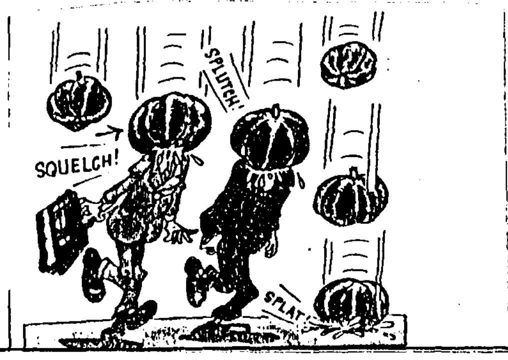
C 1



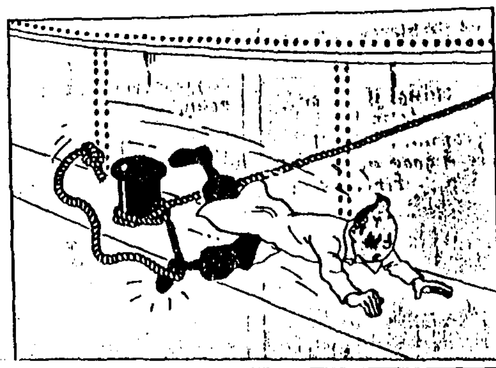
C 2



C 3

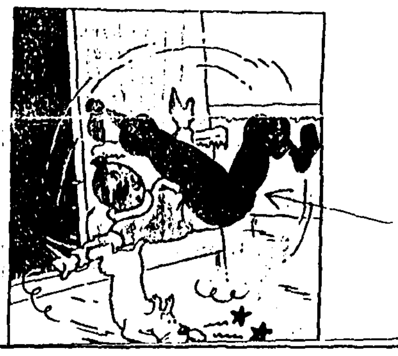


C 4

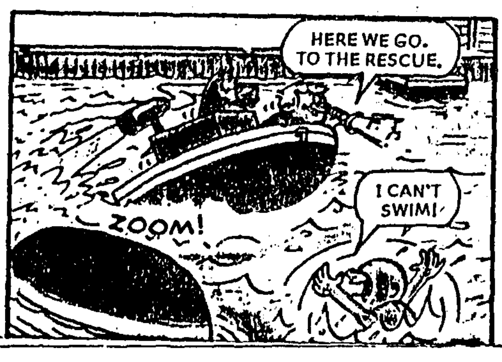


C 5

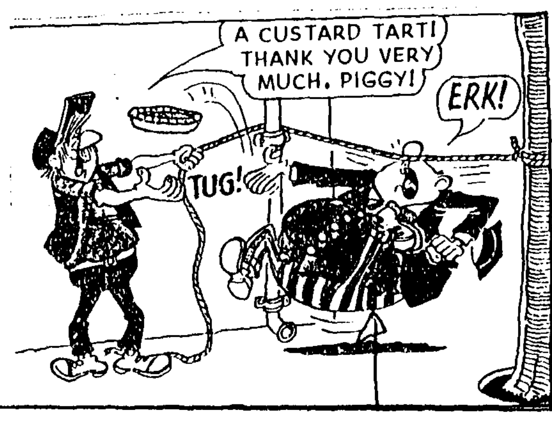
C 6



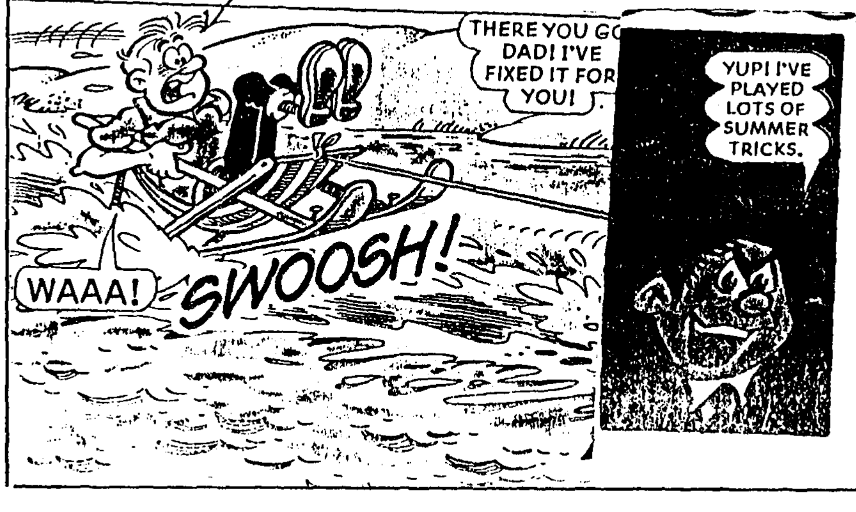
C 8



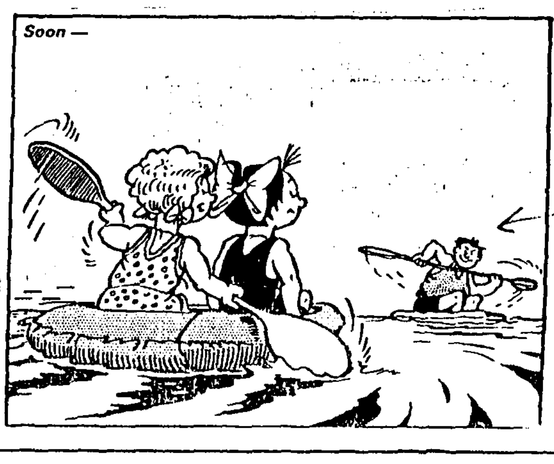
C 7



C 9



C 10



Appendix 5. Handout given to students to introduce PROLOG programming.

INTRODUCTION TO PROLOG PROGRAMMING

A. System startup

1. Put the program disk in drive A.
   Turn on the computer (or if the computer is already on, press Ctrl,Alt and Del together and then release).

2. Type "timer/s".

3. Type "prolog".

4. Type "LOAD APES" (note that it must be in uppercase characters.)

5. If you want to load a program from previous sessions, put the data disk in drive A and type "LOAD <FILENAME>". After the program has been successfully loaded, replace the program disk in drive A.

6. At the end of a session, save the program on the datadisk as well as on the data backup disk by putting the appropriate disk in drive A and then typing "SAVE <FILENAME>".

B. Rules and facts

A prolog program may be thought of as made up of *facts* and *rules*.

Facts are statements consisting of three terms expressed in the form:

```
      object    relation    object
```

```
For example:  John isa boy
          Mary is-in-condition happy
          pochai-pills aggravate peptic-ulcer
```

Notice that if the fact expressed can only be intelligibly expressed in more than three words, some of the words may be hyphenated to form a term so as to conform to the above format.

Rules are stated in the following format:

```
    statement1 if
        statement2 and
        statement3 and
              . .
              . .
```

where each statement is in the same format as a fact statement. Statement1 is called the *goal statement* or *consequence*, and the rest are *condition statements* connected by AND.

The OR connective is not supported as part of a rule. Instead, if a goal can be arrived at by more than one set of conditions, this can be expressed by listing each set of valid conditions independently.

That is, a program in the form of

```
rule1
rule2
rule3
  ..
  ..
```

can be read as rule1 OR rule2 OR rule3 OR ..

C. <u>Adding program lines</u>

Program lines (both facts and rules) can be added in the following format:

```
add( program statement)
```

For example: add(aspirin suppresses pain)

D. <u>Querying the knowledge base</u>

Queries to the knowledge base can be done through two commands:

confirm - which would get back a response of YES or NO

find - which would provide a list of appropriate answers.

For example:

```
confirm(alcohol suppresses pain)

find(_thing: _thing suppresses pain)
```

You can always ask for the reasons for a result by typing "WHY" or "HOW".

For the *find* query, you can ask for more answers by typing "more".

You can terminate the dialogue by typing "stop".

Appendix 6. Summary of the classificatory schemes given to the F.6 students on their first programming session. (These classificatory schemes were used by their own group during the classificatory task.)

SUMMARY OF CLASSIFICATION GROUPS

This is a summary of the classification types as discussed during the interviews on the comic strip pictures of motion. As can be seen from the summary, the ways of classification are very similar.

	BS	VW	WW	CL	SH	KH	CH	WF	VN	KL	CW
Types of motion											
- circular						*					*
- parabolic					*						*
- straight line								*			
- rotational								*			*
- bouncing											*
Type of positional changes											
- periodic				*				*			
- change of height	*		*			*					*
- forwards/backwards						*					
- will return to ground		*									
Direction of motion											
- upwards and/or downwards	*	*	*	*	*		*		*	*	
- horizontal	*		*				*			*	
Purpose of motion											
- on purpose			*	*					*		
- tricked by other people			*								
- to maintain balance							*				
- to jump across obstacle											*
Cause of motion											
- self			*		*			*		*	*
- external object					*			*		*	*
- machine							*				
- gravity			*								
- whether the cause itself is in motion or not	*										
- direct or indirect					*						
- reduced friction (slipperiness)							*				
Control of motion											
- self		*		*		*			*	*	
- external		*		*		*			*	*	
- gravity										*	
Whether external object/ instrument is needed	*										

	BS	VW	WW	CL	SH	KH	CH	WF	VN	KL	CW
Factors affecting motion											
- gravity						*		*			
- tension						*					
- floatation						*		*			
- air current	*										
- water current						*					
Source of force											
- self-supplied				*						*	*
- external force		*		*	*	*				*	*
- whether in direct contact with object					*						
Sudden changes		*	*	*	*						
cause:											
- slipperiness											*
- friction											*
- blocked by another object						*					
Type of force											
- springy, elastic											*
- moment											*
- biological											*
Forces acting											
- number		*				*		*			
- relative directions		*									
Speed			*						*	*	
Nature of object											
- living/non-living			*	*							*
- whether capable of changing its own cause of motion				*				*	*		*
- elastic, springy											*
Consequence(effect) of motion											
- change of shape				*							
- carry something else along				*							
- only affect itself				*							
"Naturalness" of motion with respect to the type of object	*	*		*						*	
Whether motion makes use of specially designed instruments											
- to change (increase or decrease air reistance)									*		

	BS	VW	WW	CL	SH	KH	CH	WF	VN	KL	CW
Whether the motion can be anticipated or not			*	*	*						
Whether moving object is attached to another object				*					*		
Whether the motion can be sustained or is only short lived	*										
Whether the motion is safe	*						*				
Whether motion is possible	*								*		

Key to the initials:

- BS Brenda Siu
- VW Vince Wong
- WW Wong Wai Ling
- CL Lin Chu Chang
- SH Sandra Hui
- KH Ho Kai Wing
- CH Charmaine Hon
- WF Fong Wai Leung
- VN Vivian Ng
- KL Li Kit Yung
- CW Wong Chi Wai