University of Massachusetts Amherst
ScholarWorks@UMass Amherst

Doctoral Dissertations 1896 - February 2014

1-1-1991

Case studies of cycles in developing a physics lesson.

Aletta Isabella Zietsman
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_1

Recommended Citation
Zietsman, Aletta Isabella, "Case studies of cycles in developing a physics lesson." (1991). *Doctoral Dissertations 1896 - February 2014*. 4827.
https://scholarworks.umass.edu/dissertations_1/4827

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations 1896 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

UMASS/AMHERST
312066013297005

CASE STUDIES OF CYCLES IN DEVELOPING A PHYSICS LESSON

A Dissertation Presented

By

ALETTA ISABELLA ZIETSMAN

Submitted to the Graduate School of the

University of Massachusetts in partial fulfillment

of the requirements for the degree of

DOCTOR OF EDUCATION

May 1991

Education

© Copyright by Aletta Isabella Zietsman 1991

All Rights Reserved

The study was partially supported by a grant from the National Science Foundation, #MDR-8751398. Any opinions, findings and conclusions or recommendations expressed in the manuscript are those of the author and do not reflect the views of the National Science Foundation.

CASE STUDIES OF CYCLES IN DEVELOPING A PHYSICS LESSON

A Dissertation Presented

By

ALETTA ISABELLA ZIETSMAN

Approved as to style and content by:

Klaus Schultz, Chairperson of Committee

John J. Clement, Committee Member

J. Michael Royer, Committee Member

Marilyn Haring-Hidore, Dean

School of Education

DEDICATION

To my dear, gentle parents - this one's for them. Without their boundless love for, and belief in their wayward offspring, nothing could have been. Vir al die verlang, die een is vir julle twee!

ACKNOWLEDGEMENTS

My sincere thanks go to:

The International Fulbright Scholarship Foundation and Delta Kappa Gamma International who gave generous financial support much appreciated.

The seventh grade participants and their great science teacher, Phil Covelli: the collaboration was a joyous learning experience for me.

The members of my committee: Professor Klaus Schultz, the chairman, for his thoughtful comments, cheerful support and encouragement, and for keeping me on the straight and narrow; and Mike Royer, for introducing me to the exciting new fields of cognitive psychology and cognitive science, and finding time to answer the questions of a psych novice. Special thanks to the director of the research, John Clement, of whom it can surely be said that "he wears his learning lightly and shares it generously": for sharing his knowledge of cognitive and qualitative research; for his consistent, patient support and gentle criticism; and for the pure fun this study has been.

Peter Hewson, who showed me many years ago at Wits that there is more to teaching physics than instinctive "moves": for his continued support, advice and friendship.

My SRRI "peer support system": Val Veneman and Dave Brown, for their thoughtful contributions to the early diagnostic tests; Pam Hardiman, for her careful reading and valuable comments on the pilot studies; Bob Dufresne and Tom Murray, for their time and sharing their ideas; and Kwan-Min Lee, my office-mate, for company, friendship and sharing the ups and downs of graduate research.

Makaziwe, for the strong, loving hand that held mine in friendship through personal crises and various tribulations.

The "new" friends in Amherst - especially Isaac, Dorothy, Sara, Sue and Dave and Michelle - for support and so much more; thanks for the wonderful memories. Special thanks to Jack and Dani Lochhead, who went on sabbatical and left me in charge of the house, hottub, cars and cats in Conway - and saved my comps!

My siblings, most of all: Paul, Pat, Cecile, and that marvelous circle of family and friends in the southland - for keeping Africa and language alive during the past three years.

ABSTRACT

CASE STUDIES OF CYCLES IN DEVELOPING A PHYSICS LESSON

MAY, 1991

ALETTA I. ZIETSMAN

B.Sc., STELLENBOSCH UNIVERSITY, STELLENBOSCH

M.Ed., UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Ed.D., UNIVERSITY OF MASSACHUSETTS, AMHERST

Directed by: Prof. Klaus Schultz

Children's reasoning and learning about levers and simple machines were investigated in this study. The study included several cycles of design, test and clinical interview tutoring sessions and the two final cycles are presented here. The methodology combined the use of qualitative clinical interviewing data and quantitative summative data: quantitative evaluations provided an overview of the lessons' effects, while qualitative, formative lesson evaluations allowed deeper insights into learning and reasoning processes.

Three groups of participants were interviewed about the pretest, lesson and posttest. The pre- and posttests were standardized, and several new and widespread misconceptions about levers have been discovered that are less accurate or general than conventional conceptions.

In experiment 1 the pre-posttest comparison between the control group and experimental group 1 showed that there were no differences and

the instruction in experiment 2 was revised considerably as a result of the formative evaluation findings. Significant improvements were apparent for experimental group 2 with regard to conceptual change and far transfer when compared with experimental group 1 - evident in group 2 students' ability to transfer their acquired knowledge to complex and compound levers and in conceptual changes apparent in simple levers questions.

Lesson 1 was essentially a bridging lesson where "intuitive anchoring" examples were extended analogically via intermediate bridging cases to a target situation. The findings from lesson 1 suggested that reasoning from extreme case situations of levers might be instructionally useful, and this hypothesis was confirmed by results from experiment 2, where the instructional sequences based on extreme case reasoning proved to be powerful facilitators of the construction of mechanistic models by the students that fostered conceptual change and learning.

The following directions for further research are suggested: students' conceptual models have implications for teaching and learning that are poorly understood at this stage, and research on instruction that employs experts' non-formal reasoning strategies should be encouraged.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

BACKGROUND TO THE STUDY

The nature of students' intuitive knowledge of levers and simple machines was explored by means of interviews and diagnostic tests in this study. In two distinct but related experiments, these physical intuitions (assumed to be common to all participants in the study) were used in the design of an instructional sequence, and the development (or lack of development) in students' knowledge of elementary statics during and following instruction was evaluated. The study consists of "cyclical" development and evaluation processes: the findings from pilot studies informed the development of tests and the instruction designed in the first experiment; while the second experiment was similarly informed by the findings from experiment 1.

The rationale for the study will be outlined in the following sections.

A. Context of the Study

The constructivist view of learning, the philosophical framework within which the study is situated, holds that all people are committed to making sense of the world and do so by actively constructing knowledge in the process of understanding. In order to do this, individuals have to start with their existing knowledge - thus, one can never view students' minds as "blank slates". The teacher's conception of teaching is therefore also affected:

> [Constructivism] changes the teacher's view of 'problems' and their solutions... The teacher would come to realize that what he or she represents as a 'problem' may be seen differently by the student. Consequently the student may produce a sensible solution that makes no sense to the teacher. To be then told that it is *wrong*, is unhelpful and inhibiting ..., because it disregards the effort the student put in. In contrast, constructivist teachers would tend to explore how students see the problem and why their path towards a solution seemed promising to *them*.
> (von Glasersfeld, 1989)

The premise on which the levers study was based is the belief that, although conceptions held by students before formal instruction are often detrimental to learning theoretically correct physics concepts, some preconceptions may actually be in agreement with the accepted physical theory and should therefore be useful in instruction (Clement, D. Brown and Zietsman, 1989).

The study addressed two primary issues. First, students' conceptions about levers were investigated, to determine whether any misconceptions and instructionally useful physical intuitions existed in this content area. Second, the correct, instructionally useful physical intuitions thus identified were used to design a lesson on levers and simple machines.

The levers/simple machines domain was chosen for the following reasons. The operating principle for levers is fairly simple in structure, as was illustrated by Novak (1977) in his computer programme ISAAC. However, studies by Siegler (1978, 1982) and Hardiman, Pollatsek and Well (1986) have indicated that these principles are not easily derivable by students. Siegler's studies of children's knowledge of the balance-beam indicated that very young children make errors by failing to take some important attributes of the apparatus into consideration. In contrast, older children take all relevant attributes of the machine into account, but are unable to combine information about two relevant attributes. In a similar finding, Hegarty, Just and Morrison (1988) found that competency in solving simple problems about pulley systems in mechanics depends on the ability to correctly identify attributes of the system relevant to the system's function and to combine such information successfully in a quantitative manner. Hardiman et al. (1986) suggest that students' understanding of the balance beam cannot be described in terms of the acquisition of rules of increasing complexity as indicated by Siegler and Hegarty et al., and describe a variety of heuristics used by students to make sense of a balance beam.

Thus, it seems sensible that students' naive conceptions should be an important first consideration in the design process of instruction in elementary statics.

B. Research Relevant to the Study

1. Students' Naive Knowledge of Physics

> So how do you go about teaching them something new? By mixing what they know with what they don't know. Then, when they see vaguely in their fog something they recognize, they think 'Ah, I know that'. And it is just one more step to 'Ah, I know the whole thing'. And their minds thrust forward in the unknown. And they begin to recognize what they did not know before and they increase their powers of reasoning. (Picasso, in Gilot and Lake, 1965).

Picasso's ideas concerning the teaching of new knowledge sound very much like those proposed, with more specificity, by researchers in education and educational psychology. Indeed, research in physics education in the past few decades has established beyond doubt that students bring to physics instruction conceptions of the world that are well established and often inconsistent with the established theories (for comprehensive summaries of this research see Driver and Erickson, 1983; Gilbert and Watts, 1983; and a comprehensive bibliography by Pfundt and Duit, 1985).

Early research focused on the gathering of evidence for the existence of misconceptions, and the best researched content areas in this context are probably the force and motion ideas that students hold before and after formal instruction. In kinematics, students often confuse the concepts of position and velocity and those of velocity and acceleration (Johansson, Marton and Svensson, 1985; Trowbridge and McDermott, 1980, 1981; Zietsman and P. Hewson, 1986). Some of the earliest studies on students' conceptions was concerned with mechanics, and Helm's (1980) research showed that not only students, but also secondary school teachers hold misconceptions about force. Some of the most common, although not the only, force/motion misconceptions reported are: the so-called "impetus" misconception - that is, if a body is moving there must be a force acting on it in the direction of the motion; constant motion requires the application of a constant force; and, in accord with the previous two conceptions - if an object is not moving, then there is no force acting on it. (See for example Champagne, Klopfer and Gunstone, 1982; Clement, 1982; diSessa, 1981; P. Hewson, 1984; McCloskey, 1983; McCloskey, Washburn and Felch, 1983; Minstrell, 1982; Nussbaum and Novick, 1982; Viennot, 1979; Watts, 1983; White, 1983.)

The existence of misconceptions was also reported in several studies investigating the heat and temperature, energy, electricity, gravity and density content areas (Duit, 1981; Engel and Driver, 1982; Fredette and

Lochhead, 1980; Gunstone and White, 1980; Osborne, 1981; M. Hewson, 1986; Solomon, 1982). Misconceptions in these areas included ideas such as: *gravity* operates only within the earth's atmosphere; *energy* has to do with living and moving things; and *electric* current flows around an electrical circuit in one direction and some of the current is used up by each consecutive component.

Many commonalities in these naive or preconceptions of students have been identified in this research and perhaps most important, in-depth studies suggest that these conceptions are not just incorrect pieces of knowledge, but conceptions that are more meaningful to the students than the "correct" information taught in schools and colleges, thus playing havoc with science instruction:

> [students' prior knowledge] is logically antagonistic to the content to be learned and often persists after physics instruction. (Champagne, Klopfer and Gunstone, 1982:32).

As indicated in the paragraphs above, most of the earlier research on students' naive physics conceptions has focused on where these ideas depart from formal theories in physics. The wealth of data from this research resulted in other questions, for example questions about the design of instruction to most successfully facilitate conceptual change; questions about the nature of students' alternative views; and questions about teachers' conceptions of the teaching of physics. In the research about the nature of students' misconceptions a debate has developed between the proponents of the common sense "theory" approach (McCloskey, 1983; McCloskey, Washburn and Felch, 1983) and those viewing students' naive knowledge as "fragmented" elements in a relatively unintegrated system (diSessa, 1985; Guidoni, 1985). Research on teachers' conceptions of physics teaching is in the early stages. Using research which has demonstrated effective physics teaching, P. Hewson and M. Hewson (1988) have made some progress toward an analysis of an appropriate conception of physics teaching. More research has been reported in the conceptual change and physics teaching context. Driver (1987) outlines the different approaches reported in the literature, e.g. conflict-based teaching strategies (Novick and Nussbaum, 1982), conceptual exchange teaching strategies (P. Hewson and M. Hewson, 1983) and bridging strategies, that is building experiential bridges to new conceptions (D. Brown, 1987; D. Brown and Clement, 1987; Clement, 1986; Minstrell, 1982).

The levers study can be situated in two of the larger research areas outlined in the previous paragraphs: a documentation of students' naive conceptions of levers (statics content area) and the design of the levers instructional sequences. Students' misconceptions were documented but in addition, their "correct" conceptions about levers were investigated. With regard to the last, much less is known about the correct intuitive conceptions that students may bring to the classrooms. Indeed, many teachers intuitively use generally accepted examples to build on in instruction. These examples are mostly developed "on line", in the process of teaching and are not documented, except perhaps for personal records. The levers study adds to an ongoing systematic investigation and documentation of students' intuitive, theoretically correct preconceptions (D. Brown, 1987; Clement, et al., 1987).

2. <u>Research Related to the Instructional Design</u>

a. <u>Intuitive Knowledge and Anchors</u>. Intuitive knowledge seems usually to be defined in terms of what it is *not*: it is not viewed as rational reasoning which entails the use of reason, logic and analysis and it is not mere observations. A dictionary (Oxford English Dictionary) defines intuition as "immediate apprehension by the mind without reasoning". Goldberg suggests that "understanding and conviction may be shallow unless the knowledge is also intuitively absorbed" and cites Descartes as saying

> I understand not the fluctuating testimony of the sense, but the conception which an unclouded and attentive mind gives us so readily and distinctly that we are wholly freed from doubt about that which we understand. (Goldberg, 1983:34.)

The intuitive knowledge of concern to the levers study is more specifically named *physical* intuitions, defined by Clement (1989b:346) as "knowledge structures which can provide an interpretation of a physical phenomenon". Physical intuitions share much with the above description of intuitions in general, in that they are considered to be *elemental*, hence not requiring external explanations or justifications; *general* to the extent that they can be activated by a certain range of other "states"; intrinsic or *self-evaluated* in that people do not rely on others to decide whether an intuition is correct; and *concrete* in the sense that physical intuitions provide "direct knowledge" about a physical object (Clement, 1989b).

Clement (1989b) suggests that experts use physical intuitions as anchoring assumptions upon entering a new content domain, and it is this function of physical intuitions that is most relevant to the levers study. The assumption is that students could also use their physical intuitions as anchoring conceptions (similar to the experts' anchoring assumptions and referred to as anchors in the rest of this report) in a new content domain; and that by analogical reasoning and reasoning from extreme cases, these intuitions could be extended to explain unfamiliar physical phenomena.

b. Analogies in Learning. As indicated before, some of the study's instructional strategies utilized analogical reasoning. The aim was to ground instruction in students' physical intuitions of levers compatible with current physical theories, and to extend the understanding of the anchors to new conceptions. With regard to the plausibility of learning by analogy about levers, Siegler cites Gibson's account of perceptual learning as a possible speculation about the encoding of children's knowledge, and suggests that perceptual learning, as described by Gibson, would have to come through "some process of analogy" (Siegler, 1978: 144). Hence, children's experiences with seesaws and simple levers could enable them to learn - by analogy - about balance beams.

Learning by means of analogical reasoning has been investigated by researchers from various disciplines and most conclusions indicate gains in instruction based on analogical reasoning not equalled by more "traditional" didactic instruction (see for example A. Brown and Kane, 1988; D. Brown, 1987; D. Gentner and D. R. Gentner, 1983; Gick and Holyoak, 1983; Royer and Cable, 1976).

Some differences seem to exist between the instructional techniques based on analogical reasoning most frequently reported in the literature and the technique to be used in this study. The most common use of analogies in instruction seems to be where a base analogy (a knowledge structure) is presented to the students in text or verbal instruction. This base structure has isomorphic structural relationships with the target knowledge structure so that "structure-mapping" could occur (Dupin and Johsua, 1989; D. Gentner and D. R. Gentner, 1983). An often used example of this would be the solar system presented as an analog to the structure of the atom. Aspects of analogical transfer important in this context are that a mental construction of the base and target knowledge structures be formed where the base structure is a relevant analog to the target; and that mapping of the components of the base and target structures occurs (Holyoak, 1985).

The instructional technique used in the levers study has students' physical intuitions as "base knowledge structures", thus an anchoring conception already known and understood by the student. In one instance, bridging analogies are employed to transform an anchor gradually to the target situation. Structure mapping is important between some of the bridging analogies in the lesson, but students would probably not be able to map directly from the anchors to the target conceptions without the intermediate bridging analogies. Thus, analogical connections are established between anchors and targets that the students may not view as analogous to start with (D. Brown and Clement, 1989).

c. Bridging Strategies. Several studies dealing with instruction based on anchoring and bridging analogies were conducted at the University of Massachusetts in the past decade (D. Brown & Clement, 1989; Clement, 1987; Murray, Schultz, D. Brown and Clement, in press), with apparent success. This analogical teaching strategy, or bridging strategy could be described as follows:

Target Problem. Students are presented with a target problem with the aim to draw out misconceptions. The target examples are usually researched in diagnostic tests preceding the lesson development.

Anchoring Example. A much easier case analogous to the target example, an anchoring example, is then suggested. The presence of an anchoring conception, defined theoretically as an intuitive knowledge structure in rough agreement with accepted physical theory, would also have been confirmed in previous diagnostic tests.

Bridging Examples. Students may reason correctly about the anchor but still view the target situation as completely different. Questions about bridging analogies are then posed to the students, where the "bridges" are examples that are conceptually between the target and the anchor.

CHAPTER II

METHODOLOGICAL ISSUES

> The problem is this: why kids who are obviously so bright, and who are trying so hard, fail to understand the simple things we try to explain to them in schools. (Muller, 1986:5.)

Muller's statement summarizes several of the problems teachers and researchers in learning and teaching of physics have been struggling with. He also highlights one misconception many teachers hold about those bright children - perhaps what we are trying to explain may seem simple to us, but what are the children's thoughts on the "simple things"?

The levers study attempts to address some of the issues related to Muller's statement that are prevalent in physics education at the present time. First, to learn more about children's everyday, naive knowledge of the statics domain in physics; and second, to investigate the use of some of the children's intuitive preconceptions as group anchors, extreme cases and bridging examples in the design of instruction.

A. Research Questions

More specifically, the general research questions that guided the research can be stated as follows:

(1) Although the principle of levers can be stated quite simply, and can be used with apparent ease to make predictions concerning the behavior of levers, most students are probably not capable of stating this

principle using only their own intuitions. Therefore the following questions are of interest, both for their own sake and for their pedagogical implications:

(a) What intuitive ideas about levers, that are in agreement with the accepted physical theory, do students have before instruction in elementary statics?

(b) What misconceptions do students hold about levers before instruction in elementary statics?

(2) Assuming that the above conceptions have been identified and that a lesson grounded in the students' naive preconceptions has been developed, the following issues are of particular interest in a formative evaluation of the lesson taught in one-on-one tutorial sessions:

(a) The extent to which the lesson changed the students' naive knowledge of levers;

(b) The weaknesses and strengths in the lesson;

(c) The extent to which the students were able to transfer the knowledge of levers acquired during the tutoring interviews.

B. Methodology

1. Research Design

The pilot studies and levers study were comprised of several cycles of diagnostic testing, instructional design and instructional evaluation. The evaluation components of the levers study were both qualitative and quantitative: the summative evaluation of each of the experiments was quantitative (to the extent allowed by the research design), while the formative evaluations are in the form of qualitative case studies of the pretest, posttest and lesson interviews.

Two experiments or cycles of the lever study shall be described in this report. The design for each of the experiments were basically the same, except that the second experimental group's "control" group was experimental group 1. The designs are outlined in Figure 2.1 below. (|\| indicates the evaluations at the end of the experiments.)

	Experiment 1				Experiment 2			
Experimental Group 1	Pretest	Lesson1	Posttest	\				\
Control Group	Pretest		Posttest	\				\
Experimental Group 2					Pretest	Lesson 2	Posttest	\

Figure 2.1

Outline: Experiments 1 and 2

The experiments were therefore in a classic pretest, intervention, posttest format. In the summative evaluation of experiment 1 a conventional control/experimental group comparison was performed and the findings from this evaluation and the formative evaluation were implemented in the design of lesson 2.

The first experimental group was used as a control group for the second on the basis of the following reasoning: The pre- and posttests were the same throughout the study; care was taken to keep the experimental conditions the same, e.g. the format of the lesson and the time frame for the three interviews; and although experimental group 2 students attended more classes in biology and botany, one could accept that their formal knowledge about physics was the same as that of the group 1 students. Similar summative evaluations were thus performed for the two experiments.

2. Qualitative Case Studies

Most of the proposed research can be considered "naturalistic" in the sense that one has attempted to enter the "world" of the students as it exist. The research was "descriptive" since text was the most important form of data. The focus of the transcript analyses was on the preconceptions and learning experiences of the participants, with a working assumption that they were trying to make sense of their experiences and in doing so created their own knowledge.

A qualitative approach to the particular problems which this study addressed seemed more appropriate for the following reasons. An empirical-rational research mode works best under at least the following three conditions: one, when all the variables that affect the subject matter could be controlled or predicted; two, when one can measure, quantify and define with precision; and three, when one has complete and adequate information. None of the above were true with regard to the problems in this research, and it seemed more sensible to employ a "grounded theory" approach (Glaser and Strauss, 1967) which stresses the active interplay between collecting data and generating theory, rather than having a predetermined theory and going out to test it. Glaser and Strauss (1967) described the processes in such a grounded theory approach as follows: important categories may emerge from analyses of the data (text in this investigation), such categories are then pursued and made firmer through further research. The research could be described as "interpretive" since one had to discern and articulate subtle regularities within the data. Thus, detailed descriptions of context and what the participants said or did formed the basis for inductive and aductive rather than deductive forms of analysis. Primary activities in this research mode were the reduction, organization, manipulation and display of the data, combined with the generation of hypotheses about cognitive structures and processes which can explain the data.

The research activities in the proposed study were interpretive (in the sense described before) and formative, in that the data generated and refined hypotheses about learning mechanisms. A problem in educational research is that hypotheses oriented to statistical testing are not complex or cognitive enough to be sufficiently insightful and provide opportunities of giving an explanation of the most important processes. Case study methods allow one to generate more insightful, structural (as opposed to empirical) hypotheses. Such hypotheses can then suggest separable, simpler, and more testable hypotheses and experiments in a later stage, as well as principled instructional strategies. The hypotheses generated by the case study also provided existence proofs for key learning processes (J. Clement, personal communication, October 1989).

3. Data Analysis

The interviews were transcribed, summarized and observations described. The summaries were in the form of "maps" of the students' progress through the interviews in which the explanations were coded and categorized at different levels. The coding, e.g. when an explanation was considered to be a misconception, followed after discussions with other researchers. The students' conceptions will be described in the qualitative analyses in each experiment.

An excerpt from a "map" of one student's progress over the first part of the lesson is given in Figure 2.2. The abbreviation $m.c_c$. denotes a misconception; "sharing" conveys the idea that two people share a load in a symmetrical carrying situation; and "fulcrum-helps" indicates the model that students constructed of the fulcrum as "helping" in holding a load level. A description of the map follows below Figure 2.2.

Knowledge Categories	Reasoning in a Lesson Sequence				
Principles					
Models		sharing	fulcrum helps	fulcrum helps	fulcrum helps
Non-generalizable Conceptions					
Misconceptions	$m.c_c$.				
	Target	Anchor	Bridge	Extremes	Target

Figure 2.2

Mapping a Student's Progress

The map shows that the student apparently changed his misconception about the target question, and perhaps as a result of his participation in the bridging sequence, since his answers for the rest of the sequence are correct. His explanations include a sharing model and a fulcrum-helps model. The issue concerning the change in the student's misconception can be resolved by a more in-depth analysis of the students' explanations, but the purpose of the maps should be clear: they provided a qualitative overview of the development of the students' ideas over time and could be regarded as the first level of analyses at which data was generated and hypotheses emerged.

4. Participants

Seventh grade students (none of whom have had any formal instruction in physics) at Amherst's Regional Junior High School participated in the research. All the seventh grade general science students (N = 60) were asked to complete the diagnostic test and 12 refused. Twenty-eight of the students diagnosed as holding misconceptions were approached for participation in the interviews. The teacher was asked to judge these students' "ability" in terms of their understanding of science concepts and eight were included in each of three categories labelled "high", "average" and "low" conceptual ability. The students were randomly assigned to the three experimental groups.

C. Limitations of the Study

The artificial interview context probably represents the most important limitation of this study. There is little resemblance between the classroom environment and video- and audiotaped clinical interviews about abstract and novel lever problems. Although I have no intention of generalizing to more traditional learning and teaching environments, one may reason that the clinical interview context puts the researcher at a disadvantage. If students can maintain interest and remain motivated for two hours under such conditions, one could reasonably expect similar and perhaps better results in a "normal" classroom context.

In addition, the groups are small, and any empirical results should therefore be regarded with caution.

Ideally the protocols should have been coded by more than one person to allow for some interscorer reliability.

D. Definitions and Explanation of Terms

1. Anchoring Conceptions

Conceptions are called anchoring conceptions (or briefly, anchors), when new content knowledge can be "anchored" in a student's intuitively correct conception. An anchoring conception is defined theoretically as an intuitive knowledge structure which is in rough agreement with accepted physical theory, where intuitive refers to self-evaluated knowledge - the strength of the student's belief is not determined by appeal to an outside authority, but by himself (Clement, 1988).

A physical problem situation is considered to be an *anchoring example* if a student's response is correct and accompanied by a high confidence in his answer; thus the anchoring example is considered to be a source of evidence for the existence of an anchoring conception in the student's mind (Clement, D. Brown and Zietsman, 1989). The anchors that were used in the study were general or common to the whole group of participants. In other words, the instruction was grounded in group anchors rather than individual anchors. Group anchors can be useful in instruction: for example, many students refuse to believe that static objects can exert forces, but they do believe that a spring will exert a constant force on a person's hand as he holds the spring compressed. This intuition about springs can be built upon as an anchor when teaching that inanimate objects can exert forces (D. Brown and Clement, 1987).

The following sub-category of anchoring examples is important to this study: *symmetrical* refers to the essence of this type of anchor - that is, all variables in the system that are important to the students are in symmetrical relationships. For example, given that a load of 20 lbs is held level in the center of a light, strong board, each hand will exert a force of 10 lbs to keep the system in equilibrium.

2. Extreme Cases

These are situations where one of the variables in a physical system is taken to a limit. For example, in one of the lever situations in the class II teaching sequence, a load is placed on the board as near to the fulcrum as possible without actually being on the fulcrum, thus setting both the leverarms at extreme values. The principle of levers should still apply to the extreme case situations.

3. Benchmarks

A benchmark is a specific extreme case, where the variables have exact values and for which a person has an exact quantitative answer. For example, for a seesaw type lever, a load of 20 lbs will require a force of 2 lbs to keep it in equilibrium if the load leverarm = 1 ft and the effort leverarm = 10 ft.

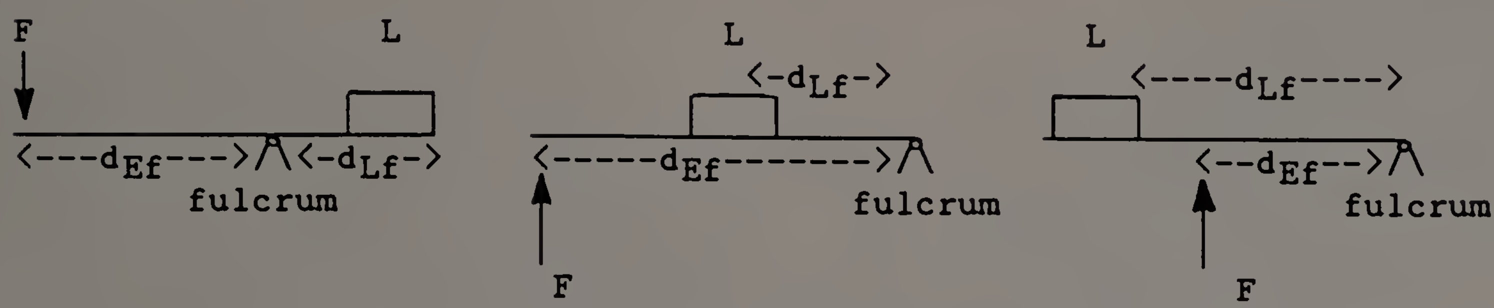
4. Levers

a. Principle of Levers. The principle of levers can be stated as follows: if a force, usually referred to as the "effort", is applied by pushing or pulling on one end of a lever, the lever swings about the fulcrum to produce a useful action at another point. The fulcrum could therefore be described as the "turning point" in a lever. The lever moves to raise a weight or overcome a resistance, both called "loads". The point on the lever where the force is applied is just as important as its magnitude.

The principle of levers, that relates the effort (E) and the load (L), states that the force times its distance from the fulcrum (d_{Ef}) equals the load times its distance from the fulcrum (d_{Lf}), that is:

$$F \times d_{Ef} = L \times d_{Lf}$$

Three classes of levers are distinguished by the relative positions of the applied force, the fulcrum and the load, as shown in Figure 2.3.



(a) Class I (b) Class II (c) Class III

Figure 2.3

Three Lever Classes

Class I levers have the fulcrum placed between the effort and the load, and the effort is usually magnified in this lever class.

Class II levers have the fulcrum at the one end of the machine and the force applied at the other. The load is somewhere in between. Since the distance from the fulcrum to the effort is greater than that from the load to the fulcrum, multiplication of the effect of the force still occurs.

Class III levers have the fulcrum at the end of the machine, but the positions of the effort and the load are reversed. The load to be raised or overcome is always at the one extreme end of the machine, while the effort is applied between the fulcrum and the load. A third class lever magnifies the distance moved.

b. <u>Symbols for Lever Elements</u>. The representations used in Figure 2.3 will be used throughout the study, thus

 represents the fulcrum, hinged to the board;

20 indicates a load of 20 lbs;

$\uparrow$ or $\downarrow$ denotes the force exerted (usually by a person); and d_{Ef} and d_{Lf} always refer to the two leverarms; that is the effort-fulcrum distance and the load-fulcrum distance respectively.

The drawings are almost exact representations of the equipment used in the experiments. The board on which the load rested was always hinged to the fulcrum, to prevent students' concerns (expressed in the pilot studies) with superficial effects such as the board "jumping off".

5. Models

Instructional research studies are increasingly concerned with the teaching of meaningful conceptual models to facilitate learning. One may view the models as advance organizers that provide meaningful assimilative sets (Mayer, 1975), and the benefits of such models are said to range from the improvement of retention to the improvement of transfer abilities in students. The effects of conceptual models on physics learning are being investigated in conventional instructional settings (e.g. D. Brown and Clement, 1989; Mayer, 1989) and interactive computer learning environments (e.g. Smith, Snir, Unger and Grosslight, 1990; White, 1990).

It is useful to explain how the term "model" will be used in this study, although as with all definitions and models in this study, it will probably be amended by the end of it.

An important characteristic of the explanatory models is that one expects the students to construct them. The instruction is aimed at that construction - all the situations are carefully designed to facilitate or trigger the conceptual building blocks, but the model as such is never presented to the children.

In general, the models students may construct in this study, will be described as "intuitively anchored". This means that the model is grounded or anchored in a physical intuition held by the students. A non-observable mechanism is hypothesized by the students to underlie and explain the physical situation (J. Clement, personal communication, April 1990).

6. <u>Near and Far Transfer</u>

One of the purposes of this study was to investigate transfer from the target conceptions in the lesson to real physical phenomena; in other words, whether the instructional sequences were understood by the students to the extent that they could apply this knowledge to other situations. This is in accord with Royer's (1987) suggestion that the ability to transfer newly learned information could be considered an index of understanding.

Transfer situations used in the posttests and within the lesson include instances of near and far transfer, for example:

<u>Problem in Lesson</u>

What force has to be exerted by the man to hold the board with the 20 lb on it level?

<u>Transfer Problem</u>

What force has to be exerted to hold the load in the wheelbarrow up?

Figure 2.4

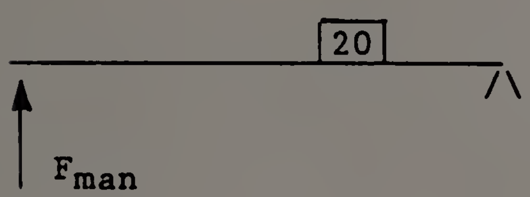
Near Transfer Problem

Near transfer situations would be in contexts perceptually the same as the target situations.

Far transfer situations would be in different perceptual contexts from that of the target situation (e.g. in Figure 2.5).

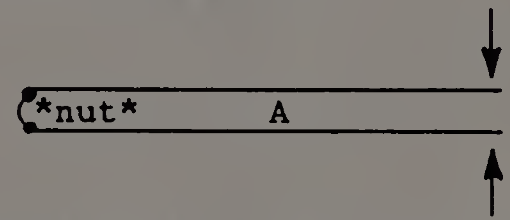
<u>Problem in Lesson</u>

What force has to be exerted by the man to hold the board level?



<u>Transfer Problem</u>

What force has to be exerted to crush the nut?



[Given that a 20 lb force is needed to crush the nut in A]

Figure 2.5

Far Transfer Problem

The example in Figure 2.5, a nutcracker, has two class II levers, one fulcrum (the hinge) shared by the two levers and a force exerted on each of the two levers, but in opposite directions. This machine is therefore not perceptually similar to the simple class II levers in the teaching sequence. Additional transfer problems that depict complex levers at work in the real world were included as far transfer problems.

The transfer problems in the study were not used in the same sense as Bassok and Holyoak, that is where "transfer is simply the result of applying information about a known category to a new instance" (Bassok and Holyoak, 1989:159). Rather, one of the purposes of this study was to find out what students learn from the anchor-bridging analogies and the extreme case examples, and how what they may learn could be transferred to the novel machines in the transfer problems.

7. Normative

The word "normative" will be used to indicate an ideal state, e.g. a normative understanding suggests an understanding that is aligned to currently accepted physical theory. In the International Dictionary of Psychology (Sutherland, 1989) the following meaning is given:

> normative 1. Setting principles or standards of how people or other systems ought to behave, e.g. both GAME THEORY and SIGNAL DETECTION THEORY are normative since they describe how an ideal system would behave.

E. Advance Organizers From the Pilot Studies

1. The Pilot Studies

The major goals of the preliminary research were to develop a test to diagnose students' misconceptions in the content domain and to establish, by means of diagnostic testing and clinical interviewing, anchoring examples in the content domain. Finally, the interview and test results were used to develop instruction about levers.

The point of departure in the development process was the design of a diagnostic test with input from teachers, physicists and researchers in science education. The test was administered to 32 participants and informal test interviews were conducted with 8 students. On the basis of information gained, the diagnostic test was revised, a lesson about levers developed and a pilot study conducted to evaluate both the revised test and the new lesson. The pilot study's diagnostic test was administered to 34 seventh grade students, twelve of the students holding misconceptions about levers were interviewed about the diagnostic test and four of those interviewed participated in tutoring interviews about levers.

The results of the pilot study were used to revise the diagnostic test and the lesson again. The third version of the diagnostic test was administered to 28 seventh grade students and three students participated in tutoring interviews in a final revision process preceding the levers study.

The cyclical process of design of instruments, evaluation, revision and redesign is in accord with the philosophy underlying this research; that is, that the researchers should be informed repeatedly by the participants, and that hypotheses are therefore generated, not just evaluated during the research processes.

2. Some Results from the Pilot Studies

a. Anchoring Examples. Several anchoring examples were identified in the pilot studies, and one, the situation with two people or two hands holding a light, strong board with a load in the center level, was used in the pilot lesson. Participants in the interviews gave acceptable, naive explanations for this anchoring example, e.g. that the load would exert equal forces down on the two people holding the board.

b. An Extendable Anchor. Students' knowledge of the anchoring example appeared to be extended to other lever situations in parts of the lesson. One may infer this since students who could not perform a simple task about levers before the lesson could answer the same question correctly at the end of the lesson. Second, they could explain their correct answers and third, these explanations related to the anchoring example.

c. Naive Conceptions. The results of the pilot study's diagnostic test and the interviews suggested that some students were able to formulate their own, naive conceptions of "how levers work". Although these conceptions do not nearly convey the same physical ideas as the principle of levers, students were able to predict correct answers to problem situations.

The naive, but qualitatively appropriate conceptions indicated that the children formulated at least three identifiable conceptions instead of the one principle of levers stated before, namely:

The Fulcrum Acts as an "Active Aid." This conception proved to be particularly helpful in the lesson, namely the idea that the "fulcrum helps", e.g.: the wheel in a wheelbarrow helps one lifting the load; the fulcrum in the crowbar "takes some of the weight", and the table "holds" part of the weight resting on the board. It appeared that this belief allowed the students to conceptualize the force they were asked to estimate for different points on a lever. Although they did not use the term leverarm, they seemed to be focusing on these variables and were able to estimate the forces' magnitudes by using their naive conceptions and their "fulcrum helps" explanation.

A Conception for Class II Levers. These conceptions related the load-effort separation and an increase in effort, for example: "the further you (effort) are" from the fulcrum, "the easier". This conception was also applicable to some class I problems and the students seemed to be consistent in their applications of the idea.

A Special Rule for Class III Levers. There was enough evidence from the protocol analysis to hypothesize that the students were using a different idea to explain the way class III levers work, namely "the nearer the effort is to the load, the less effort is required". This conception seemed to be more rule-like in character, in the sense that no plausible explanation supported the idea. Closely related to this idea was their notion that "the closer the effort is to the load, the more control" the person exerting the effort will have over affairs.

d. Categorizing Levers for Instruction. There was scant evidence from the pilot study for the importance of the categorization of levers in the three different classes to enhance students' understanding of the subject matter, and it was suggested as a conjecture. The results indicate that the students used different conceptions to reason about class II and III levers. In transfer type questions, however, levers were categorized purely on the basis of their functions, hence evidence that confounded the suggestion that students may view at least type III levers as fundamentally different from the other types. This could be an important issue, since traditional quantitative instruction in levers tends to present just a general principle of levers, assuming that students will be able to transfer the general principle to all kinds of lever-like situations. Thus, if it is true that students view lever types as being different from each other in some important sense, one may suggest that such a finding would be useful in developing instruction that takes account of the students' naive conceptions.

F. Instruments: Pre- and Posttests and Lessons

The test questions were administered at least three times to different participants to iron out problems with language and students' comprehension of the lever situations and drawings. Thus, although the tests were not formally validated, one may be reasonably sure that the range of questions should draw out naive or misconceptions, or in the case of the far transfer questions, suggest an understanding of the principle of levers to explain the phenomena satisfactorily.

The pre- and posttest questions will be described in the course of the discussions on the summative evaluations for experiment 1, while in depth discussions of the lesson's three sections will precede the formative evaluation. I have decided on this format since one will have to refer to the different problems frequently within the evaluations and it seemed more sensible to keep the descriptions and references as close together as possible. The complete pre- and posttest are attached as Appendix C, and the explanation sections of the two lessons as Appendices D and E.

The diagnostic test combines most of the problems that appears in the pre and posttests and in the lesson. The problems are not reviewed separately, since I shall not discuss students' performances on this test in detail, but only in terms of overall scores. The diagnostic test is attached as Appendix B.

CHAPTER III

EXPERIMENT 1: RESULTS AND DISCUSSION

A. Introduction

The research design for Experiment 1 is similar to a conventional control/experimental group approach. A traditional summative evaluation is performed in the comparison of pretest and posttest scores of the experimental and control groups. However, the evaluation of the instructional intervention is primarily formative. The formative evaluation should generate structural hypotheses about the effects of lesson segments, particularly with regard to the processes of knowledge construction that were facilitated by the lesson's design and underlying philosophy.

The report of experiment 1 is organized into five sections. First, the *diagnostic test results* are summarized. Second, a discussion of the *students' preconceptions* is followed by a *summative evaluation* in two sections: a pre-posttest comparison of the control group and experimental group responses to the simple or generic levers questions, and to the far transfer problems. The *formative evaluation* of the lesson is then presented, and finally *general findings and recommendations* for the changes to experiment 2's lesson.

B. Transcript Conventions

Excerpts from the students' protocols will often be given, as illustrations and for clarification. The following conventions apply throughout the discussions of results.

/	indicates a pause shorter than 2 s
//	indicates a pause longer than 2s but shorter than 4s.
...	phrase from a larger quotation
{	indicates simultaneous speech
[text]	researcher's notes or observations (noted during live interviews or while analyzing video tapes)
S:	student
I:	interviewer

C. Diagnostic Test Results

The diagnostic test was administered to 48 seventh grade science students at the Amherst Regional High School. The classes were all taught by the same teacher and none of the students had been taught physics in any formal sense.

The results are given in Table 3.1. The belief score is the percentage of students who answered correctly with high confidence; the lever class is indicated; and d_{Lf} and d_{Ef} refer to the leverarm varied in the question.

The primary purpose of this written test was the identification of students who held misconceptions in one or more of the questions. Twenty-eight students were diagnosed as holding misconceptions about at least two questions in the test. The misconception scores were much higher than anticipated; for example six students predicted only two of the questions correctly. The high scores on the class III lever questions were expected: the pilot studies indicated that students have a definite "rule" for predicting answers to this type of problems, although they were not able to formulate a satisfactory explanation to accompany their predictions.

Table 3.1

Diagnostic Test Results

Question	% Correct	Belief Scores
1. Class I, d_{Lf}	83	79
2. Class II, d_{Lf}	67	59
3. Class I, d_{Ef}	41	33
4. Class II, d_{Ef}	43	31
5. Class III, d_{Ef}	72	60
6. Class III, d_{Lf}	79	64
7. Anchor 1, Class II	64	47
8. Anchor 2 Class I	87	72
9. Anchor 3 Class II	79	69

The belief score for question 7's anchor 1, the symmetrical situation where two hands are holding a light strong board with a 50 lb load in the center level, was lower than a set limit (belief score criterion: 70%, [Clement et al., 1989]). As a result, both anchors 1 (question 7) and 3 (question 9) have been included for the class II lever teaching.

D. Students' Preconceptions

The preconceptions stated below are interpretations of the students' statements in the protocols of the pre- and posttests, as well as in the lesson's target questions. Since conceptions in the posttest cannot in fairness be referred to as preconceptions, only those instances where posttest explanations support the preconception description and are particularly illuminating, will be mentioned.

1. The Control Misconception

This misconception was noted in the pilot studies and the instruction used in this experiment is aimed at changing the conception. The misconception was seen as a need to be near the load, to prevent it from falling, or to "wobble" etc. A typical student explanation would be:

> 005 S5: You just have more control over it [the load] if you're holding it closer to you.

The analyses in this experiment have added some ideas to this "control" aspect. It may be that students are not only concerned with control in the manner described above, i.e that the load may fall, or wobble or "tip over"; but that some may be holding a naive view of forces exerted in the situations. I shall describe these as "naive force diagrams". It should be obvious that the student may not have the actual diagram in his "head", but the force diagram idea presents the most succinct explanation of the conception.

Suppose a student has to consider the situation in Figure 3.1 below and has to decide where (at A or B) it would be easier to hold the board with the 50 lbs load on it level.

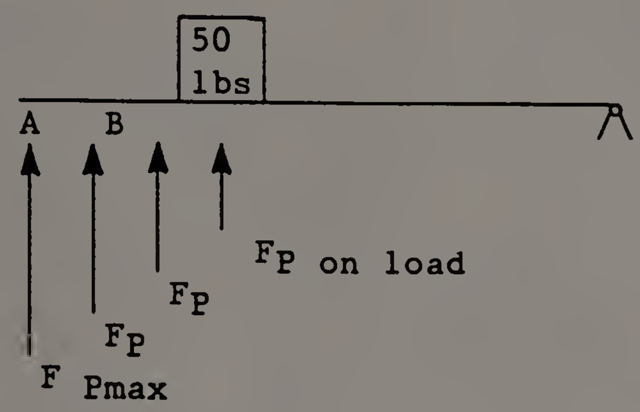


Figure 3.1

Control Misconception Force Diagram

Students explain that the person exerting the force (at A) has "less control over the weight in position B", and many indicate that this is because he can exert a maximum force only at the point where he is pushing. Thus, were the load in position A and the person pushed up at A, he would be able to exert that maximum force F_{Pmax} and hence have more control.

If however, as in the case above, he pushes at A and the load is at B, only some of this maximum force is exerted on the block, resulting in less control and making it more difficult to hold the block level than when the load was positioned at A, for example, as control students C4 and C6 explained:

> 024 C4: I think it'd be easier in case B ... Because once again you're closer to the load...If there's less room, it'd be easier because it's more direct [referring to force?].

010 I: Can you say a little bit more / You said B is closer?
011 C6: Yeah, because of //
012 I: Closer to what?
013 C6: I think you would need it more to be around this area [pushing up motion with hands].
014 I: Where you would be pushing up?
015 C6: Yeah, it may be a little better.

S3 explained that:

020 S3: I think it'd be easier right here because the pounds, the weight would be closer to you and you can keep it up easier...

And S5:

010 S5: I think B, because it's shorter / It's [load] closer to you ... // [at] B it's closer to you, so if you're holding it, it would be easier, because the weight...is right there.
108 : The load is closer to where you're holding the board level...because it's [A] kind of far away,...you have to kind of *hold* it more, because it's so far [student's emphasis].

I infer from the excerpts that there is the notion of a "dissipating" force in these explanations, as represented by the decreasing lengths of the arrows in Figure 3.1. Some examples that may be better illustrations of the dissipating forces idea were observed in the posttest protocols.

The revolving door problem, where two persons are pushing equally hard from opposite sides, but in different positions, elicited the following explanation from C4:

005 C4: I don't know if it [door] would move or not. Because if they are both pushing equally hard //
006 Uhmmm // I think if it did move any way at all, it would move counter-clockwise / Because Beth is pushing right near the middle of the door / And // And Ann is pushing right near the edge of the door, so / The edge of the door // You'd move quicker because it's the edge of the door.
007 Uhm / Because / What we're trying to move is the *edge of the door* so that we can get out, and *it's easier to push from the edge*, here, than it would be to push from the middle where Beth is [my emphasis].

I interpret her statements as a "dissipating force" idea, thus that the nearer one is to what you're trying to move (for C4 the edge of the door), "the easier it would be"; since the force is applied exactly where it is needed - none of it is "wasted". This idea is conceivably a major basis of the control misconception.

Nutcrackers. In a comparison of two nutcrackers with different effort-fulcrum distances (the one has longer handles), some of the students held the control misconception. The explanations include some of the best examples of the notion of a dissipating force encapsulated in the control misconception:

> 010 C4: I think it would be easier in B, because there is less that you have to move. You just have to move what's directly in front of you // Not the whole sticks, handles, whatever.

> 011 C5: This [B] is shorter, when you push down on the handle, the nut's right there [pointed to where effort is applied].

> 016 C6: The pressure that you're applying is closer to the nut...

> 013 S1: When I crack a nut I usually like to get my hands close to the nut.

The more complex nature that the control misconception may have for some students could influence the teaching in a manner not foreseen at the design stage of Experiment 1, thus an issue to keep track of in the evaluation.

2. The Non-Generalizable Conception

There seems to be an interesting shift of the students' focus in the non-generalizable conception, when compared to the control misconception. I infer that students appear to be considering a dissipating effect of the *weight* of the load on the force exerted by the person, rather than focusing on the person's force (as described in 1 above).

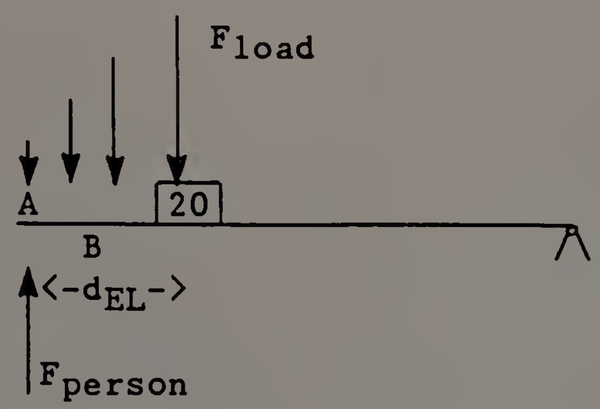


Figure 3.2

Force Diagram of the Class II Non-Generalizable Conception

In Figure 3.2 the force exerted by the load on the person, F_{load}, is not "all *on*" the person (or not "direct", another description used by students). Thus, the weight of the load is "distributed" over d_{EL} (the effort-load distance), and the greater d_{EL}, the smaller F_{person}. The converse of this statement is also applicable: the nearer the load is positioned to the person, the more "direct" the load's weight is on the person and the less "board" there is to have the load's weight "distributed" over.

Examples from the protocols to illustrate this interpretation are given below:

020 C5: I think it would be easier to hold at [A] because it's not so much out *on* you [students' emphasis].

037 S4: If you pushed up here [B], it'd be more weight on your side / Pushing down more.

037 S3: It's easier to pull up when the load's closer to you...

124 S6: The longer one, the more distance you're away from the weight...the less you have to...push.

The students' statements about the non-generalizable conception suggest a consideration of two forces: the person's and the weight of the load. This view is closer to that of the physicist, who would consider the force exerted by the fulcrum in addition to the two already mentioned.

The lesson in experiment 1 was designed to build on this non-generalizable conception. This view predicts the correct answer for a class II lever and is not a misconception in the strongest sense. Again, one should consider that this more complex analysis of the class II non-generalizable conception may have implications that were not considered during the design of the lesson.

E. Summative Evaluation

The students were interviewed about seven questions in the pre- and posttests: three simple lever questions (in both the pre- and posttest) and four far transfer questions in the posttest. The interviews were conducted over two days: the pretest and sequence 1 of the lesson on day one and lesson sequences 2 and 3 and the posttest on day 2.

One could hypothesize that the experimental group students should hold posttest conceptions that are more compatible with a physicist's principle of levers than the control group students' postconceptions. Learning may, however, also occur in either the pre- or posttest and there is always the possibility that a child may learn about levers in his everyday life. Several issues arise when one has to decide about criteria for the evaluation of learning. First, it is not sufficient to look only at the correctness of students' answers to the test questions, since the pilot study data indicated that students are likely to hold a non-generalizable conception which gives correct answers to all class II lever problems; but is considered a misconception when applied to class I and III levers. Second, changes in students' responses to the simple lever question (pre- to posttest), are probably not sufficient to indicate an understanding of levers. First, no transfer is required from the pre- to the posttest situations (they are the same), hence this seems more like an issue of consistency in students' use of conceptions; and second, the situations depicted in these question are abstract, schematic levers - thus, no transfer to real levers is required. The far transfer questions in the posttest depict real, complex levers and meaningful explanations to these questions are believed to indicate a greater depth in students' understanding of levers.

With the issues above in mind, the summative evaluation will be discussed in terms of students' performances on the simple lever questions and the far transfer problems. The pre- and posttest questions and expected answers, based on interviews and data from the pilot studies, are provided as an orientation to the discussion of the test results in each section, and this is followed by qualitative analyses of the results.

1. Simple Levers: Quantitative Analysis

The levers are referred to as "simple", since the apparatus in the drawings contains only three essential elements of a lever, namely an applied force (the *effort*, indicated by an arrow) to move a *load* (the 50 lbs) around a fulcrum (a triangle in the class I lever and tables in the other two questions). The three simple levers questions are given in Figure 3.3.

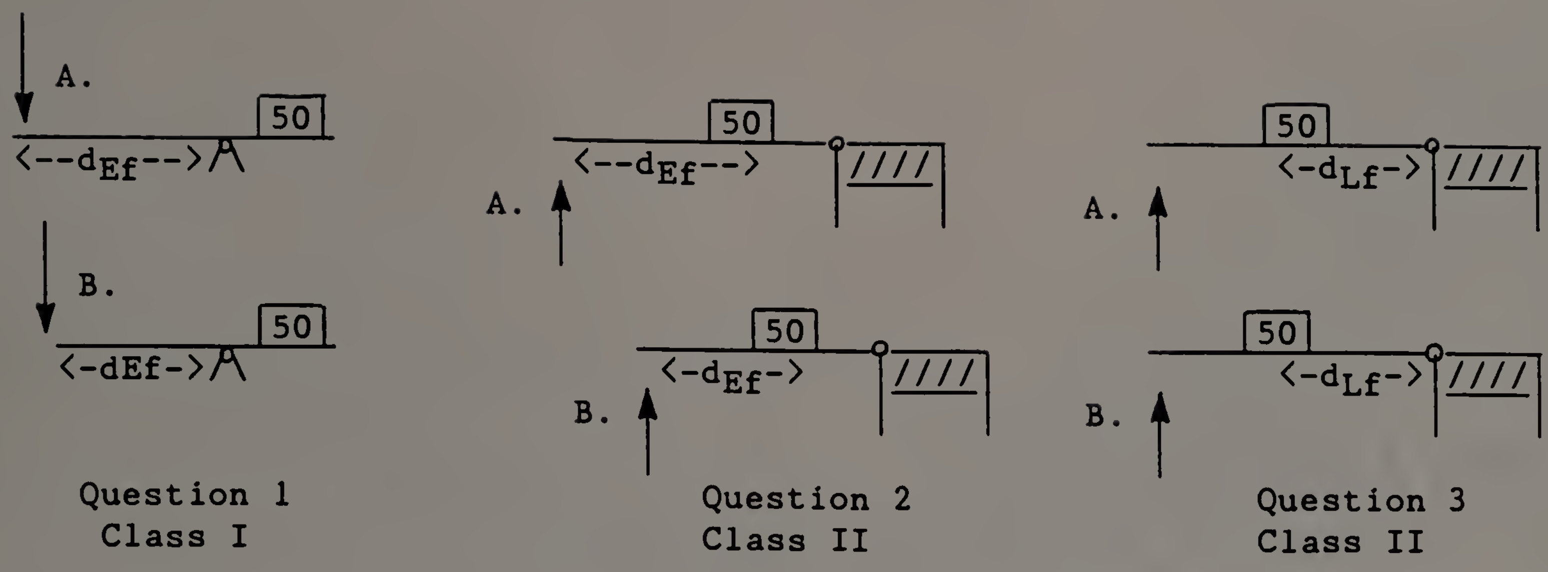


Figure 3.3

Pretest / Posttest Simple Lever Questions

In all three problems the board is hinged to the fulcrum, and students are asked to imagine strong, inflexible, very light boards. Students have to compare the two levers in each question and decide where it would be "easier" to hold the board with the 50 lb load level.

These simple levers questions are the questions with the lowest correct scores in the diagnostic test. The scores are given in Table 3.2.

Table 3.2

Diagnostic Test Scores: Simple Lever Questions

Diagnostic Test Question	% Correct	Belief Score
3. Class I; d_{Ef}	41	33
2. Class II; d_{Ef}	43	31
4. Class II; d_{Lf}	67	59

The absence of class III levers in the pre- and posttest may seem like an omission. There are however, class III levers in the far transfer questions. The expectation is that students who have attained an understanding of the principle of levers would be able to analyze the complex and compound class III levers in those situations. In contrast, students who hold the non-generalizable conception, indicative of a limited understanding, would probably focus on surface features in their explanations. Thus, an evaluation of students' responses in the far transfer questions would probably provide the most conclusive evidence of learning about class III levers.

There are obvious limitations with regard to a quantitative analysis of the data in this project. The groups are small and one cannot assume that the population is normally distributed. The students were selected from a larger group diagnosed as holding misconceptions in an area of physics. However, since the students were randomly assigned to the three groups, the experimental conditions were randomly assigned to the groups and the populations were continuous, one may use assumption-free tests on the predominantly ordinal data available.

Scores, obtained from students' answers and their confidence in the answers, were computed by: one, assigning positive and negative values to correct and wrong answers respectively; two, assigning a number (1 to 4) to the confidence level (rated from "a guess" to "sure" on a four point scale); and three, multiplying the confidence level number with the appropriate symbol to indicate a correct (or not) answer. Thus, a student who guessed a wrong answer would score -1 on a question, whereas a student who was sure that he was right about a wrong answer, would score -4.

In Tables 3.3 and 3.4 the experimental and control group scores on the simple lever question in the pre- and posttest are given.

Table 3.3

Simple Lever Questions: Pre- and Posttest
Experimental Group

Students	Pretest Question			Posttest Question			Summed Changes in Scores
	1	2	3	1	2	3	
S1	-3	+1	+3	-3	-3	-3	-10
S2	+3	-3	+4	+4	+4	+4	+8
S3	+4	+3	-3	-4	+4	+4	0
S4	+3	-3	+3	-3	+3	+3	0
S5	+3	-3	-3	-2	-2	-2	-3
S6	-3	-3	-3	-4	+4	+4	+13

Table 3.4

Simple Lever Questions: Pre- and Posttest
Control Group

Students	Pretest Question			Posttest Question			Summed Changes in Scores
	1	2	3	1	2	3	
C1	-1	-1	+1	+1	+1	+1	+4
C2	+4	+4	+4	+4	+4	+4	0
C3	+3	+3	+3	+3	+3	+3	0
C4	-3	-3	+3	+3	+4	+4	+14
C5	+3	-3	+3	-3	-3	+3	-6
C6	-3	+3	+3	-3	+2	-3	-7

The changes in scores suggest the differences in understanding from the pre- to posttest, and the two groups are compared with respect to the summed changes in scores for each student.

The Mann-Whitney U-test was used to test the hypothesis that the control and experimental students were identical with respect to their performances on the pre- and posttests. A null-hypothesis is accepted, since at a set level of significance ($p < 0.05$), $U_{calculated} = 16$ has a probability of occurrence under H_o of $p = 0.41$.

It seems, from the pre-posttest data for the simple levers alone, that the instructional intervention had no measurable impact on the experimental group students' understanding of levers. It may even be that some of these students were adversely affected by the instruction. A detailed qualitative analysis should provide more evidence towards these findings.

2. Simple Levers: Qualitative Analysis

There are several ways to define changes in the students' ideas from the pre- to the posttest questions. Firstly, one could consider students' changes in explanations for each question from the pre- to posttest situation. It seems that this approach would generate "fragmented" data, thus too specific to contribute any meaningful inferences about a change in a student's ideas about levers in general. I shall therefore only consider pretest to posttest question changes for one question where the analyses suggest an interesting phenomenon. Here one would focus on changes across students, i.e. how many students changed their conception from question $1_{(pre)}$ to question $1_{(post)}$.

Second, one could consider changes in a student's conceptions of a lever class. Thus, a student has reached a different understanding when the same conception (but different from the pretest explanation) is used to explain the class II lever questions in the posttest. This clearly requires a "within student" protocol analysis.

Third, one may consider an overall change in a student's explanations. Thus, by definition, comprehensive or overall conceptual change would have occurred when a student has consistently used a non-generalizable conception to explain all three questions in the pretest, compared to the consistent use of a misconception in the posttest explanations. Again, a description of such conceptual changes requires a "within" student analysis of protocols.

Finally, *normative* conceptual changes, that is, comprehensive conceptual changes towards the physicist's view of levers in each question in the posttest, will be seen as instances where learning has occurred, as a result of: one, participation in the lesson; two, the pretest interview; three, the posttest interview; or, finally, if the above cannot be established from the protocols, as a result of external interactions.

All these criteria will be used in the data analyses.

a. Summary of Results. A summary of the students' responses to the pre- and posttest questions are given in Tables 3.5 and 3.6. The abbreviations will be used throughout the report of data, and are to be interpreted as follows:

- or + - correct/wrong response

$m.c_c.$ - control misconception

$m.c_s.$ - symmetry misconception

non-gen - non-generalizable conception

$d_{Ef} >$ - effort-fulcrum distance greater

and } principle of levers

$d_{Lf} <$ - load-fulcrum distance smaller

model - the "fulcrum-helps" model

[ ] - the same explanation for consecutive questions

For example (from Table 3.5), S1 gave a wrong answer to question 1 and used the control misconception in her explanation, she guessed a correct answer for question 3 and explained her correct choice for question 3 using the non-generalizable conception.

Table 3.5

Responses of Experimental Group Students

Student	Pretest Questions 1	2	3	Posttest Questions 1	2	3
S1	− m.c.$_c$	+ guess	+ non-gen	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$
S2	+ non-gen	− m.c.$_c$	+ model	+ $d_{Ef} >$	+ [$d_{Lf} <$; model	+]
S3	+ $d_{Ef} >$	+ non-gen	− m.c.$_c$	− m.c.$_c$	+ non-gen	+ non-gen
S4	+ $d_{Ef} >$	− m.c.$_c$	+ non-gen	− m.c.$_c$	+ non-gen	+ non-gen
S5	+ non-gen	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$
S6	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	+/− $d_{Lf} <$	+ [$d_{Lf} <$ and $d_{Ef} >$;	+ model]

Table 3.6

Responses of Control Group Students

Student	Pretest Questions 1	2	3	Posttest Questions 1	2	3
C1	− m.c.$_c$	− m.c.$_c$	+ non-gen	+ non-gen	+ non-gen	+ non-gen
C2	+ non-gen	+ $d_{Lf}<$	+ $d_{Lf}<$	+ $d_{Ef}>$	+ non-gen	+ non-gen
C3	+ $d_{Ef}>$	+ non-gen	+ non-gen	+ $d_{Ef}>$	+ non-gen	+ non-gen
C4	− m.c.$_c$	− m.c.$_c$	+ non-gen and model	+ non-gen	+ non-gen	+ non-gen
C5	+ $d_{Ef}>$	− m.c.$_c$	+ non-gen	− m.c.$_s$	− m.c.$_c$	+ non-gen
C6	− m.c.$_c$	+ non-gen	+ non-gen	− m.c.$_c$	+ non-gen	+ non-gen

b. Comprehensive Normative Conceptual Changes. Inspection of Tables 3.5 and 3.6 indicates four cases of normative conceptual change; two in the experimental group (S2 and S6) and two in the control group (C1 and C4). As stated before, the learning may have been facilitated by one (or a combination of) the following factors: students' participation in the pretest interview, the lesson interview, the posttest interview or external factors (say ordinary experiences in the real world).

Learning in the Lesson. In the experimental group, students S2 and S6 changed to consistent applications of a qualitative principle of levers, supported by a fulcrum-helps model. S2 had a mixed bag of conceptions in the pretest, while S6 consistently applied the control misconception. While one may speculate on the amount of learning that occurred in S2's pretest interview, it is probably safe to say that S6 could only have come to the understanding evident in his answers as a result of the lesson interview, since none of the control group students produced these types of posttest explanations.

Learning in the Pretest. The explanations of students C1 and C4 also suggest normative conceptual changes. It is difficult to determine the source of C1's change to a consistent use of a non-generalizable conception in the posttest, since she described all her answers and explanations as "pure guesses". However, it may be significant that she changed her mind during the explanation of question 3, the final class II lever question. This significance is inferred from the protocol analysis of C4 for question 3 (presented below), for whom this question provided a conflict that subsequently changed her mind on her previous answers. (Question 3 is represented again in Figure 3.4 below for reference.)

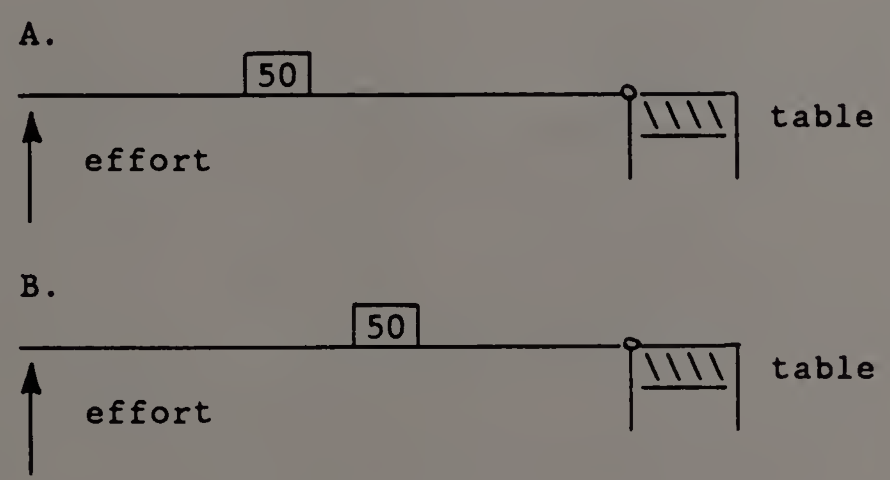


Figure 3.4

Question 3: Change in Load-Fulcrum Distance,
Effort-Fulcrum Distance Constant.

As for all the simple lever questions, students had to decide which board would be easier to hold level with the 50 lb load on it, given that the board is hinged to the table. C4 decided initially that B would be "easier to hold":

030 C4: I think once again it'd be easier in case B, because when the weight is farther away from the table itself then / it puts more stress on the board //
031 I: OK.
032 C4: And / Oh wait // Actually //
033 I: Uhhm?
034 C4: I think it'd be easier in A, because you're closer to the weight and uhm // like I said before with the seesaw // Things like that would be easier // Uhm / I'm not sure.
035 I: What is puzzling you now?
036 C4: // It seems that in B it'd be easier because it's [load] closer to the table, and so the table is holding more weight than you are.
037 But / In A, if it would be so, then it would change what I said for all the other answers / all the others.

When she became aware of a conflict between her (correct) answer for this question (line 032) and her previous answers, C4 repeated the control misconception. This may indicate an epistemological commitment (to consistent explanations [P. Hewson, 1985]), but her intuition about the table "holding more weight" when the load-fulcrum distance is smaller, seems strong enough to cast doubt on all her previous answers. A more extreme version of the comparison of levers in question 3 is used in the lesson to reinforce a "fulcrum-helps" idea established in the anchor-bridge sequence for class II levers, and it is therefore interesting that C4 constructed the model spontaneously from this single question. S2, in the experimental group, had an intuition similar to C4's, but apparently saw this fulcrum-helps idea applicable only to this specific question:

023 S2: I think it'd be B, because the block, the weight is closer to the table, which means that you have less to hold up, you have the table there and you are holding less...

In contrast, C1 did not suggest a "fulcrum-helps" model, only that "it might be easier if you had it [load] in the middle [of the board] than right towards the edge [as in drawing A]".

It is, however, interesting that these consistent, normative conceptual changes seemed to be facilitated by this particular question, resulting in the only observable instances of learning from the test.

Other Factors Facilitating Learning. It was also suggested that learning could have occurred as a result of external influences. However, the control group data shows that there is evidence for a better, although limited understanding of levers in the responses of students C1 and C4 only. As was shown, both students changed their conceptions during the pretest interview, with more evidence of learning in C4's protocol than in that of C1. I shall therefore assume that there are no discernible influences on learning except from the pretest interview in some cases or from the lesson.

c. Non-Normative Changes. There is evidence of an actual regression in one student's (S1) understanding of levers. S1 changed from a very tentative non-generalizable conception to describe the behavior of class II levers (pretest questions 2 and 3) to a consistent use of the control misconception. This regression can only be attributed to the lesson.

d. Limited or No Changes. Limited changes in the explanations of two students (S3 and S4) were observed. Both students used the non-generalizable conception to explain the posttest class II lever comparisons (from applications of this conception and the control misconception in the pretest). However, this is scant evidence for normative conceptual change; both students probably had very little or limited understanding of levers in general, as is indicated by the continued existence of the control misconception in their explanations for the class I question.

The lesson made no dent on S5's understanding of levers either; there is no change in her misconception about class II levers and in addition, she probably became more convinced of the misconceptions since she extended the conception's use to the class I lever situation. It seems at this stage of the analysis that S1, the regression candidate, and S5 may be the most interesting examples of the lesson's failure.

C2 (in the control group) appears to have the makings of a naive principle of levers. There is evidence (in both pre- and posttest) of reasoning about the leverarms. The naive principle appears to be used in an inconsistent fashion and it is not possible to get a clear idea of his overall conception of levers from the simple levers data alone.

e. Experimental Group: Class I. Even at this early stage one could suggest that the instruction on class I levers has failed. In the experimental group, three students held the control misconception in the pretest class I lever question and all but S2 and S6 held this misconception in the posttest question. The data shows that three students changed to this misconception, with one unchanged. No such changes occurred in the control group - students did not acquire misconceptions. Again, this is an early indication that something is amiss in the class I teaching sequence.

3. Conclusion

The pre- and posttest analyses of the simple levers questions provide three organizers for the rest of the evaluation.

Three students (one from the control group and two from the experimental group) used a qualitative principle of levers in the posttest. It should be interesting to compare these students' far transfer abilities and to determine the origins of their learning.

There is evidence that three students changed their control misconceptions as a result of the pretest. This is considered an interesting finding, since the question facilitated the use of a fulcrum-helps model in two of the three students involved. This one question seems to have produced the results expected from a substantial part of the teaching sequence for class II levers. The problem triggered the same intuition expected from the extreme case version. One may therefore anticipate interesting findings from the lesson's extreme case bridges.

The analysis of the simple levers questions provided little evidence about students acquiring an adequate understanding of levers. There are, however, some early indications of the intervention's very limited success in bringing about conceptual changes in students. The class I levers sequence seems to be a failure: all the experimental group students, except two ended up with the control misconception. This effect was not observed in the protocols of the control group participants.

4. <u>The Far Transfer Questions</u>

The four far transfer questions are difficult. The assumption in this evaluation will be that a correct answer, accompanied by an explanation approximating one required by the physical theory, would illustrate a "deeper" understanding of levers than was required in the simple levers questions. Successful solutions to these problems require transfer on the basis of deep, structural principles rather than naive conceptions or surface features (A. Brown, 1990). I propose that the knowledge structure that would be transferred to enable a person to solve these four questions, may be described as abstract and qualitative: abstract since the essence of a description of the behavior of levers is contained in the conception; and qualitative since a correct solution does not require any computation.

The levers lesson is designed to facilitate the construction of a qualitative, abstract principle of levers by the students. The principle should be grounded in an explanatory, causal model of levers, and this model should be constructed by means of analogical reasoning from anchoring conceptions and limiting cases presented in the lesson. One expects that the students who have constructed a qualitative levers principle via the processes outlined above, would be more successful in their analyses of the far transfer questions than those with a more limited understanding, e.g. students who used acceptable, but insufficient non-generalizable conceptions in their simple levers explanations.

The expectation of more meaningful transfer by students who hold both the causal explanatory fulcrum-helps model and the principle of levers, is in agreement with A. Brown's (1990) argument that children's ability to transfer on the basis of "higher level causal relations" rather than on the basis of surface features (lower level relations), is an indication of the depth of their understanding of the conceptual knowledge in the domain.

The discussion of the transfer issue will be organized as follows: first, a statement of the criteria to evaluate transfer; second, a quantitative overview and brief discussion of the results; third, a look at learning, in terms of transfer, across groups (control vs experimental); and finally, a qualitative discussion of children's thinking about the transfer problems.

a. Definitions: Transfer or Not? It seems necessary to distinguish between children's transfer of correct physical conceptions or principles, and the "transfer" of misconceptions and naive, non-generalizable conceptions. The following definitions are considered sufficient for the level of protocol analysis of knowledge transfer in this study.

Instances where there is evidence that students used misconceptions not apparent in their pretest explanations, thus where it looks as if they had acquired those conceptions in either the lesson or from external interactions in the period between the pretest and the posttest, will be referred to as *negative transfer*.

Instances of far transfer will be defined as those that provide evidence that students have a "deeper" understanding, i.e. that students have used the qualitative, abstract principle of levers in their explanations. The students could have acquired this "deeper" understanding as a result of their participation in the pretest interviews; their participation in the tutoring interviews; or other, external interactions.

Analyses of the student protocols may lend credence to the proposal that only students who participated in the lesson would have acquired both the principle and the causal model underlying the principle to allow far transfer. Protocols of experts solving two of the far transfer problems suggest that, at the very least, students will have to be able to recognize (and distinguish between) the levers in the machines before they could apply any principle. Lever recognition could simply be described as transfer on the basis of identical elements in the machines and the simple levers, i.e. the identification of a load, an effort and a turning point (fulcrum). However, from descriptions of the complex and compound levers in the far transfer questions (following in sections d. to g.), it should be evident that a fairly sophisticated structural analysis has to be performed to distinguish the levers in the machines. Thus, one criterion for transfer will be that students have to analyze the machines (i.e. break down into different lever components) and refer to these levers explicitly.

A student's ability to recognize levers in the machines, combined with the consistent application of an accepted conception to the problem, will be viewed as far transfer. The criteria for the evaluation of instances of far transfer is essentially normative in that the correctness of the solution and the acceptability of the explanation is taken into account.

The expectation is that instances of far transfer would only be evident in the protocols of students who have constructed a qualitative principle of levers grounded in a causal, explanatory model of levers. It should be possible to obtain evidence of the use of a principle of levers from students' explanations; but one will have to trace the development of the model by students during instruction to provide evidence that they have in fact constructed such a model. Thus: it may be impossible to illustrate the existence of the fulcrum-helps model from the posttest protocols only - the principle may have become so "automatized" or well assimilated that the causal explanations are not necessary in students' analyses of the far transfer question.

b. Quantitative Analysis. A summary of the conceptions used across the four far transfer questions are presented in Tables 3.7 and 3.8. The abbreviations are the same as those used in the simple levers analysis: "m.c_c." refers to the control misconception; "m.c." to a misconception peculiar to the problem (discussed in c. below) and "surface feature" indicates an explanation predominantly consisting of superficial factors in the problems.

Table 3.7

Conceptions of the Experimental Group Students: Far Transfer Questions

Student	Revolving Door	Nutcracker	Shadoof	Nail Clippers
S1	− surface feature	− $m.c_c.$	+ surface feature	− $m.c_c.$
S2	+	+ [used principle of levers throughout]	+	−
S3	−	+ [used surface feature explanations throughout]	+	−
S4	+ surface feature	− m.c.	− m.c.	− surface feature
S5	− surface feature	+ non-gen	+ d_{Ef} >	− surface feature
S6	+	+ [used principle of levers throughout]	+	−

Table 3.8

Conceptions of the Control Group Students: Far Transfer Questions

Student	Revolving Door	Nutcracker	Shadoof	Nail Clippers
C1	− surface feature	− surface feature	+ m.c.	− m.c.
C2	+ d_{Ef} >	+ non-gen surface feature	+ non-gen	− surface feature
C3	− surface feature	+ non-gen	− m.c.	− surface feature
C4	+ $m.c_c.$	− $m.c_c.$	− $m.c_c.$	− surface feature
C5	− surface feature	− $m.c_c.$	+ surface feature	− surface feature
C6	− surface feature	− $m.c_c.$	+ non-gen	− surface feature

By simply counting the number of surface feature and misconception applications across both groups (in Tables 3.7 and 3.8), it may appear that the experimental group has fared better than the control group (distribution shown in Table 3.9).

Table 3.9

Number of Conceptions in Explanations

	surface features and misconceptions	non-gen conceptions	principle of levers
Experimental Group	14	1	9
Control Group	20	3	1

One may actually reject a null-hypothesis ($X^2 = 8.46$, df = 2, $p < 0.02$), and infer that their participation in the tutoring interviews resulted in a better understanding of levers by the experimental group students. Such an analysis may be misleading, since the differences were mainly contributed by two students in the experimental group, and a comparison of the total scores obtained by each student in the two groups may be more informative.

The scores calculated from the responses given in the far transfer questions will be used in the comparison of the groups. The "correctness" of the students' answers will be determined both by the explanation and acceptability, since the use of a misconception or a focus on surface features may give one a correct, acceptable answer (e.g. C4's response to the door problem). The level of confidence indicated for each answer will be scored as before (for the simple levers); that is, 1 for a "guess" through to 4 for "I'm sure". However, the explanation will determine the sign (+ or -) associated with the response.

An answer will be scored "correct" (= +), when the explanation accompanying the correct response is either a non-generalizable conception or the principle of levers; and for the nail clippers problem only - when the student provided evidence of lever recognition and the use of either a non-generalizable conception or the principle of levers in the explanation.

A response will be scored "incorrect" (= -), when the response is correct but accompanied by a surface feature or a misconception explanation and when the response is incorrect and accompanied by a misconception, surface feature or non-generalizable explanation. (I rule out the possibility that the use of the principle of levers could result in the wrong answer except for the nail clippers problem.)

The summed scores for each student in the two groups are given in Table 3.10.

Table 3.10

Summed Scores: Far Transfer Questions

Student	Experimental Group	Control Group
1	-11	-4
2	+13	+8
3	-16	-6
4	-15	-10
5	+1	-4
6	+14	-6

There is no significant difference between the two groups, since at a set $p < 0.050$, and for $n_1 = n_2 = 6$; $U_{calculated} = 11$ has a probability of occurrence under H_o of $p = 0.155$ (Mann and Whitney, 1947). What appeared to emerge from the simple levers data is therefore confirmed by the far transfer questions' data: the experimental group students did not gain a sufficient understanding to support transfer of learning.

c. Qualitative Analysis. I proposed at the beginning of this evaluation section that students with a deeper understanding, acquired as a result of their participation in the tutoring interviews, would be the most successful in the far transfer questions. This deeper understanding should be evident in a student's use of a qualitative principle of levers across all the posttest questions. The possibility that students may construct such an understanding of levers as a result of their participation in the pre- or posttest interviews or as a result of other, external interactions, was also posed. The use of conceptions in students' explanations across the pre-and posttests are shown in Tables 3.11 and 3.12. The idea is to trace a student's use of conceptions over time to identify, to some extent, the origin of the conceptions used in the far transfer question explanations.

There are at least six interesting cases to consider, namely S2, S4, S6, C1, C2 and C4. The analysis of students' responses to the simple lever questions suggested that the three control group students have somehow acquired an understanding of levers that is at least compatible with that of a student from the experimental group. I shall discuss two issues here: is it at all possible to determine the origin of these acceptable conceptions; and, are the three control group students' understanding comparable with what was described earlier as a deeper understanding of levers?

The explanations were coded in the following manner:

* -> surface feature

- -> control misconception

-2 -> symmetry misconception (only for S4)

n -> non-generalizable conception

+ -> principle or fulcrum-helps model

|\| -> instructional intervention (tutoring interviews)

Table 3.11
Trace of Conceptions Used over Time: Experimental Group

	Pretest Simple Levers				Posttest Simple Levers			Far Transfer Questions			
								II/III	II/II	III/I	II/III/III
	I	II	II		I	II	II	door	nuts	shadoof	clippers
S1	-	-	n	\	-	-	-	*	-	*	-
* S2	n	-	n	\	+	+	+	+ +	+	+ +	+ +
S3	+	n	+	\	-	n	n	*	*	*	*
* S4	+	-	n	\	-2	n	n	*	-2	-2	*
S5	n	-	-	\	-	-	-	*	n	n	*
* S6	-	-	-	\	-	+	+	+ +	+	+	+

Table 3.12
Trace of Conceptions Used over Time: Control Group

	Pretest			Posttest						
	Simple Levers			Simple Levers			Far Transfer Questions			
	I	II	II	I	II	II	II/III door	II/II nuts	III/I shadoof	II/III/III clippers
* C1	-	-	n	n	n	n	*	*	*	*
* C2	+	n	n	+	n	n	+	n *	n	*
C3	+	n	n	+	n	n	*	n	-	*
* C4	-	-	n	n	n	n	-	-	-	*
C5	+	-	n	-	-	n	*	-	*	*
C6	-	n	n	-	-	-	*	-	*	*

Learning as a Result of External Interactions. In C2's protocol there was evidence of the use of at least part of the principle of levers in the simple levers questions. One might therefore assume that he had a meaningful or deeper understanding of the content, apparently constructed in the three week interval between the administration of the diagnostic test and his interviews. However, when one compares C2's progress with that of S2 and S6 who also used the principle of levers, differences are observed. I am inferring that S2 and S6 have both constructed the principle of levers in the lesson interviews, since these conceptions appear for the first time in the posttest protocols in this analysis. (The posttest interviews were conducted immediately after the lesson interviews.)

There is evidence that both S2 and S6 were able to distinguish levers within the far transfer questions, while C2 mentioned "more leverage" in a general sense in the door problem and the nutcracker problem. Probing his use of the term "leverage" revealed that he meant "more power", perhaps an acceptable naive description of a lever. However, C2 used both non-generalizable conceptions and surface feature elements in his explanations, while neither S2 nor S3 relied completely on these naive descriptions only. Thus: even though C2 seemed to have acquired (via some interactions in his "real world") an acceptable view of levers, one can infer from the evidence above that his understanding was still less robust and flexible than those of S2 and S6.

The Robustness of Conceptual Change facilitated by Question 3. Two other interesting incidences of knowledge construction, apparently as a result of the class II simple levers questions in the pretest, were suggested by the pretest protocol analyses. C4's pretest responses indicated that the second class II lever question facilitated conceptual changes for her, in quite clearly causing conflict between a "fulcrum-helps" model and her control misconception. Again one should question the plausibility and fruitfulness (P. Hewson, 1985) of this new knowledge for the student, and again one may infer that this understanding was limited: she reverted to the control misconception and surface feature explanations in her responses to the far transfer questions. The protocol of the other student, C1, also suggested conceptual change as a result of the class II lever questions. Her responses to the far transfer questions show that she was not able to transfer this new understanding.

Summary. Thus, in accord with the definitions of transfer in a. above, one may conclude that students who acquired a seemingly acceptable understanding from sources other than the tutoring interviews, were less able to transfer this knowledge to far transfer situations. This lack of transfer indicates a limited understanding.

There is inadequate evidence available to infer that more meaningful knowledge was constructed by participants in the instructional sequences. However, an analysis of the processes by which the two successful students constructed their understanding will be important towards the lesson's redesign.

The profiles of the unsuccessful experimental group students are equally important to provide ideas toward improving the lesson. At this stage one can infer that S1 and S4 became more convinced of, and acquired misconceptions respectively, as a result of the instruction. The above are all findings that are useful towards the finer, formative evaluation of each of the lesson's sequences.

A thorough analysis of the students' reasoning about the far transfer problems should yield more information about instances of far or negative transfer and other phenomena that may be have been lost in the grosser analyses before.

d. Students' Ideas: Complex Levers - a Revolving Door. A brief analysis of the levers in the revolving door is given in Figure 3.5.

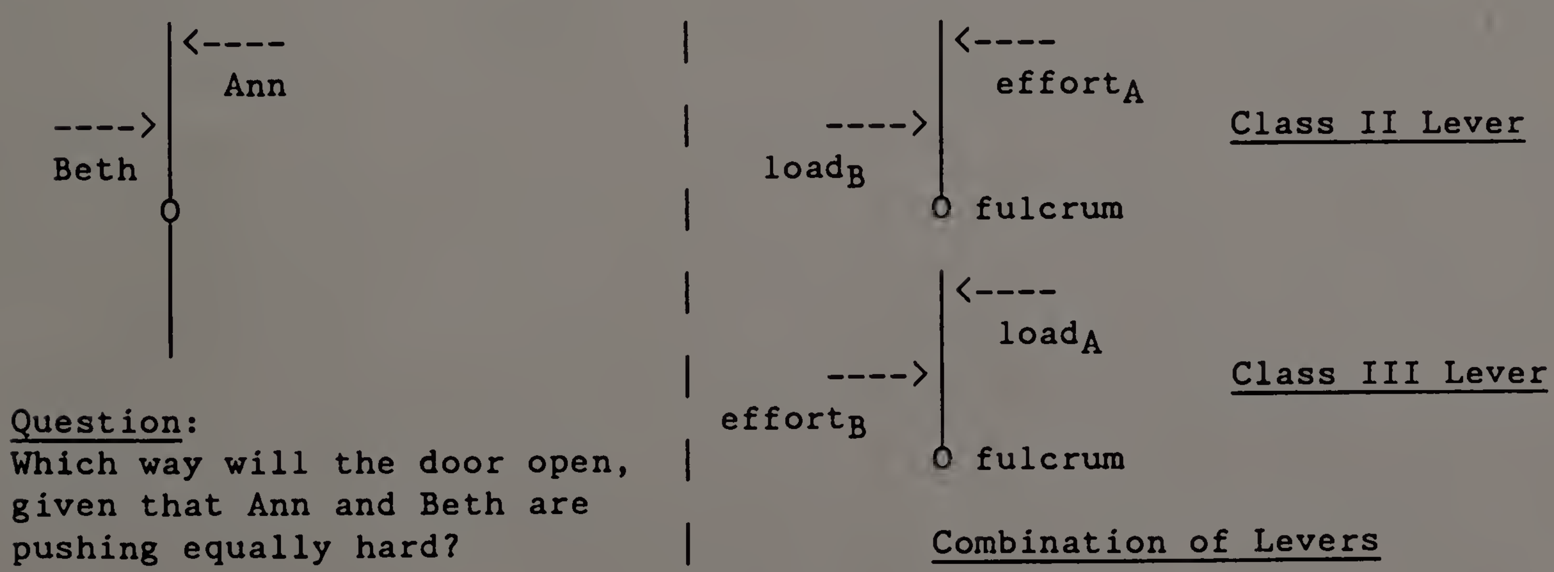


Figure 3.5

Far Transfer Question: Revolving Door

The door will move counter-clockwise. If the analysis above is carried out, it follows that Beth's effort is less than that of Ann (where effort = force x d_{Ef}), thus giving Ann's effort an edge over Beth's.

Surface Feature Recognition. The students' explanations suggested mostly transfer as a result of surface features recognition. The children apparently use a well-known, probably often experienced fact, i.e. that two equal and opposite forces balance each other. The facts, that the forces are equal and exerted in opposite directions, are the only details of the problem that the students appeared to notice. The more significant aspect of the problem is the different points of force application. All but four students (S2, S6, C2 and C4) used the opposite-and-equal-forces surface feature idea to arrive at their incorrect conclusion.

Students' frequent use of simulations (with their hands) and body language in general to augment their explanations was an interesting aspect of the conversations about the revolving door question, since this rarely happened in the other problem contexts. Most pressed both hands together, or pushed with one hand on top of the table's edge and the other directly below in an attempt to convey their ideas. The level of confidence in this answer was always high, never below "I'm fairly confident", except for student C1, who insisted throughout the interview that she was "guessing", although her door-answer was quite eloquently stated. I prompted her more than others on this question, since she noticed that the forces were applied at different points, yet decided that:

005 C1: I think the door won't move [pushed hands against each other].

006 I: The door won't move?
007 C1: Yeah // Yeah, the door won't move.
008 I: What did you say about Beth "is more on the inside"? [Reference to an earlier, mumbled statement.]
009 C1: Beth is more lower towards the inside and Ann is like, higher towards the outside. So I think / It won't move at all.
010 I: Won't move at all? Even though they're not exactly in the same place?
011 C1: Yeah.
012 I: So what makes it stay still?
013 C1: Just like / They're both pushing on it [again pushed hands together].
014 I: Pushing on opposite sides? Is that what you're showing me with your hands?
015 C1: Yeah, yeah.

Other students' reasoning were similar to her statement in line 013, with the same gestures indicating opposite and equal forces:

007 C3: It [the door] can't move if they are pushing equally hard, they'll just be going against each other [pushed hands together].

001 C6: The door will not move / Because they're both pushing [pushed finger tips together]. There's the door here [pointed to sketch] and they're both pushing the door, so I don't think they're gonna move.

004 S1: So I think that it [door] won't move. Since they are pushing the same, like equally hard? [Pushed on top and below the table.]

005 S3: Well, if they're both pushing the same, that means it can't go forwards / Counter-clockwise or clockwise, because they're both pushing the same amount of force.

009 S4: If they are both using the same amount of pressure to open / To try to open the door, instead // If they were the same then neither is going to be able to open the door. Neither one can knock the other one down...

Far Transfer. Three students (S2, S6 and C2) use the principle of levers, or part of the principle to explain why the door would move counter-clockwise. C2 already showed some glimpses of understanding of the principle in the simple levers pretest and said here that:

001 C2: I think (a), the door will go counter-clockwise.
002 I: Will go counter-clockwise?
003 C2: Ah / Because Ann has more leverage than Beth does and that

gives her more strength.
004 I: She has more leverage?
005 C2: It's [pointed to Ann] farther away from the hinges.

It is clear that leverage, for C2, depends on the effort-fulcrum distance, an acceptable naive idea.

Both S2 and S6 used the lesson's lever terminology; indicated that one may consider Beth and Ann as either load or effort in analyses reminiscent of that in Figure 3.5; and used the levers principle to solve the problem:

001 S2: A, the door will go counter-clockwise. It's like the stuff I've been talking about [referred to lesson], there is a longer distance from A to the hinge or turning point/ But / But Ann and Beth are both forces, and Ann is farther out, so then the door goes clock-wise.

And:

001 S6: Uhm // Well, actually both of them [Ann and Beth] are constituted both as loads and forces in this. So. Beth is closer to the load on the turning point's side / But Ann is / Ann has the load in the middle. So, Beth is gonna have to push harder / To push the door / So, the door would go counter-clockwise.

I am claiming that S2 and S6 learned the principle from the lesson: S6 used phrasing similar to "closer to the load on the turning point's side" in his reasoning about class III levers; S2 explicitly referred to "stuff" he talked about just before in the lesson interview; and both students used lever terminology.

Summary. There was evidence of transfer of knowledge in three students only, one of whom was in the control group. All the other students used inappropriate reasons (surface feature ideas and misconceptions) to explain their answers, even though they indicated familiarity with the situation and high confidence in their answers. Since the non-generalizable conception appeared frequently in the simple levers posttest, I had expected more reasoning from non-generalizable conceptions. A first requirement to a successful solution of this problem is the componential analysis of the levers in the situation (into the class II and III levers). A fairly simple transformation to permit analogies to the simple levers should be performed next (turning the levers horizontal) and in addition, Beth and Ann should be "seen" alternately as loads and efforts. This is quite sophisticated reasoning, much like Clement's (1988) description of analogical reasoning in experts, and most students probably could not "get" to a situation where their non-generalizable conception seemed suitable.

e. Students' Ideas: Compound Levers - Nutcrackers. The nutcracker problem is probably the easier of the four; the machine is a compound lever, consisting of two class II levers; and requires one transformation of the simple levers since one lever is upside-down, as shown in Figure 3.6 below.

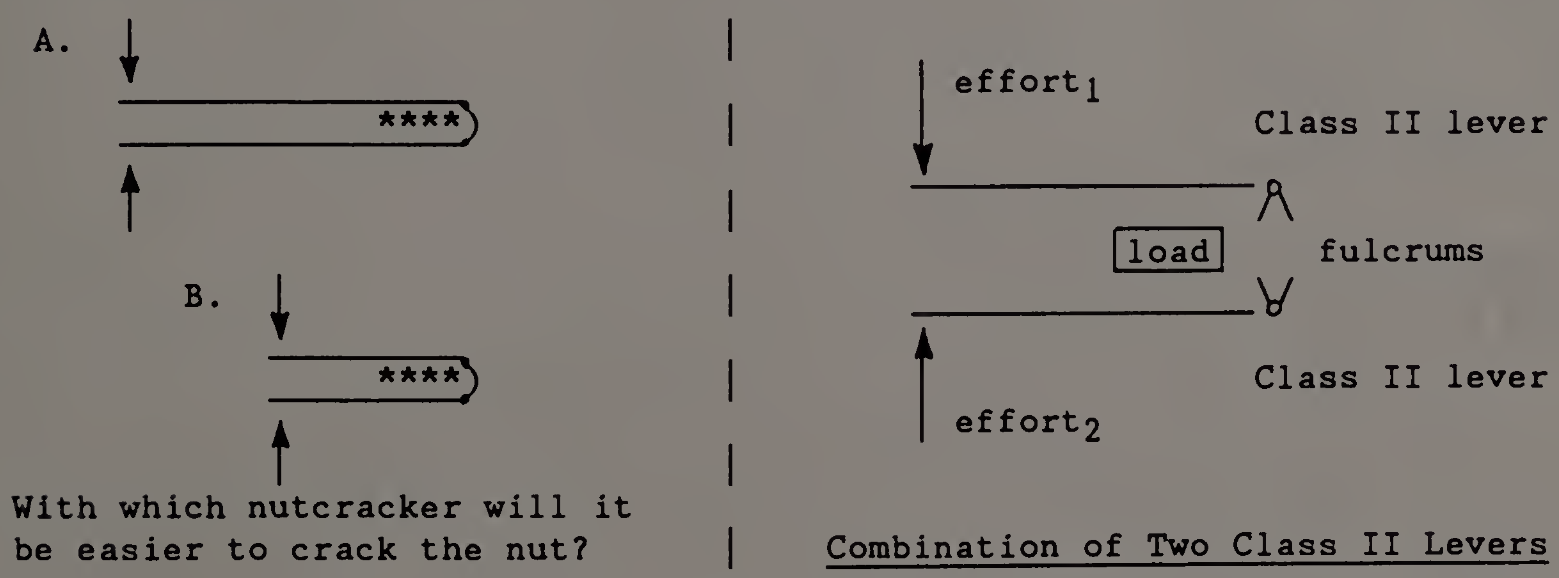


Figure 3.6

Far Transfer Question: Nutcrackers

Nutcracker A has a longer effort-fulcrum distance and the load leverarms for A and B are the same, hence it will be "easier" to crack the nut with A.

Surface Feature Recognition. C2 added a surface feature reason to his explanation that A, since it has "...longer handles..." and thus "...more leverage..." (line 011) would require less force. He thought that there was "...less space between the handles..." (line 019) in A, making it easier to "...get them together..." (line 019). He repeated this even after suggestions that he may measure the distances in A and B to verify equality.

C1 was aware of the different length of the handles for A and B, but "...since it's the same kind of nut, it does not matter what size the nutcracker is..." (line 031).

Students' responses in the pilot studies led me to anticipate that most students would focus on a "comfort" surface feature, but only S3 thought that "...there's more place to hold [in A], then you can put more force on it..." (line 009).

Negative Transfer. S4 considered differences in leverarms irrelevant. This misconception appeared for the first time (in the pre-posttest analysis) in the posttest, leading to the inference that he had acquired a misconception in the teaching interview. This is clearly an instance of negative transfer.

Far Transfer. Again, only S2 and S6 used the principle of levers and lever terminology, and there is also an example of analogical reasoning in S2's transcript:

> 014 S2: It [A] has longer handles, it's like the one above [revolving door], there's a longer distance from the turning point to where you're pressing down.

012 S6: So [marks A]. And you've got more distance [drew line from the turning point to the force in A which he labelled].

Summary. The nutcracker is not a complex lever and the instrument was familiar to most students. However, the majority of the children (eight - five control group and three experimental group students) used either a misconception or surface feature explanations. This is an indication that even the class II levers teaching sequence may not have facilitated (in more than two students) the kind of deep understanding intended.

f. <u>Students' Ideas: Complex Levers - Shadoofs</u>.

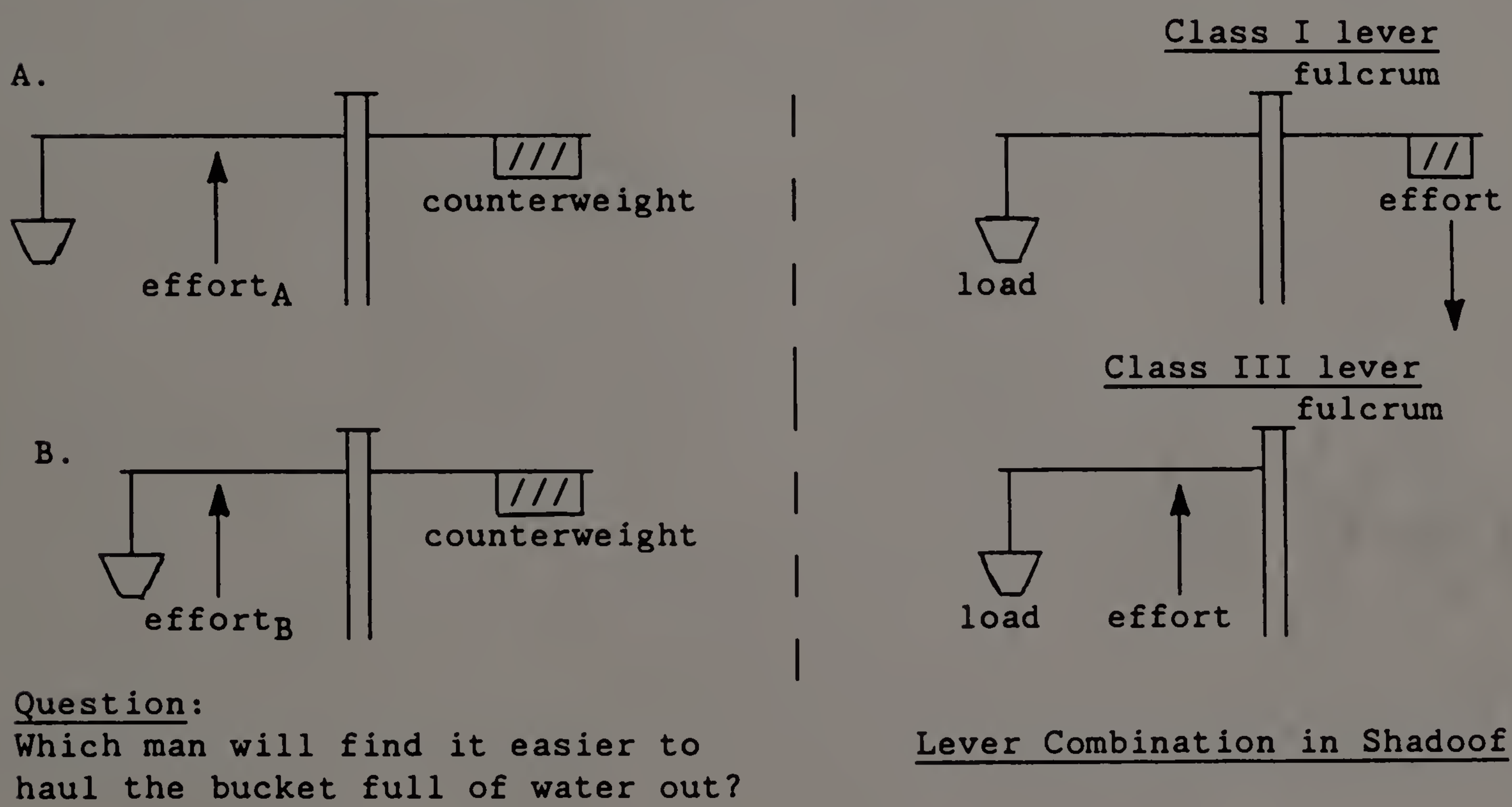


Figure 3.7

Far Transfer Question: Shadoofs

I expected that this question would be difficult: it is not a familiar machine and surface details may easily detract attention from the levers. However, only two students were confused about the shadoof's "action" and both had strong, correct naive intuitions about the "easier" job of hauling water. If one ignores the first class lever, and this can be done since the counterweight is obviously common to both machines, this becomes a simple class III lever problem, as shown in Figure 3.7 above.

Misconceptions and Surface Feature Recognition. Half of the students used surface elements to explain their answers. There was "...too much wood..." in A (S3, line 015); "...less gravitational pull..." on B (C2, line 035); "...the pole was longer, so he has more force in pushing it [the bucket] up..." (C3, line 033); difficulty in reaching the bucket in A (S1) and a conviction that the longer pole in A would break (C5), despite assurances that this was not a possibility.

The "irrelevant length" misconception appeared for the second time in S4's protocol - suggesting that this may not be a situationally dependent conception or surface feature recognition, but perhaps an understanding of levers that makes sense to him. Again, it should be noted that such an understanding could only have originated from his learning from the lesson.

Non-Generalizable Conceptions. The class III non-generalizable conception was used by five students in their explanations, including S2 and S6 who were identified as students who showed most evidence of far transfer. The explanations related a shorter load-effort distance with less effort, e.g. "...B [would be easier] because this [bucket] is closer to you..." (S5, line 025).

All of these students described, in responses to probes about the role of the counterweight, the class I lever in the machine as an aid to the person hauling the water: "...I think to make it easier for him, and also it would keep it [bucket] balanced when it came out..." (C6, line

036); and "...the weight sort of pulls down some and the man pushes up..." (S3, line 021).

Transfer by Analogical Reasoning. C6 made an interesting analogy to the nutcracker. He gave the wrong answer to the nutcracker problem and explained, employing the control misconception that:

> 016 C6: The pressure that you're applying is closer to the nut, it's not as far away as in the longer one.

I have mentioned before that the control misconception becomes a non-generalizable conception when applied to class III levers (such as the shadoof), and C6 noticed this:

> 025 C6: I think it'll be easier for B to pull it [bucket] out, because // He or she is closer to the bucket and they have // Ah, *sort of like the nutcracker*, they have, ahh // They're closer to it [bucket] so they'll be able to lift it up easier. [My emphasis added.]

S5 used an analogy in a more positive sense. She recognized (after a probe) the class I lever:

> 028 I: OK. Can you just tell me one more thing about that particular problem, uhm / What does that weight do?
> 029 S5: It, it / It pulls it [shadoof] down. It's like / Kind of like a seesaw kind of / It helps pull it down, so if you push up, the weight goes down. Because otherwise if you pushed it up, then the weight / Then you could not get this [bucket] off anyway. 'Cause if you let it go, it would just go 'boom'. [Accompanied with much body language and simulations.]

Far Transfer. There is evidence that S2 and S6 used lever terminology and that they were able to distinguish between the two levers in the shadoof, in addition to their use of the class III non-generalizable conceptions in the explanations.

S6 distinguished two loads (the bucket and the counterweight); drew one leverarm in on the given sketch and decided that "...B would be easier because the turning point is...closer to the load..." (line 019). Although the counterweight should be described as an effort rather than a load, the use of lever terminology as well as his drawing and statement is probably enough evidence to specify far transfer.

S2 was the only person who was able to distinguish the two levers, and he labelled the force, load and turning points on the drawing. Although he used a non-generalizable explanation - "...person B is closer to the bucket..." (line 032), I regard his answer as far transfer because of his clear identification of the levers.

Summary. The children fared remarkably well in understanding this question. They held strong intuitions about the "logic" of the apparatus, and the fact that almost half of the students used the class III non-generalizable conception in their explanations, lends support to the pilot study finding that the class III levers are well understood in an intuitive way. Far transfer was again observed in the protocols of only two students and not as specifically as required by the definition, since the principle of levers was not used. I infer from this that students may resort to those explanations that make more "gut sense", i.e. in this case the class III non-generalizable conception, when faced with complicated and unfamiliar situations. This last inference is in accord with the now generally accepted view of children's naive knowledge of physics (see Driver and Erickson, 1983).

g. Students' Ideas: Complex Levers - Nail Clippers. The nail clippers comparison is by far the most difficult of the problems used in this investigation. It is difficult to distinguish loads and efforts in the two levers, the most "visible" is the nail, as load, in the lower lever. It is only after one, the identification of the load, effort and fulcrum for all four different levers that one can determine a shorter effort-fulcrum distance in the lower lever of clipper B; and two, making an assumption of equal magnitudes of the forces acting on the lower levers in both A and B, that the solution becomes apparent.

It is unreasonable to expect thirteen-year old children to perform this kind of analysis, and since the problem was included more as an exercise in lever recognition, evidence of the latter will be regarded as sufficient to specify far transfer.

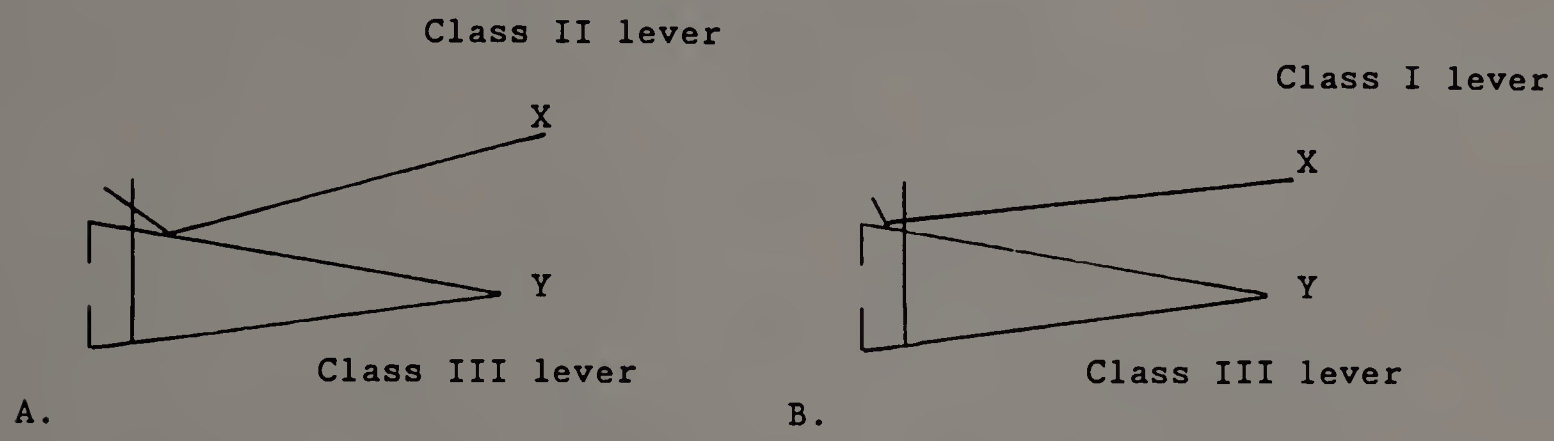


Figure 3.8

Far Transfer Question: Nail Clippers

Surface Feature Recognition. Most students focused on a striking superficial difference between the two clippers, namely the different angle formed between the top and lower levers in A and B. They described this as a difference in "space", and pointed the relevant distance out after probing. I interpreted their "space" as references to the vertical distance from the top lever's end to the end of the lower lever (XY in the sketches in Figure 3.8). Since there is this extra "space":

> 070 C4: ...there's more room to pull it down and so you get more strength, more power out of it...

> 025 S3: I'd say A, it's up on a higher degree of an angle, so you could push down harder.

039 S4: Because in A since it's higher you get more power to push down, more time and I'm sure I'm right.

035 S5: The lever [in A] is up more, so you have that much more room to push it down.

C2 and S6 mentioned the greater vertical space, but were not able to articulate a causal relationship between this factor and the increased force they perceived.

Two students, C1 and C5, thought the clippers were exactly the same, that they "do the same thing" (clip the same nail) and were therefore equally effective in cutting the nail.

Transfer on the Basis of Analogies. There was one example of "negative" analogical transfer: negative in the sense that the observed (erroneous) analogy was used to support an argument based on surface feature details. S5 spontaneously ventured further information (following her statement in line 035 above):

041 S5: See, on that one it's closer down [B], so it'd be that that one [A] has that more force. *It's like when when you're trying to pull out a nail* [flapped pages backwards, referred to crowbar question in lesson], it's easier if you have it up high, than when you have it down here [as in B]. Then you only have that [indicated distance XY] much force.

There are other fascinating aspects to her explanation, e.g. the idea that force is somehow proportional or represented by distance. It is perhaps important to note that students may err when reasoning by analogy in this content area. The underlying principle is the same for all three classes of levers, and until students are able to analyze situations more "expert-like", they may utilize analogies in the manner illustrated by S5. One can envisage a misconception becoming more and more entrenched if students could find enough negative analogies among machines (of different lever classes) to bolster their beliefs.

Far Transfer. S2 distinguished the two levers in A and B, identified fulcrums and loads correctly, but became lost in the analysis at the point where the expert would start assuming equal forces. He finally decided that clipper A would be easier, since the effort-fulcrum distance in A is larger than that in B, thus never taking the lower class III lever into consideration. However, I am satisfied that he was able to transfe2 the principle he constructed during the tutoring interview to a situation very remote from the simple levers.

S6 recognized the top lever, correctly identified the two loads in each clipper as well as the efforts and fulcrums for the two top levers in A and B. He used the load-fulcrum distance difference to decide that clipper A was the better nail clipper. He did not "see" the lower lever though, even after probing. I maintain, however, that his analysis of the top lever, the use of the principle of levers as well as lever terminology, are sufficient reasons to regard his answer as an example of far transfer.

5. <u>Summative Evaluation: Conclusion</u>

The summative evaluation indicates that most of the experimental group students may not have added to their naive knowledge of levers from their participation in the lesson. There was only one instance of significant difference between their performances in the posttest when compared to the control group students, i.e. it seemed that the experimental group students used less surface elements and more principles in their explanations for the far transfer questions.

The qualitative analyses suggest that the class II teaching sequence was more successful than the other parts of the lesson, and that the teaching sequence on class I levers may be particularly ineffective. The protocol analyses also show a general inability of the majority of the experimental group students to transfer their newly acquired knowledge of class II simple levers to the far transfer questions. This may indicate that even the effectiveness of the class II teaching sequence (suggested by the quantitative analysis) could be superficial.

The summative evaluation alerts one to possible problems areas in the lesson, especially the frailty of the apparent success of the first teaching sequence, and a possible major mishap in the second sequence.

F. Students' Learning Processes

1. Introduction and Definitions

The lesson's instructional goals were to change students' misconceptions about levers and to add to students' non-generalizable preconceptions of levers to construct a conception more compatible with a physical theory view. Both these statements describe conceptual change processes and the following definitions are proposed for use in the analyses of the lesson's effects.

a. Conceptual Change. Suppose one is considering an individual holding an existing conception C that is not in agreement with accepted physical theory, who is then faced with a new conception C' in some instructional setting. Obviously, the person can reject C', thus keeping his conceptions essentially unchanged. More important to the interests of the analysis of the levers lesson are those instances in which a person would change C. Conceptual change can happen in a number of different ways that are not independent, since one may give rise to another in complex patterns (Hewson, 1981) e.g.: C' may be added to and become integrated with C; or the old conception C can be "taken over" or replaced by C' so that C is rejected or greatly reduced in scope.

b. Observable Criteria. Several observable criteria were used to identify conceptual change instances in the protocol analyses. The following issues were considered in the data analysis: one, all changes in conceptions (both desirable and not) are probably related to the instruction; and two, students' reasoning and the processes of knowledge construction are important for possible improvement of lesson 1. Both statements above imply that analyses of all identifiable instances of conceptual change are needed.

A baseline was established from which to identify conceptual changes, i.e. a criterion for instances *not* regarded as conceptual change: suppose a student responds to a problem with an answer and supports this with an explanation. At the strongest level of rejection, this person's responses for all problems that are viewed as analogous to the target problem, will remain unchanged, thus indicating no conceptual change. Situations where the student's answer has changed, but his explanation remained substantially unchanged will also be regarded as no-change cases.

In addition to the baseline criteria, *normative conceptual changes* are to be distinguished. Some of the observed conceptual change instances seemed to be more normative than others, but since I was also interested in changes that are not normative, the following criterion for normative conceptual change is proposed. Given two target questions in a sequence, and assuming that both the answer and the explanation in the first question were unacceptable; a student should respond correctly, with an explanation more compatible to that of a physicist's, to the second target question.

2. <u>The Class II Levers Lesson</u>

A diagrammatic representation of the teaching sequence is presented in Figure 3.9; followed by a short description of expected results from the tutoring; processes of conceptual change and the students' responses to the near transfer questions.

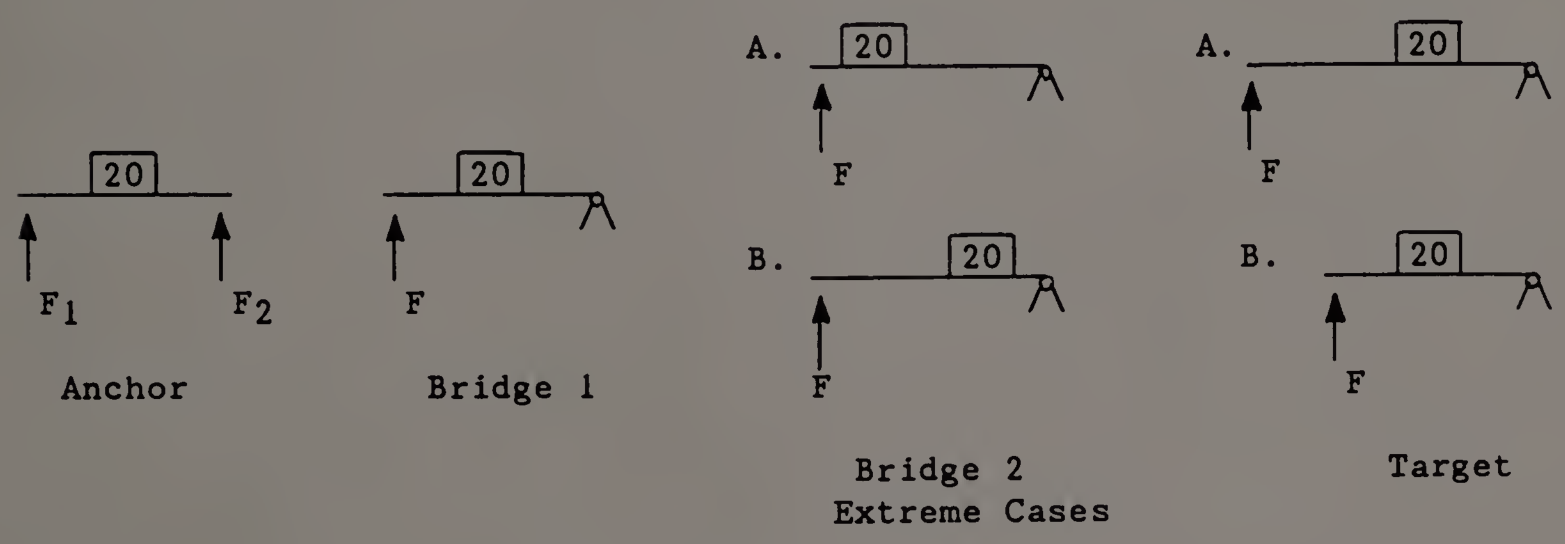


Figure 3.9

Bridging Sequence for Class II Levers

a. <u>Intended Results</u>. Students had to estimate the forces exerted in the anchoring and bridge 1 examples. The target, bridge 2 and the near transfer questions are all comparisons of two situations and the question is always "which person would find it easier" to perform the task illustrated.

Most students would probably hold a misconception about the situation depicted in the target problem, that is man B would have to exert less force since he has more control over the load. The anchoring situation, about which the students would probably have a useful intuition of a sharing model, is then extended (analogically) to bridge 1. One expects that students would start constructing a fulcrum-helps model from bridge 1, given that in the pilot study they were able to appreciate an analogous case (to the anchor) in bridge 1 and thought that the fulcrum would push up (or support or hold) with 10 lbs. In bridge 2 this fulcrum-helps model is consolidated. Finally, by using their fulcrum-helps model and the class II non-generalizable conception students should change their initial misconception.

b. Summary of Conceptions Used. In Table 3.13 a summary of the students' responses in the lesson sequence is given. The abbreviations are the same as in the previous tables, and the fulcrum-helps model is sometimes abbreviated to "f-h model".

The summary in Table 3.13 shows that five students either changed their misconception about the target problem, or added to a non-generalizable conception for the target problem. The sixth student (S1), retained her original conception, but there is evidence for a different reasoning strategy in one of her solutions in the sequence.

Table 3.13

Conceptual Change: Class II Levers

	Target	Anchor	Bridges 1	Bridges 2	Target	Near Transfer 1	Near Transfer 2
	–	+	–	+	–	+	–
S1	$m.c_c$.	sharing	$m.c_c$.	intuition	$m.c_c$.	non-gen	$m.c_c$.
		model					
	–	+	+	+	+	+	+
S2	$m.c_c$.	sharing	fulcrum-helps		non-gen	[principle]	
		model	model			f-h model	
	+	+	–	+	+	+	+
S3	non-gen	sharing	m.c.	non-gen	non-gen	princ	$m.c_c$.
		model		f-h model			
	+	+	+	+	+	+	+
S4	$m.c_c$.	sharing	sharing	non-gen	princ	non-gen	surface
		model	[fulcrum-helps model]				feature
	–	+	+	+	+	+	–
S5	$m.c_c$.	[sharing]		f-h	non-gen	non-gen	$m.c_c$.
		model		model	f-h model		
	–	+	+	+	+	+	–
S6	$m.c_c$.	sharing	[fulcrum-helps model]		[non-gen, princ]		$m.c_c$.
		model			[f-h model]		

c. New Conceptions Acquired. The protocol analyses indicate that five students' posttutoring non-generalizable conceptions of the target situation were markedly different from the control misconception and remarkably alike across the students. The non-generalizable conception was discussed as a common preconception in D. above. It is interesting that one can expand this concept of the students' naive non-generalizable force diagrams when the posttutoring conception is considered.

In their posttutoring explanations of the target problem, the students were focusing on the distance from the load to the person applying the force, and in some cases the distance from the load to the fulcrum. The "dissipating force" notion is apparent in the explanations of S2 and S6 particularly (my emphasis added to distinguish the relevant phrases):

033 S2: I guess let's say that A would be easier. Because again I think the lengths [from] the weight to the person might make it a little bit easier, may take a little bit of the pressure off.

020 S3: I would say A, there's more board to, push up, well there's more board right here [pointed to force-load distance], there's more space and it's [load] more toward [pointed to the fulcrum].

068 S5: OK, I would think that it would be // Probably person A [who has to exert less force], because this [load] is farther away from him.

029 S6: Hmm, person A [would exert less force since] the block [fulcrum] is supporting more, because it's [load] closer to the block [fulcrum] on this one. And / it's [load] closer to the man in example B, so the person would have to hold up more.

S4 had a more complex explanation, but also included a statement similar to those above. He argued that person A would exert a 10 lbs force and B a force of about 15 lbs:

102 S4: And here [case B] I'd give this [person] to push maybe 15 lbs, and this block [fulcrum] maybe not so much.
103 I: The triangle block [fulcrum]?
104 S4: Yeah the block, maybe 5 or 10 lbs. I'll say 5. 'Cause this block [load] is towards him [person B].
105 I: Towards him?
106 S4: Yeah.
107 I: So that makes a difference?
108 S4: Yeah, it's a lot heavier.

I suggest that the excerpts above illustrate the emergence of students' beliefs that a third force is being exerted; that is, by the fulcrum (triangle in S4's explanation). This is the kind of knowledge construction one had hoped for in designing the lesson; the construction of a schema that approximates the physicist's view of three forces in equilibrium in the static situations. Previous research concerned with students' conceptions of forces exerted by inert objects showed that a common misconception (that such objects do not exert forces), is quite resistant to change (D. Brown, 1987: Minstrell, 1982). I propose that, in the levers lesson, students' construction of a conception that inert objects could exert forces depended on their holding an explanatory, causal model of the object (i.e. the fulcrum). This is the idea that the fulcrum is "sharing the load" or holding up some of the weight, as illustrated by S6's statement in line 029 above and the following explanations (S4 thought that the load was almost in the center of the board in the target problem's case A, and confirmed this impression with a question):

092 S4: This [load in A] is in the middle - almost in the middle?
093 I: I think - it looks like it is in the middle.
094 S4: Yeah. Oh, OK, they'd have to share the same amount of weight.
095: I: What would have to share the same load?
096 S4: That person and this [pointed to the fulcrum].
097 I: Could you perhaps write in how much [force] that is?
098 S4: Yeah. OK, I guess 10 and 10, because the block [fulcrum] is in the middle.

His idea that person A and the fulcrum share the load indicates that S4 recalled and used the symmetrical anchor situation. He estimated, by a process of semi-quantitative reasoning from a "fulcrum-helps" model, the forces exerted in case A and case B, and only then formulated the non-generalizable conception in line 104 (p. 81). In the two excerpts from S4's protocol (line 92 to 102) the entire process of the knowledge construction intended is reflected: a progress from the control misconception to a view of one force (the person's ability to exert maximum forces and "control"); to two forces (adding the weight of the load "pressing down"); to the last idea - that the fulcrum exerts a force, thus helping the person to support the load's weight.

Another example, perhaps not as rich as S4's, of a more complex knowledge structure underlying the non-generalizable conception, is evident in S5's reasoning about the posttutoring target question. She alternated between the control misconception (that a shorter distance from the load to the person exerting the force would give one more control) and the non-generalizable conception. She finally decided on the latter and referred back to bridge 2 as the source of her change of mind:

060 S5: OK. Now // From the last one [the previous problem, bridge 2], I am beginning to think it was A [exerting less force], because the 20 lbs is more in the middle. Well, actually this [B] board is shorter, so // But this [distance from load to A] is farther.
061 I: When you say this is farther, tell me, what do you mean?
062 S5: OK. Well, this is the distance [indicated with pencil from A to load].
063 I: So man A is farther?
064 S5: From the 20 lbs than man B is, but then this board [B] is shorter than that board [A].
065 I: And what difference does that make?
066 S5: It makes it easier to hold // But I am not sure. So // I think that it would be // Probably person A, because the block [load] is farther away from him and I know the board is longer, but in comparison // Uhm, if you had a board // Well, actually // This board [A] is longer so it gives it more // It's like a crowbar. So I am going to say A and I guess I am fairly confident.

S5 was clearly vacillating between the two explanations, but there are two reasons in support of the non-generalizable conception: her qualitative analogy (line 066 - like a crowbar) and the problem before (bridge 2 mentioned in line 060). She articulated a fulcrum-helps model in her explanation for her answer in bridge 2, but she did not mention this again in the final target question. However, her reference to the "last one" in line 060 could be interpreted as a reference to the model as well.

d. <u>Instances of Conceptual Change</u>. Students S2 and S6 changed their misconceptions in the manner anticipated when the lesson was designed; that is, their knowledge development (as far as can be determined from the protocols) closely resembles the "ideal" progress. Therefore, although they did not explicitly refer to the fulcrum-helps model in their posttutoring target explanations, one could hypothesize that the non-generalizable conceptions applied there resulted from a process of model construction, building on the anchor where two people share the load. The sharing conception from the anchor example is extended to bridge 1 and consolidated in bridge 2:

007 S2: [anchor]
About 10 lbs on each hand I guess. Because it's [load] in the middle so you're holding up about half on each side.

017 : [bridge 1]
I guess it is still 10, because it is still the same thing, 'cause the block on the other side is still holding it up.

020 : [bridge 2]
I'd say that's B, like on this, it's [load] closer to the triangle and I think the triangle is holding up more.

And S6:

005 S6: [anchor]
Each hand has to push up 10 lbs // Like 10 lbs and 10 lbs is 20 lbs, and if you're not counting the board that should be right.

017 : [bridge 1]
He has to push up 10 lbs because this block [fulcrum] is supporting 10 lbs.

029 : [bridge 2]
Hmm / I think / It would probably be B, because / Yeah, it's B because the block is supporting most of the weight in this case.

Novel and Successful Processes. In contrast to the processes outlined above, both S4 and S5 initially rejected one of the bridging examples (bridge 1). Both concluded that the man would have to exert a 20 lbs force, since the fulcrum could not push up. S5's reasoning is particularly interesting, because she contradicted herself within the first two sentences of her explanation: [After a long silence, a probe to establish whether the drawing was clear to her, emphasis added.]

032 S5: Yeah, because this [fulcrum] is there *pushing up*. OK // Uhm, I am thinking he could either have to hold up 20 lbs because this [fulcrum] is not holding up anything or he could be holding up 10 lbs and the other 10 lbs could just be resting on this [fulcrum], because that's what the other person was holding [reference to anchor]. So that's what I guess it is; he's holding up 10 lbs.

She became more convinced of this fulcrum-helps model, in that the fulcrum now "held" and "pushed" rather than the "...10 lbs just resting..." on it. As shown before (p. 83), the model was probably a major factor in the conceptual change evident in her protocol.

S4 also stated that the fulcrum in bridge 1 could not push up:

059 S4: Uhm, 25 lbs [Al's force - *25* a slip of the tongue?]], because ah // I'm not really sure. I said how much for this one? [Turned back to anchor.]
Ten [lbs] each.
060 Yeah, 20 lbs because you replaced this person [in anchor] with one of these [fulcrum], right?
061 I: Yes.
062 S4: Well, now you have to double that because the person is gone, but this, this end / I am not sure why I put that [pause 10s].

After a probe to reconsider and another review of both the anchor and the bridge 1 examples:

064 I: Ah, suppose you and a friend were holding this [anchor problem] then you said you would?
065 S4: Push up 10?
066 I: And he would?
068 S4: Push up 10.
069 I: Now you friend goes away and leaves his end on the table, then you would push up, you said 20, double that?
070 S4: [Pause 5s.] Ten. I'd have to push up 10....
071 S4: ...all you're doing is taking a person and put in this block [pointed to fulcrum] and the block can hold up as much weight as a person.

Again, as in the case of S5, this fulcrum-helps model is apparently a major factor in the conceptual change process.

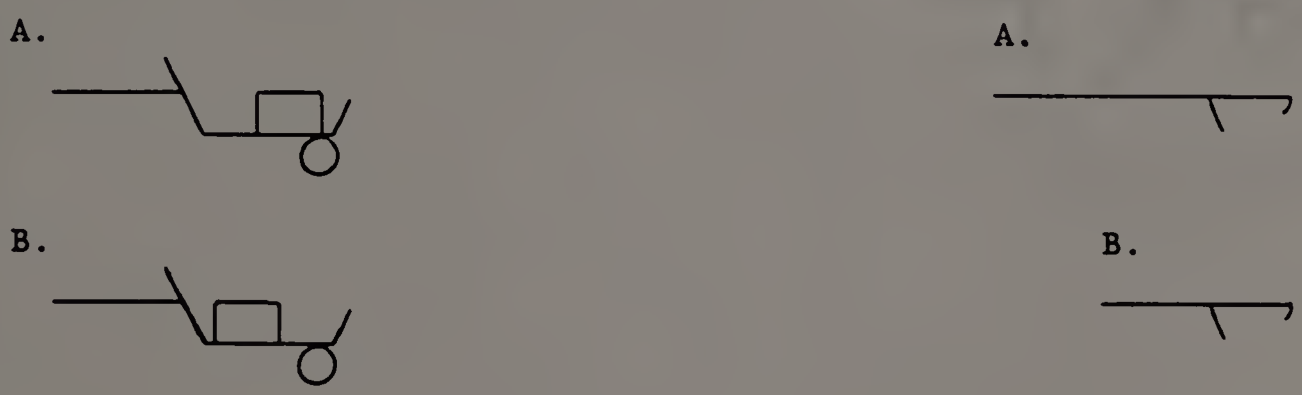
S3 also rejected the analogy between the anchor example and bridge 1. However, his reasoning about bridge 2 and pushing motions with his hands to indicate the fulcrum's action, may be sufficient evidence that he reasoned from a fulcrum-helps model about the extreme cases.

003	S3:	[anchor] I'd say they'd both have to push up 50%, each would have to push up the same amount to keep it level.
012	:	[bridge 1] I'd say he'd have to push up with / Uhm, 20 lbs.
013	I:	20 lbs - can you tell me more about that?
014	S3:	Well, there's no-one on this side [right hand], so it doesn't have any strength to push up - so it's up to Al to push up the 20 lbs.
019	:	[bridge 2] You'd probably only have to push up 5 lbs on this [B] 'cause it's really in [load towards fulcrum] and right here [A] you'd have to push up at least 20 lbs.

S3's statements in lines 014 and 019 are contradictory - in bridge 2 the man pushes up only 5 of the 20 lbs in bridge 1, indicating that 15 lbs must be held or supported by something else. His reference to the load being "really in" and a lifting motion with his hands accompanying this statement may suggest that the fulcrum exerts the "missing" 15 lbs force.

No Conceptual Change. S1, the student making the least overall progress, viewed the anchor and bridge 1 situations as completely different. The tutoring strategy used to facilitate the analogy between the two situations for another student (S4), failed for her. She used the control misconception throughout, except for the extreme case bridge 2 and the wheelbarrow near transfer question (which is perceptually near to the extreme case situation). This intuition was not sufficient to change her misconception though, perhaps suggesting more bridging *back* to the anchoring example and bridge 1 for students who appear to believe so strongly in the control misconception.

e. <u>Near Transfer Questions</u>. The near transfer questions are perceptually "near" the simple lever examples in the sequence. The aim was to provide a different, but simple context with the hope that successful transfer to these two questions would reinforce the students' conceptions constructed in the sequence, and to provide a measure of the robustness of their newly acquired conceptions. The questions are given in Figure 3.10. As in most of the comparison situations, the students had to decide which machine would be "easier" to use in the manner suggested in the drawings.



Wheelbarrows Bottle openers

Figure 3.10

Near Transfer Questions: Class II Levers

The children all answered the wheelbarrows question correctly, including S1 who used a non-generalizable conception in her explanation. This question is perceptually near to the extreme cases bridge 2 about which she had a strong, correct intuition and may explain this deviation in her otherwise consistent use of the control misconception. The students all recognized the wheelbarrow as a lever and identified the turning point (fulcrum), force exerted and load correctly.

The bottle openers, however, was a more difficult problem. The fulcrum is not a fixture of the machine and difficult to determine. The load (the cap) would not in everyday language use be referred to as such. Only S2 was correct in his prediction, used the longer effort-fulcrum distance principle to explain his answer, and correctly identified the fulcrum, load and applied force. Four other students applied the control misconception, and one (S4), explained that the openers were equally "easy" to use.

The potential fragility of the new knowledge is illustrated by the students' lack of transfer to a situation where it was not obvious that the "fulcrum helped". This finding would already suggest that the apparent success of the class II teaching sequence may be context or situationally dependent.

f. Summary. The class II bridging sequence was successful in producing normative conceptual changes in four of the six students. In two students the learning processes were similar to those aimed at in the instructional design. An important, unexpected finding from analyses of the other two normative change protocols was the conceptual change brought about by the extreme case situations in bridge 2. Students who rejected the anchor-bridge 1 analogy apparently constructed the explanatory model of the fulcrum solely on the basis of the extreme situations.

A problem in this sequence is that some students were not able to extend the anchoring example to what was believed to be an analogous case (bridge 1). This problem may be circumvented by the extreme case comparison in bridge 2, which was apparently powerful enough to suggest a fulcrum-helps model to the students most reluctant to believe that inert objects can exert forces.

The most significant finding in the analyses of the class II teaching sequence is clearly the instructional usefulness of extreme cases. Second, the students' failure to transfer a conception compatible to physical theory to the bottle openers question, suggests that the knowledge constructed during this sequence may be fragile, perhaps requiring more bridging situations and applications in the lesson to reinforce the fulcrum-helps model and the emerging principle of levers.

3. <u>The Class I Levers Lesson</u>

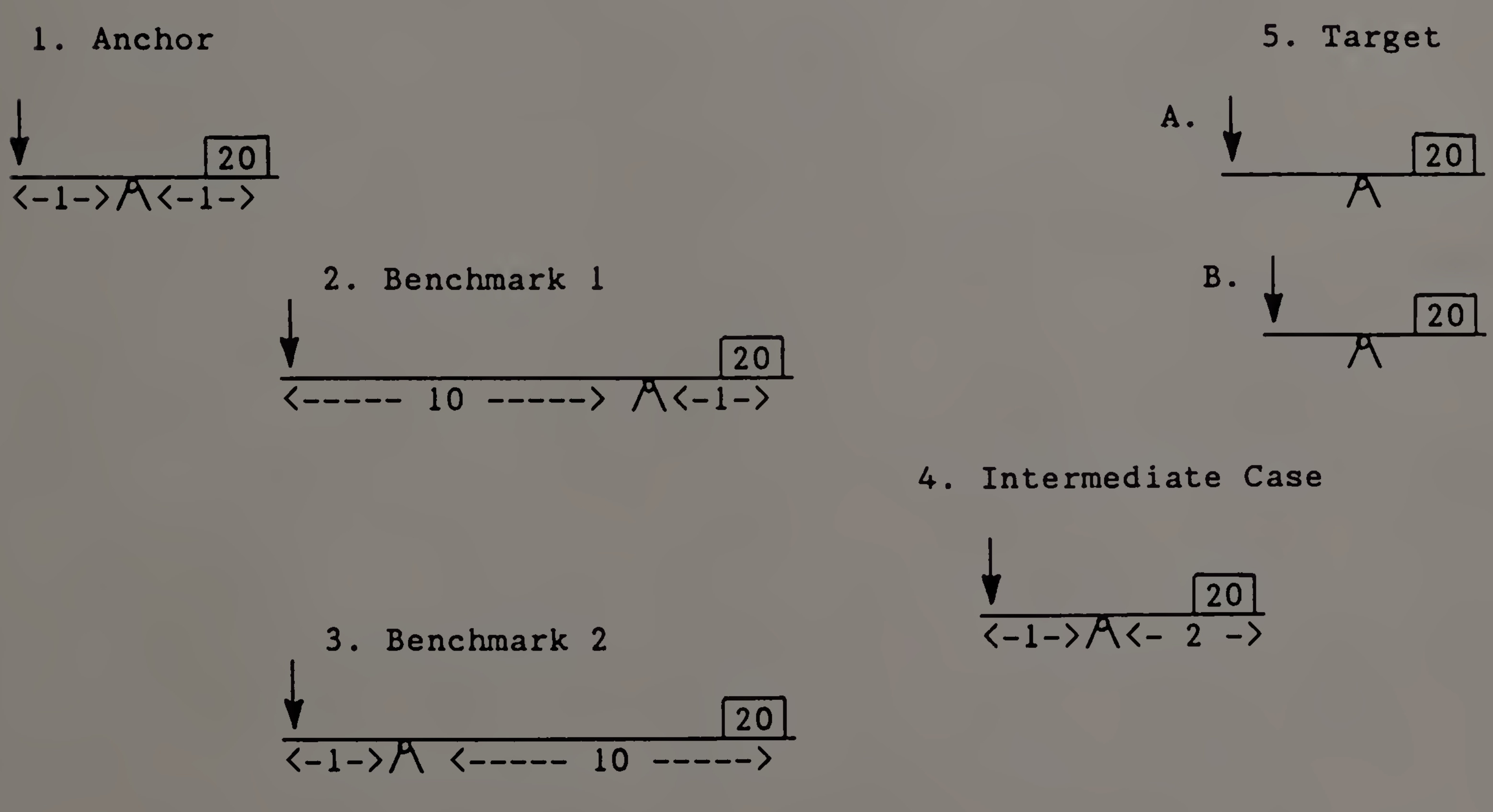


Figure 3.11

Benchmark Sequence for Class I Levers

a. Intended Results. The intention was that students would use the benchmark examples as points of reference. Presumably they would have an intuition about each of the particular benchmark situations, and would be able to argue from these extreme reference points about intermediate cases. The two benchmarks in the sequence were chosen because the changes in the leverarms are so extreme and obvious that, when coupled with the students' intuitions about the situations, they should focus on the importance of the leverarms.

b. Students' Conceptions. A summary of the students' responses is given in Table 3.14. There are two new abbreviations: S4 and S5 both acquired a symmetry misconception ($m.c_s.$), and the principle of levers appears for the first time (abbreviated to princ).

Table 3.14

Conceptual Change: Class I Levers

	Target	Anchor	Benchmarks 1	Benchmarks 2	Benchmarks 3	Target	Near Transfer 1	Near Transfer 2
S1	− $m.c_c.$	+ balance	− $m.c_c.$	+ intuition	− $m.c_c$	− $m.c_c.$	− $m.c_c.$	+ model
S2	+ princ	+ balance	+ [principle; intuition]	+	+	+ [principle]	+	+
S3	+ princ	+ balance	+ [principle; intuition]	+	+	+ [principle]	+	+ model
S4	− $m.c_s.$	+ balance	− $m.c_s.$	− $m.c_s.$	− $m.c_s.$	− $m.c_c.$	− $m.c_s.$	+ surface feature
S5	+ non-gen	+ balance	+ [unsuitable model]	+	+	+ non-gen	+ princ	− $m.c_s.$
S6	− $m.c_c.$	+ balance	+ [principle, quantitative law; non-gen concept]	+	+	+	+	+

c. Preconceptions. The students' conceptions of the target problems in this sequence are not preconceptions in the sense that there was no prior instruction in levers. One may find effects due to the tutoring in sequence 1, but since changes in students' knowledge as a result of the different components of this lesson sequence are important here, the target question explanations will be viewed as conceptions prior to instruction in class I levers.

An interesting "new" misconception appeared for the first time in the study: S4's apparent belief that equal leverarms, thus a symmetrical situation, would be preferable to (or "easier" than) all other lever possibilities:

103 S4: B would have to push less because the board is equal on the two sides of the triangle.
104 I: So if the board is equal on the two sides of the turning point it's easier?
015 S4: Yeah, you wouldn't have to push with as much weight.
106 I: What makes it more for A?
107 S4: On the left side it's longer than on here, on B.

Except for S2 and S3 who used one part of a qualitative principle of levers, the other students used the type of conceptions anticipated in designing the sequence, namely the control misconception and the class II non-generalizable conception.

d. Normative Conceptual Change. There is evidence that for three students the intended teaching goal was attained by the use of the benchmark teaching strategy.

One student's progress was indeed better than anticipated (S6). He changed a misconception and, awkwardly but satisfactorily, formulated the law of levers and proceeded to apply this law to the transfer questions following the sequence. The protocol excerpts illustrate his preconception; the conflict between his control misconception set off by the benchmark problems; the subsequent conceptual change; and his statement of the law of levers.

First, consider the statement of his control misconception in the target problem:

> 066 S6: Uhm // Person B, because [he] is closer to the load again.

Upon encountering benchmark 1, he expressed astonishment ("Wow!") at the differences in the lengths of the leverarms, and continued:

> 078 S6: Let's see / If this [force leverarm] is 10 times longer than [pointed to load leverarm, that means he's gonna have to push it down with 200 lbs.

And for benchmark 2:

> 086 S6: Uh, let me think. // He has to push with 200 lbs again.
> 087 I: So are these two the same? [Benchmarks 1 and 2.]
> 088 S6: Oh, no, no. On this one [benchmark 1] he has to push with *2 lbs*.
> 090 Because // Of the distance he has to push with. Wait // Yeah, yeah he has to push with less weight because of the distance.
> 096 And on this one [benchmark 2] he has to push with *200 lbs* weight. [Student's emphasis.]

By the time he had to explain the intermediate case question, he was confident about his answer:

> 101 S6: He has to push with 40 lbs / 40 lbs // The load is twice as far as he is.

When asked to compare the anchor, benchmarks and intermediate case again, he said:

> 105 S6: This one / This one [anchor] he's only got one distance away and this one [intermediate case] it's [load] twice as far away so that means that it weighs twice as much against the seesaw.

Finally, he explained the target situations again:

> 109 S6: [A would require less effort.]...Because of the fact that the farther away you are the less you have to push down, because like if you're twice as far away you only have to push down with half the pressure.

110 I: And when you say 'you're far away', from what do you mean, from what?
111 S6: Farther away from the turning point.

He wrote a sentence summarizing his final explanation for the situations in sequence 2, and stated the principle of levers:

115 S6: OK // The number of times longer the board is on the force's side [from the fulcrum], you divide the load by and that's how much force you need.

Two other students (S2 and S3) constructed the qualitative principle of levers. Neither of the two students held a misconception as a preconception - S3 applied the non-generalizable conception from sequence 1, and S2 provided an explanation for his correct answer informed by his experience with the analogous case of a seesaw.

082 S2: I think that A will be easier and the only reason I think that is because I / I've used a lot of seesaws with friends and stuff and I know that if you sit farther in on a seesaw then it's easier to go up.

S2 and S3 proceeded through the sequence as anticipated. Neither gave correct quantitative answers for the benchmarks, but appeared convinced that these were extreme situations in the correct direction as anticipated, e.g. for benchmark 2:

076 S3: He'd have to push a good amount.
075 ...there's less board here [fulcrum to force] which would make it a lot harder.

And, for S2:

125 S2: Can I just put more than 20 lbs - a *lot* more than 20? [Student's emphasis.]
126 I: A lot more?
127 S2: It's a lot more...
129 I don't know, I just keep thinking how hard it is to push like here.

The last remark (line 129) appears to indicate that S2's prediction is based on a direct physical intuition rather than a chain of reasoning.

Finally, both students reasoned from the benchmarks to the intermediate case and proceeded to give acceptable explanations for the target problem, focusing on the force leverarm: "...there's longer distance from the force to the turning point..." (S2, line 131) and "...how long the board is from the center where the hinges [fulcrum] are to [A, is easier]..." (S3, line 081).

e. Conceptual Change: Limited and Curious. S5's explanations for the pre- and posttutoring target problems appear to be the same, that is, non-generalizable conceptions:

Pretutoring conception:

151 S5: Yeah, because this // Like a seesaw. If I am trying to hold a person on a seesaw closer to me it's going to be harder.

Posttutoring conception:

190 S5: OK. I think person A, because the load is far away from him. And also [he] has more leverage because the board is longer.

In the benchmark situations she paid attention to the position of the fulcrum, and it is not clear whether her "longer board" that implies "more leverage" (line 190 above) referred to the effort leverarm or not. However, what is interesting is that her ordinal reasoning from the benchmarks was supported both by intuitions and, unexpectedly, by a fulcrum-helps model, e.g. at benchmark 2:

172 S5: ...He has to push down much more, because the turning point is much closer to him and so that [fulcrum] doesn't take very much of the weight...

and again in the intermediate case:

177 Well, the turning point is almost in the middle so it probably might take some of the weight.
178 I: It might take some of the weight?
179 S5: Well, not // Well, it would hold // I mean, if it's

> [fulcrum] way up here [in benchmark 2] then you have to push that much more, but if it's [fulcrum] right here, you don't have to push down quite as much // But [if] it's in the middle [anchor] so it can take some of the weight...

Analyses of the forces in any of the above systems at static equilibrium, show that her model (especially in line 172) is not in agreement with the physical theory*. However, her intuitions about the benchmark situations (although supported by a "false" model), allowed her to reason from the benchmarks to the intermediate case, and to construct an adequate conception about the force leverarm variable in the problems.

* [In benchmark 1 the force upward at the fulcrum is 22 lbs, in benchmark 2 it will be 220 lbs and in the intermediate case 60 lbs.]

f. <u>No Conceptual Change and Worse</u>. Student S1's conceptions remained unchanged: she steadfastly applied the control misconception (identified earlier in sequence 1 and surviving unscathed through all the tutoring) in both the pre- and posttutoring target problems in sequence 2. She formulated an acceptable conception in benchmark 2, and her statement suggests an intuitive certainty about this situation:

203 S1: Oh, here I think he has to push with a lot too. Because // It's so // This block [fulcrum] is close to him, and it's just a little / short?
204 I: Yes, he is only one feet from the turning point.
205 S6: So he has to push down really hard, so he can get this 20 lbs up.

Throughout the interviews S1 was slow to respond and there were invariably pauses between her reading of the problem and her response. She also often needed prompting before she started her answers. Yet, for this benchmark her response was immediate and confident; there is no discernible time lapse between the reading statement and her response to it, and one may view this as a strong physical intuition. She also held a correct anchoring conception, but it appears that for her neither of these two situations were related to the others in the sequence.

One student's (S4) pre- and postconceptions for the target situations were two different *misconceptions*. He verbalized a novel "symmetry" misconception that was never mentioned in all the pilot study interviews, for the first target problem:

> 103 S4: B would have to push less because the board is equal on the two sides of the triangle.

And for the posttutoring control misconception:

> 160 S4: B, because the weight is closer and it doesn't // Wait a minute // Yeah, I think in B, because even though it'd be heavier for this guy [B] to hold up, it'd be easier, a lot easier.

His last statement contains a contradiction, but even probing that focused on the difference between "easier" and "heavier" was not successful in eliciting his understanding, and we ended with the control misconception:

> 164 S4: I just think this would be easier, because the 20 lbs is closer to the person.
> 165 I: In terms of pushing up: who would have to push up more?
> 166 S4: Push up more?
> 167 I: Uhm, A or B?
> 168 S4: A.
> 170 S4: Yeah, 'cause it's a longer board.

The anchoring example was an anchor for him, but not the benchmarks. He decided, while responding to benchmark 2 that there was no difference between the two extreme situations: (He was speculating that less force would be required to lift the load in benchmark 2 when compared to benchmark 1.)

> 142 S4: [Pause 7s.] I think I changed my mind here [benchmark 1]. ...I don't really think there is a difference in this one and this one [the two benchmarks].

He repeated this apparently symmetry-based explanation for the intermediate case and also when asked to compare the benchmarks and intermediate case.

g. Near Transfer Questions.

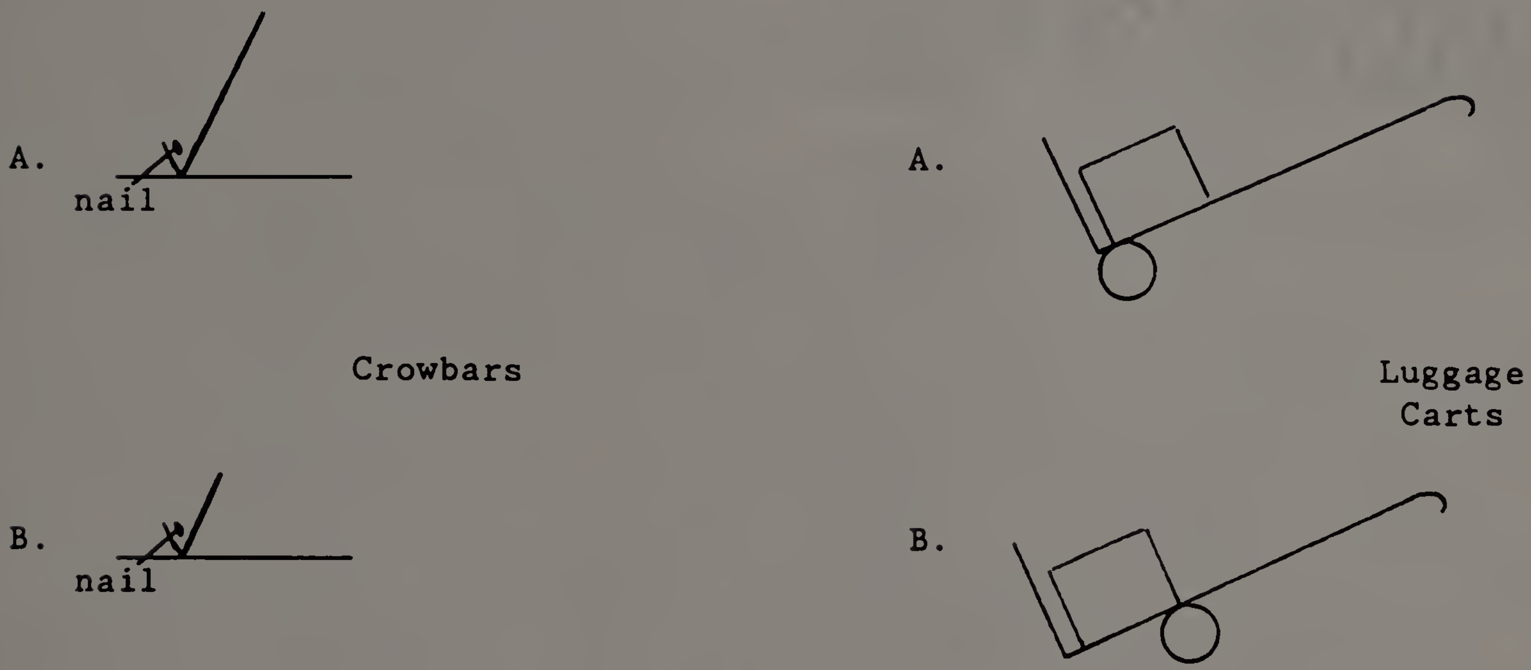


Figure 3.12

Near Transfer Questions: Class I Levers

All six students were able to identify loads, efforts and fulcrums correctly in the two near transfer questions, perhaps indicating a growing familiarity with the lever terminology. The results confirm the hypothesis that near transfer should be evident only in the protocols of students who have reached the intended understanding of class I levers. The three students who constructed a qualitative principle of levers all transferred their knowledge to the both the near transfer questions (S2, S3 and S6); although accompanied in their explanations by the non-generalizable conception (S6), and inappropriate reasoning (S3) about the fulcrum-helps model discussed in e. above.

The luggage carts problem gave rise to the kind of reasoning (S1 and S4) about the fulcrum-helps model described before as inappropriate. (This reasoning was well illustrated by S5' protocol excerpts in e. above.) In view of this, use of the fulcrum-helps model to explain class I levers should not be encouraged although it "gets" the students the correct answer and focuses attention on the leverarms. Since the drawing also caused confusion, requiring too much explanation (for S5 and S1 particularly), the luggage cart problem should probably be removed from the next lesson.

S5 indicated that she was well acquainted with crowbars and gave the principle of levers as an explanation; an indication that the crowbar was an anchoring situation for her. It is not at all clear what S5 understood at the end of this sequence: in the luggage carts problem she answered with a misconception, although this might have been due to confusion about the machine.

Finally, S1 and S4 proved to be consistent in their beliefs since both explained their answers to the crowbars question with the respective misconception used before.

h. <u>Summary</u>. There is limited evidence that the benchmark strategy may be used successfully for both changing conceptions and adding to students' existing, adequate conceptions.

A benchmark was defined as a physical phenomenon for which one has an exact, quantitative but intuitive response. The two benchmarks in this sequence turned out to be intuitions for some students, with the direction of change relationship (increase in effort) as desired, but only one student gave an exact, quantitative response. His answer indicated that the quantitative answer was not intuitive: he constructed a law and only then suggested the quantitative answers. At this stage one may therefore assume that students may have a *qualitative*, directionally correct intuition about the benchmarks and that that should be sufficient for the purposes of this sequence.

There is also evidence that the benchmark strategy's teaching success hinges upon the existence of the anchoring conceptions in students, perhaps to a greater extent than for the bridging strategies used in sequences 1 and 3 of this lesson. This is obvious when one considers the design: all the situations in the teaching section except the intermediate case, are intended anchors. An obvious drawback of this approach is that one cannot assume that these would be anchoring conceptions for all students. In this sequence three students did not hold all three the anchoring conceptions required in the design and they were the least successful. The use of empirical feedback (experiments with equipment) may alleviate this problem. The actual physical outputs in effort required in the two benchmark examples are so strikingly different that such experiences may facilitate the beliefs required.

It is not clear from the sequence 2 protocol analyses why the students performed so poorly on the class I lever question in the posttest. The emergence of a misconceptions for one student is a course for concern, but this may relate to the problem regarding anchoring conceptions discussed before.

4. <u>The Class III Levers Lesson</u>

The lever situations from sequence 3 are sketched in Figure 3.13.

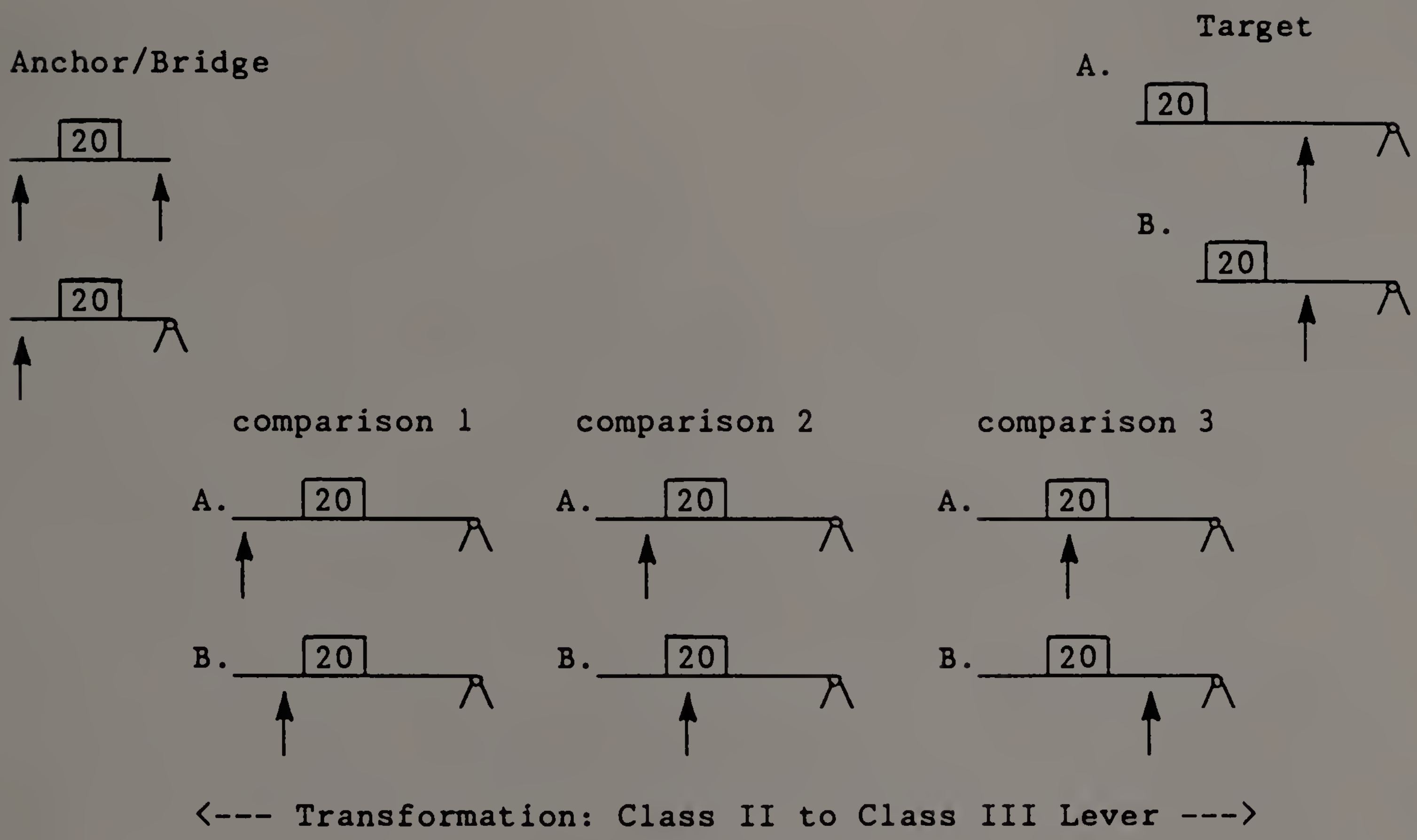


Figure 3.13

Transformation Bridging Sequence

a. <u>Intended Results</u>. Previous results indicated that most students could predict the correct answer to the sequence 3 target problem, but that few students could explain their answer and that the given explanations were naive. In view of this it was proposed that students should be able to *add* to a non-generalizable, but adequate preconception of class III levers (i.e that a smaller distance between the force and load would result in a smaller effort), rather than a replacement of the naive conception.

The assumption is that most students should have acquired at this stage of the lesson an acceptable, if non-generalizable conception of *class II* levers. The class II non-generalizable conception is important to this bridging strategy since the gradual transformation of a class II lever into a class III lever is intended.

An explanatory model of the fulcrum in class III levers may also add to students' understanding of this lever type. An analysis of the forces in a static equilibrium system for a class III lever shows that the force at the fulcrum is opposite in direction to the direction of the force at the fulcrum in a class II lever. Students might focus on the different action of a class III lever's fulcrum and develop a different model of the fulcrum: one that exerts a force against the effort - thus a *fulcrum-hinders* model rather than a fulcrum-helps model.

b. Summary of Students' Conceptions.

Table 3.15

Conceptual Change: Class III Levers

	Target	Anchor		Bridge			Target	Near Transfer	
		1	2	1	2	3		1	2
	+	+	−	−	−	+	+	−	−
S1	non-gen	share/m.c.		$m.c_c.$	$m.c_c.$	non-gen	non-gen	m.c.	surface feature
	+	+	+	−	+	+	+	+	+
S2	non-gen	[share]		m.c.	[non-gen; and fulcrum-hinders model]				
	+	+	−	+	+	+	+	+	+
S3	non-gen	share/m.c.		[non-gen]		princ	non-gen	princ	surface feature
	+	+	+	+	+	−	−	−	−
S4	non-gen	[share]		[non-gen]		m.c.	[confusion]		surface feature
	+	+	+	−	+	+	+	+	−
S5	non-gen	[share]		$m.c_c.$	[non-generalizable concept]				surface feature
	+	+	+	+	+	+	+	+	+
S6	$m.c_c.$	[share]		$m.c_c.$	[principle; models; non-gen concept]				

Pre-conceptions. The pilot study results suggested that most students would hold a non-generalizable but acceptable conception for the target problem in sequence 3. This hypothesis was confirmed: all six students applied the expected conception to the pretutoring target question. There were no markedly different statements of this conception amongst the students - all mentioned that a smaller effort was required in the case where the force-load distance was smallest.

Intended Changes. The class III non-generalizable conception seems to be a strong, intuitive belief for most students. The pilot studies' diagnostic test scores averaged over three groups (N = 118) indicated an anchoring conception with belief score 72%. This conception can be illustrated with a naive "force diagram" similar to that suggested before for the control misconception, i.e. that the shorter the load-effort distance is, the more "force" can be exerted, hence more control and an "easier" task. The problem here is of course that this belief, if encouraged, causes a "regression" to a naive view; a distraction from the importance of the leverarms and essentially bringing one full circle back to the one force, control-type naive ideas. It was proposed that the reversion to naive conceptions may be avoided by relying on the knowledge acquired by the students in the previous two sequences, plus a view of the fulcrum's force as "hindering" rather than helping.

c. Conceptual Changes. In an interesting interview, student S6 restated the law of levers and referred back to the class I lever sequence in which he had first stated this law. His non-generalizable explanation for the target problem (class III) lever was as expected. However, his initial answers for bridge 1 and 2 were "misconceptions" in that he applied the class III non-generalizable conception from the target example again with erroneous results. This may have been the effect of an attempt at consistent reasoning. He seemed aware of the conflict and alternated between two answers (line 155), and finally corrected himself (line 157), by referring first to the fulcrum-helps model (in line 159) and second, to his law-like statement in sequence 1 (line 159):

155 S6: Oh no, man A's having it easier, because the block is supplying more, that block is supporting more - the turning point.
156 I: Uhmm.
157 S6: Right. // No, it *is* man B ... he's supporting most of it because he is right under it [load]// But the turning point is here [in case A] and it is supporting some of it also // But // [Pause 5 s.]
158 I: What are you thinking about - can you tell me?
159 S6: Well, the thing we did before, the division of the length of the board//

He continued this line of thought, referred again to the fulcrum-helps model and concluded (in line 179) that man A would exert less force since:

179 S6: You're dividing the weight by / By a / A larger number [than in case B] because of this distance [effort-fulcrum].

In the remainder of the interviews he used the fulcrum-helps model and the law of levers consistently with coherent and thoughtful explanations.

d. <u>Augmented Non-Generalizable Conceptions</u>. The posttutoring conceptions held by S2 and S3 are alike and include in addition to the pretutoring non-generalizable conception, statements about the two leverarms and the fulcrum's action:

162 S3: I would stick with A [exerting less force], because... even though he is in the same place [as man B] the

weight has moved, has gone further down, which has given [B] more board from the turning point, which would make it a lot harder to pick up than right here [A].

When asked to explain a remark about "holding" the load, he said:

174 S3: ...If you're holding it, you're just holding it and there is nothing on the other side [fulcrum] that's keeping it up.

S2 explained that:

235 S2: The distance from the load to the force / I think if he
236 has less board from the load // When he is *in between* the turning point and the load that would be easier than if there was less board [student's emphasis]. [Thus, smaller effort-load distance.]
237 But when he is outside the load [demonstrated a class II lever with apparatus], I think if there's just a little bit of board that's better than no board at all. [Thus, greater effort-load distance.]

S2 and S3 both proceeded through the sequence as expected, except that S2 applied his class III non-generalizable conception to the bridge 1 situation. He never realized this, but his spontaneous explanation of the class II lever (line 237), demonstrating a correct, non-generalizable conception, suggests that this was probably not an important regression into misconceptions about class II levers.

e. No Conceptual Change. The protocols of three students show either no conceptual change or aspects of regression.

Student S5's overall progress was similar to that of S2 and S3 above, except for her lapses in bridges 2 and 3. It seems that the example where the force is applied underneath the load (bridge 2 and 3), and according to her, "crunching down" on the person exerting the force, was more compelling than the answer suggested by her adequate, non-generalizable conception, e.g. for bridge 3:

276 S5: I think it would be easier to have the load a little bit away from you, not right on top of you because then it

squashes your hand...

She responded correctly to problems where the force was not applied directly underneath the load, which may indicate appropriate, non-generalizable conceptions that were overridden in bridge 2 and 3 by the "squashing, crunching effect" she described.

Student S4 also proceeded as expected through the sequence, until bridge 3 where the class II lever in bridge 2 became a class III lever. He seemed not to notice this difference, and consistently applied the non-generalizable conception for class II levers to all the remaining class III levers, although his response to target problem 1 was correct and accompanied by the naive, non-generalizable conception for class III levers. It may be that he simply did not recognize any difference in the levers in bridge 2 and 3, or that a commitment to consistency caused this phenomena, as is illustrated in the excerpt from his response to bridge 1:

> 230 S4: OK, I think it would be A [exerts less force] *although I always say if the weight's closer it'd be harder...* [My emphasis.]

Thus, it seems that he was aware of his response earlier to target 1. He did not "always" use the non-generalizable conception mentioned, only once (for this specific target) in the entire lesson. It was expected that students would react to the kind of conflict described above by noticing the different "actions" of the fulcrum; i.e. that the fulcrum helps in class II levers and "hinders" in class III levers, and that a comparison of these models would lead to acceptable conceptions (illustrated in S6's protocol before). However, S4 did not refer to the fulcrum, although he had constructed and used the fulcrum-helps model in sequence 1 successfully.

Finally, for the student who most firmly believed in the control misconception, such consistency was rewarded. Student S1 now gave correct answers to the two target problems, but indicated no substantial change in her conception. Yet again it appears that, for her, the problems in the sequence were not related in any way. Her responses to the anchor problems were the same as in sequence 1. She was not able to extend anchor 1 analogically to anchor 2 and probably unable to construct a fulcrum-helps model as a result. Again, her intuitions about some situations were valid (e.g. target problems) and she recognized the difference between levers II and III in bridge 3:

319 S1: // But you're after the load...
323 Like you're between the block [fulcrum] and the load /
And the load is like ahead of you, in front of you...

However, it is clear that her initial control misconception remained essentially unchanged; her last statements still imply control e.g. the load "...may fall if you're not right there..." (line 325).

f. <u>Summary</u>. Again results were mixed: the protocol analyses show that three students changed their conceptions from correct answers with a naive explanation to correct answers with evidence of a more expert-like explanation; two students regressed from a naive conception (accompanied by a correct answer) to the use of an inappropriate conception, and one student's naive conception remained unchanged.

The students' reasoning in this sequence were as expected and one could probably account for instances of failure such as shown in the protocols of S1 and S4 (failure to appreciate a presented analogy and a need for consistent explanation). Again, these two students' responses inform the revision of the lesson, perhaps more so than those of the more successful students.

G. Conclusion

The impact of the formative evaluation of this experiment's lesson will be visible in the design of lesson 2 (for experiment 2) and will be discussed in depth in the description of lesson 2. The most significant findings of the evaluation are summarized briefly below.

Extreme Cases Initiated Model Construction. The few extreme case comparison situations were the most successful elements in the lesson. All these were useful intuitions for the students, including for the most unsuccessful student. These situations were often the anchoring examples for the construction of a model of the fulcrum, particularly in the two cases (S1 and S3) where the anchor-bridging analogies were not accepted by the students. Limiting cases can probably be employed at various stages in the lesson, particularly where more bridging examples or the consolidation of ideas are required.

In addition to the above, it seems that an expansion of the extreme case situations in sequence I and III of the lesson may facilitate *change* in a naive notion that inert objects do not exert forces. This is a potentially important finding, since previous research (e.g. Carey, 1985) has shown that students are often reluctant to exchange naive views; yet one single comparison situation in this lesson initiated such conceptual changes.

Models of the Fulcrum Helped. Most of the participants were able to construct the model of the fulcrum intended by the design of sequence 1, either as a result of the complete bridging sequence or as a result of the extreme cases bridge. The model provides an explanation of the behavior of the levers to the students, and was, to my mind, instrumental in the relative success of sequence 1. However, it is also clear that this model as well as the emerging principle of levers were fragile and transient structures at the end of the class II sequence. These results suggest that consolidation of the new ideas may be essential to meaningful learning.

Fragmented Aspect. Two students did very well in assimilating the different components of the lesson, and were able to "see" all the different levers as governed by one principle. However, the majority were clearly not able to relate the three different sequences, and their final knowledge structures can be described as fragmented. They were still inconsistently using non-generalizable conceptions and misconceptions; a definite indication that the levers in the lesson's problems were still seen to be unrelated in any sense.

This fragmentary aspect probably resulted from my initial views that were informed by the pilot studies; i.e. that the children naively perceived the levers differently, and that the instruction should be designed to build on these different conceptions to eventually merge them into one principle. There was evidence that the principle could be constructed, particularly after the first teaching sequence. The students somehow "lost" these useful bits of knowledge in the next two sequences, where new misconceptions were actually acquired by a couple of students. Hence, a more holistic approach in the lesson seems to be indicated.

Naive Models may be Important. The results add to the evidence that the students' pre- and misconceptions of levers may be more sophisticated than thought at the lesson's design stage. This was already suggested by the analysis of the pre-and posttest protocols, as I have suggested in before. The students may have naive "force diagram" models of the non-generalizable conceptions and misconceptions in this content area, and the fact that the non-generalizable conceptions in particular contain some elements that are in agreement with a physical theory, may hamper rather than support the instructional design. In other words, how do you convince a person that a perfectly useful conception, with which one is able to explain the behavior of levers and to predict correct answers, may not be entirely "correct"?

CHAPTER IV

EXPERIMENT 2: RESULTS AND DISCUSSION

A. Introduction

The research design for experiments 1 and 2 was the same: the pretest interviews and the first part of the lesson interviews were conducted on day 1 and the rest of the lesson was followed immediately by the posttest interviews on day 2. The pre-and posttests were standardized across the study, and care was taken to ensure the same context and equipment used for illustrations during the lesson interviews.

In view of the above, experimental group 1 is used as a control group in the summative evaluation of experiment 2. The questions that concern one here are similar to those posed in experiment 1, thus - are there discernible pre-posttest differences between the two groups with respect to conceptual change and the far transfer abilities of the students? One can assume that differences in these two areas may be attributable to either of the two interventions, given that the possibility of learning outside of experiment 2 is kept in mind, as in experiment 1.

Analyses of the protocols of those students who changed their preconceptions and were successful in the far transfer questions are presented first in the formative evaluation of lesson 2. The success of the teaching strategies is evaluated with regard to the construction of the fulcrum-helps model in the teaching sequence for class II levers;

the role of this model in students' increased understanding of class III levers; the separation of their naive variables into the leverarm variables; the processes by which conceptual change was fostered in the lesson; and the students' increased understanding of levers evident in their responses to the far transfer questions.

The apparent failures of the lesson will be discussed in the final section.

B. Summative Evaluation

1. The Simple Levers: Pre- Posttest Analysis

The simple levers questions are again given in Figure 4.1 below for reference. The students were asked to compare the two situations and predict which lever would be 'easier" to hold level with the 50 lb load on it.

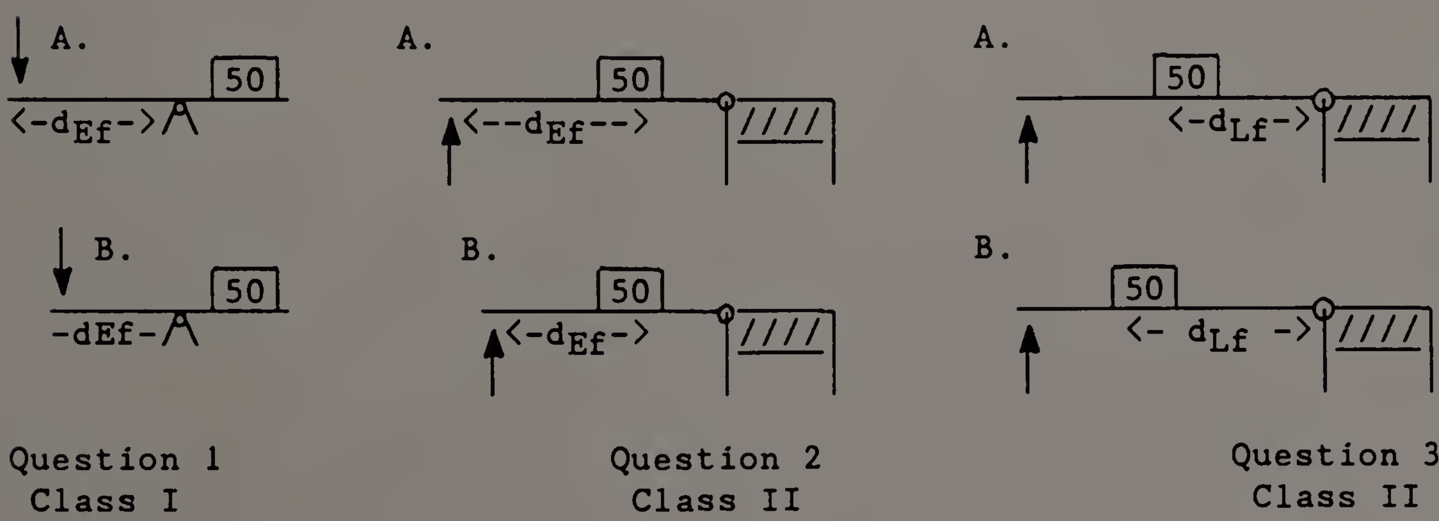


Figure 4.1

Pretest / Posttest Simple Levers Questions

a. <u>Quantitative Analysis</u>. There are at least two different issues of importance in the analysis of the pre- and posttest performances on the simple levers questions. One, are there differences with respect to conceptual change between the two groups? A comparison of the pre- and posttest scores on the simple levers questions should speak to this question. Two, is there evidence of differences between the two groups with respect to students' use of the levers principle and decrease in use of misconception and non-generalizable conceptions?

The students' scores were computed by assigning positive and negative values to correct and inappropriate (misconceptions and non-generalizable conceptions used in explanations) answers respectively; assigning a number (1 to 4) to the confidence level (rated from "a guess" to "sure" on a four point scale); and multiplying the confidence level number with the appropriate symbol to indicate an answer's status; thus the same procedure as that used in experiment 1. In Tables 4.1 and 4.2 the scores for students on the simple levers questions are given.

Table 4.1

Simple Lever Scores: Pre- and Posttest
Experimental Group 1

Student	Pretest Question			Posttest Question			Summed Changes in Score
	1	2	3	1	2	3	
S1	-3	+1	-3	-3	-3	-3	-4
S2	-3	-3	+4	+4	+4	+4	+14
S3	+4	-3	-3	-4	-4	-4	-10
S4	+3	-3	-3	-3	-3	-3	-6
S5	-3	-3	-3	-2	-2	-2	+3
S6	-3	-3	-3	-4	+4	+4	+13

Table 4.2

Simple Lever Scores: Pre- and Posttest
Experimental Group 2

Student	Pretest Question			Posttest Question			Summed Changes in Score
	1	2	3	1	2	3	
E1	-4	-3	-3	-3	-2	+3	+12
E2	-3	-3	-4	+4	+3	+4	+21
E3	-3	-1	-2	+4	+3	+3	+16
E4	-3	-3	-2	-3	-2	-2	+1
E5	-3	-3	-3	+3	+3	+3	+18
E6	+3	+3	-3	+4	+4	+4	+9

Using a criterion of $p < 0.050$, the difference between the groups is significant (for $U_{calculated} = 7$, $p < 0.047$), indicating that the experiment 2 students were at least able to give better predictions for the posttest simple lever comparisons.

Use of Conceptions. An analysis of the conceptions used in the students' explanations may again provide a more illuminative quantitative evaluation than the comparisons of performances on pre-posttest questions above. The effects of the two lessons are therefore compared in terms of the increase in use of appropriate conceptions and the decrease in the use of alternative conceptions.

Summaries of students' explanations are provided in Tables 4.3 and 4.4. The abbreviations are the same as before, and one new abbreviation is used: $m.c_a$ indicates a misconception as a result of a negative analogy (appearing in the protocols of students E1 and E2).

Table 4.3

Experimental Group 1:
Conceptions Used for the Simple Levers Questions

Student	Pretest Questions 1	Pretest Questions 2	Pretest Questions 3	Posttest Questions 1	Posttest Questions 2	Posttest Questions 3
S1	− m.c.$_c$	+ guess	− non-gen	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$
S2	− non-gen	− m.c.$_c$	+ model	+ [principle; model, principle]	+	+
S3	+ principle	− non-gen	− m.c.$_c$	− m.c.$_c$	− non-gen	− non-gen
S4	+ principle	− m.c.$_c$	− non-gen	− m.c.$_c$	− non-gen	− non-gen
S5	− non-gen	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$
S6	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	− [principle; model, principle]	+	+

Table 4.4

Experimental Group 2:
Conceptions Used for the Simple Levers Questions

Student	Pretest Questions 1	Pretest Questions 2	Pretest Questions 3	Posttest Questions 1	Posttest Questions 2	Posttest Questions 3
E1	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	− m.c.$_a$	+ model	+ model
E2	− m.c.$_a$	− non-gen	− non-gen	+ [principle; model,principle]	+	+
E3	− non-gen	− m.c.$_c$	− non-gen	+ [principle of levers]	+	+
E4	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$
E5	− m.c.$_c$	− m.c.$_c$	− m.c.$_c$	+ [principle and model]	+	+
E6	+ [naive: leverage]	+	− non-gen	+ [principle of levers]	+	+

A cursory inspection of the summaries seems to suggest that lesson 2 was more effective in changing students' conceptions. There are four apparent changes from naive conceptions (misconceptions and non-generalizable conceptions) to the use of the principle of levers and a fulcrum-helps model; one person (E1) seems to have changed her ideas about class II levers and one person (E4) who did not change a predominantly misconception view. In contrast, only two students in group 1 added to or changed their naive preconceptions.

The comparison between the two groups can be made on a more rigorous basis than in the experimental/control group comparison in Chapter III. The lesson's goal for both experimental groups was essentially the same: to facilitate change in students' naive conceptions to a view of levers more in agreement with that of a physicist's. In this analysis the non-generalizable conception as well as misconceptions will be considered inappropriate or naive. The pre- and posttest use of conceptions by students in the two experimental groups are given in Tables 4.5 and 4.6.

Table 4.5

Inappropriate Conceptions Used: Simple Levers Pre- and Posttests

	Pretest	Posttest
Experimental Group 1	15	13
Experimental Group 2	18	4

A hypothesis that lesson 2 was more effective in reducing the use of naive conceptions by students is accepted (for df = 1; $X^2 = 4.38$; $p < 0.05$).

Table 4.6

Appropriate Conceptions Used: Simple Levers Pre- and Posttests

	Pretest	Posttest
Experimental Group 1	3	5
Experimental Group 2	0	14

Again, a hypothesis that lesson 2 was more effective in teaching a view more compatible to that of a physicist, can be accepted (df = 1; and $X^2 = 6.09$, $p < 0.02$).

Summary of Conclusions from Quantitative Analysis. There are significant differences between the two experimental groups' performances on the simple levers questions. Indications are that lesson 2 was more effective in teaching about class I and II levers and in facilitating changes from naive conceptions to appropriate conceptions of levers in students.

Again, one should view these results with caution: the groups are small, thus making quantitative results more tenuous and the students' responses to the simple levers questions alone cannot be taken as an indication of a deep understanding of levers, as stated before in Chapter III.

b. Qualitative Analysis. In the experimental group 1 comparison to the control group in Chapter III a thorough qualitative analysis of the protocols was provided to enable a discussion about possible instances of learning in the tests. Inspection of Table 4.4 shows that there are no discernible changes in the group 2 students' conceptions in the pretest. One can therefore accept that the changes evident in the summary table are a result of either participation in the tutoring interviews or of external learning experiences. A qualitative analysis of the simple levers pre- and posttest protocols could not provide evidence towards deciding the last two issues (causes for conceptual change) and is therefore not provided.

2. The Far Transfer Questions

The simple levers analysis seems to suggest that lesson 2 was more effective in bringing about conceptual changes. Students' success in the far transfer questions are seen to be a measure of a deeper understanding, and of the robustness of the newly constructed conceptions. I shall assume at this stage that the lesson was entirely responsible for the changes in conceptions and for any subsequent far transfer abilities. Only an analysis of the target problems within the lesson will indicate whether or not any external experiences could have brought about some conceptual changes.

The experimental conditions, such as time lapse between the pretest and the posttest were the same for both experimental groups: the posttest interviews were conducted immediately after the lesson interview in both experiments. One may therefore compare the groups with respect to their performances on the far transfer questions, and hypothesize that differences in students' far transfer ability may be attributed to either of the two lessons.

a. Quantitative Analysis. The groups will be compared with respect to: one, the use of appropriate (e.g. principle of levers) and inappropriate (e.g. non-generalizable) conceptions across the four far transfer questions; and two, students' summed scores on the questions computed by taking into account the correctness of the explanation, the level of lever recognition and a student's confidence in his answer. Both of these measures should give an additional overview of possible differences between the second and first lesson's impact on students' learning.

Comparison of Appropriate and Inappropriate Conceptions Used. A summary of the conceptions used across the four far transfer questions is presented in Tables 4.7 and 4.8. The abbreviations are the same as those used before.

Table 4.7

Summary of the Conceptions Used in the Far Transfer Explanations: Experimental Group 2

Student	Revolving Door	Nutcracker	Shadoof	Nail Clippers
E1	− surface feature	− $m.c._c$	+ non-gen	− surface feature
E2	+ [principle and model used in explanations]	+ (continued)	+ (continued)	− surface feature
E3	+ [principle of levers used in all the explanations]	+ (continued)	+ (continued)	− (continued)
E4	− surface feature	+ principle	− [surface features	−]
E5	+ [principle of levers in explanations]	+ (continued)	+ (continued)	+ lever recognition
E6	+ [principle of levers and model used in explanations]	+ (continued)	+ (continued)	+ (continued)

Table 4.8

Summary of the Conceptions Used in the Far Transfer Explanations: Experimental Group 1

Student	Revolving Door	Nutcracker	Shadoof	Nail Clippers
S1	− surface feature	− $m.c._c$	+ surface feature	− $m.c._c$
S2	+ [used principle of levers throughout	+ (continued)	+ (continued)	−]
S3	− [used surface feature explanations throughout	+ (continued)	+ (continued)	−]
S4	+ surface feature	− m.c.	− m.c.	− surface feature
S5	− surface feature	+ non-gen	+ principle	− surface feature
S6	+ [used principle of levers throughout	+ (continued)	+ (continued)	−]

For this analysis, the principle of levers and a model of the fulcrum will be considered acceptable in students' explanations, while misconceptions, non-generalizable conceptions and surface feature explanations are viewed as inappropriate.

From the two summary tables above it appears that there were fewer explanations based on inappropriate conceptions in group 2 than in group 1 (distribution shown in Table 4.9).

Table 4.9

Conceptions Used in the Far Transfer Explanations: Experimental Groups 1 and 2

	Surface Features Non-gen and Misconceptions	Levers Principle Model, and Lever Recognition
Experimental Group 1	15	9
Experimental Group 2	8	16

One may reject a null-hypothesis and assume that there are significant differences between the two groups with regard to explanations containing inappropriate or appropriate conceptions ($df = 1$; $X^2 = 4.09$ and $p < 0.05$), thus that lesson 2 facilitated the use of appropriate conceptions in the participants. An analysis of the responses to the different questions may indicate further differences between the two groups.

Comparison of Groups. Each student's scores in the far transfer questions will be used in a comparison of the groups. The "correctness" of the students' answers will be determined both by explanation and appropriateness, since the use of surface features and a misconception may give one a correct, acceptable answer. The confidence level shall again be used to indicate students' commitment to an answer, ranging from 1 for a "guess" to 4 for "sure".

However, the explanation will determine the correctness (+ or -) associated with the response. An answer will be scored "correct"(= +), when one, the explanation accompanying the correct response is based on the principle of levers, and two for the nail clippers problem only - if the student provided evidence of lever recognition and gave an acceptable explanation. An answer will be scored "incorrect"(= -), when one, the response is correct but accompanied by an inappropriate explanation, and two, when the response is incorrect and accompanied by an inappropriate explanation.

In addition to the above, students' spontaneous recognition of levers will be scored. Instances where the elements of a lever (load, effort and turning point) are identified spontaneously (i.e. without interviewer probes) will be regarded as evidence of lever recognition. I suggested in the discussion of transfer in Chapter III, that a prerequisite for an understanding of levers would be the ability to analyse the complex and compound levers in the far transfer questions into the basic components, thus an ability to recognize levers. There are two criteria here: students should *spontaneously* identify the levers, and all three the basic elements (effort, load and fulcrum) should be identified *correctly*. A student could therefore score an additional three points per question.

The students' summed scores on the far transfer questions are given in Table 4.10.

Table 4.10

Summed Scores: Experimental Groups 1 and 2.
Far Transfer Questions

Student	Experimental Group 1	Experimental Group 2
1	-11	0
2	+22	+12
3	-16	+26
4	-15	+2
5	+1	+22
6	+20	+22

A null-hypothesis can be rejected at a set level of $p < 0.050$, since for $n_1 = n_2 = 6$, $U_{calculated} = 7$ has a probability of occurrence under H_o of $p = 0.047$ (Mann and Whitney, 1947). It does seem therefore that the group 2 students were more able to identify levers and to give more acceptable explanations for the far transfer questions than the group 1 students. I interpret this to mean that participants in the tutoring interviews on lesson 2 indicate a deeper understanding of levers than group 1.

b. Qualitative Analysis. In view of the quantitative finding above, that the participants in experiment 2 were apparently more successful in the transfer of knowledge, it may be more useful to provide a detailed discussion of the analyses of the far transfer protocols in conjunction with a discussion of the learning processes in the lesson. In other words, to try to relate more directly the ability to transfer knowledge to the processes by which the knowledge was constructed as observed and interpreted from the lesson protocols.

Only one qualitative analysis will be presented, and that is a trace of the development of conceptions over the course of the experiments.

The coding is the same as that used in Chapter III:

* surface feature
- misconceptions (control, symmetry and due to a negative analogy)
n non-generalizable conception
+ principle or fulcrum-helps model
|\| instructional intervention (tutoring interviews)

Table 4.11

Trace of Conceptions Used over Time:
Experimental Group 1

	Pretest Simple Levers				Posttest Simple Levers			Posttest: Far Transfer Questions			
								II/III	II/II	III/I	II/III/III
	I	II	II		I	II	II	door	nuts	shadoof	clippers
S1	-	-	n	\	-	-	-	*	-	*	-
S2	n	-	n	\	+	+	+	+ +	+	+ +	+ +
S3	+	n	+	\	-	n	n	*	*	*	*
S4	+	-	n	\	-	n	n	*	-	-	*
S5	n	-	-	\	-	-	-	*	n	n	*
S6	-	-	-	\	-	+	+	+ +	+	+	+

Table 4.12

Trace of Conceptions Used over Time:
Experimental Group 2

	Pretest Simple Levers				Posttest Simple Levers			Posttest: Far Transfer Questions			
								II/III	II/II	III/I	II/III/III
	I	II	II		I	II	II	door	nuts	shadoof	clippers
E1	-	-	-	\ \	-	+	+	*	-	n	*
E2	-	n	n	\ \	+	+	+	+ +	+ +	+ +	*
E3	n	-	n	\ \	+	+	+	+ +	+ +	+ +	+ + n
E4	-	-	-	\ \	-	-	-	*	+	*	*
E5	-	-	-	\ \	+	+	+	+ +	+ +	+ +	+ + +
E6	n	n	n	\ \	+	+	+	+ +	+	+ +	+ +

The posttest changes evident in the group 2 students' explanations, particularly those instances of overall change indicated for E2, E3, E5 and E6, appear to be normative since all four students changed to appropriate conceptions. In addition, group 2 students used the appropriate conceptions in a more consistent fashion across the far transfer questions than the students in experiment 1. One may also hypothesize that there is evidence of an increased understanding of levers in some group 2 participants, since all four successful students in experiment 2 were able to transfer their changed conceptions to situations that are perceptually far from the simple levers in the lesson and pretest. In contrast, only two students in experiment 2 performed comparably well.

This analysis over the pre- and posttest domain supports the quantitative findings: there is evidence that lesson 2 was more effective than lesson 1.

C. Lesson 2

I shall present in this section a personal model of the learning processes expected in the new instructional sequence. The lesson's design was informed by the findings from lesson 1, and it is therefore useful to review the major suggestions.

The students' naive conceptions may be more complex and general than the pilot studies indicated. Two of these naive conceptions are described as non-generalizable, meaning that in some lever problems the students are able to predict the correct answers and explain their answers satisfactorily with the non-generalizable conceptions. Thus, these are truly alternative rather than misconceptions and present a unique problem: how to reconcile the alternative knowledge structures with the accepted physical theory views, when the existing view cannot be considered "wrong". The complex nature of the class III lever non-generalizable conception presents another dilemma - it is essentially the control misconception from the class I and II levers. Hence, students have strong intuitions for one class of levers that are useful and aligned with the physical theory view, but that are, at the same time, misconceptions for the other two classes of levers. The repercussions for the construction of a unified, acceptable principle of levers are clear.

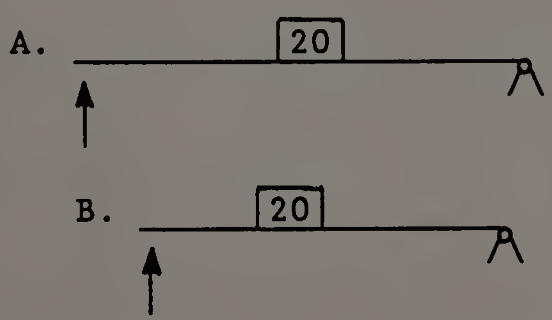
One goal of the first lesson was to encourage the construction of a model of the fulcrum's action. This qualitative model was "lost" after the first teaching sequence - only two students grasped the significance of this way of thinking about the fulcrum in the rest of the lesson. Since there is enough evidence to suggest that most students were able to construct the qualitative model, one should be able to use this knowledge more effectively in the rest of the lesson.

The use of the fulcrum-helps model in a more integrative fashion should also circumvent the difficulties with the fragmentary aspects of the first lesson. There the idea that one could use students' naive conceptions to construct an appropriate conception in each of the three classes and then aim for unification resulted in more fragmentation — two students actually acquired misconceptions in this process.

Finally, the students' strong, intuitive beliefs about extreme cases in lever situations, coupled with their ability to reason from the limiting cases to more general situations and to construct general qualitative models on the basis of the extreme cases alone, are unexpected and potentially very useful findings.

1. Drawing out Misconceptions

The lesson starts with a target example, similar to that in lesson 1. The diagnostic test results indicated this to be the most likely situation to draw out the control misconception (belief score = 31%). The target question is given in Figure 4.2.



Target Question:

Which person would find it easier to keep the board with the 20 lb load on it level?

Figure 4.2

Target Question: Class II Levers

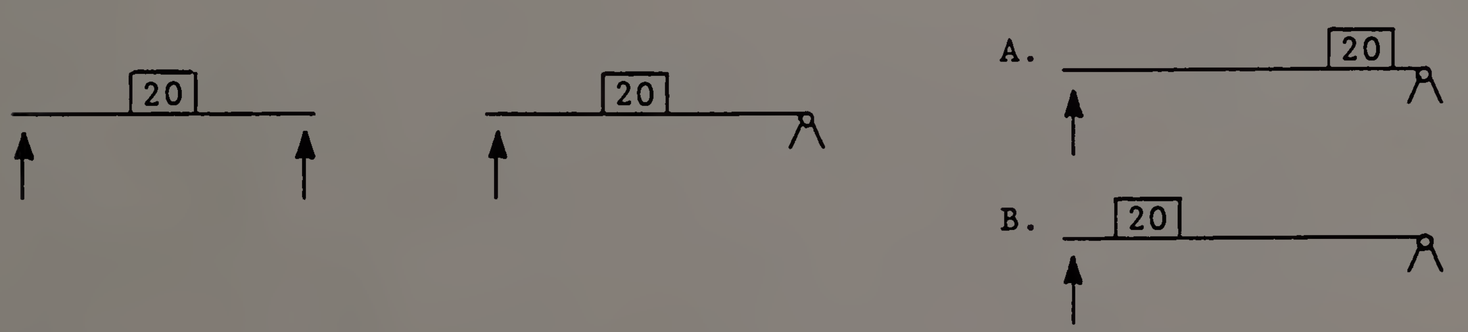
Most students should give predictions to the target question based on the control misconception or the non-generalizable conception for class II levers - thus, similar to experiment 1.

2. Constructing a Qualitative Model

The lesson 1 findings showed that the students are likely to construct a fulcrum-helps model in either the bridging sequence used in experiment 1 or by means of the extreme case examples.

a. The Bridging Sequence. Similar to the technique used in lesson 1, a useful intuition of sharing a load is extended analogically to a bridging example, and then consolidated in the extreme case situations.

The idea that the fulcrum is "helping" or "holding" 10 lbs in the bridging example like the second person in the anchoring example, is therefore reinforced by the belief that, in case A in the extreme case situation, the person will have to push up very little, since the fulcrum is "holding" most of the load. The sequence is presented in Figure 4.3.



Anchor Bridge Extreme Case Examples

Figure 4.3

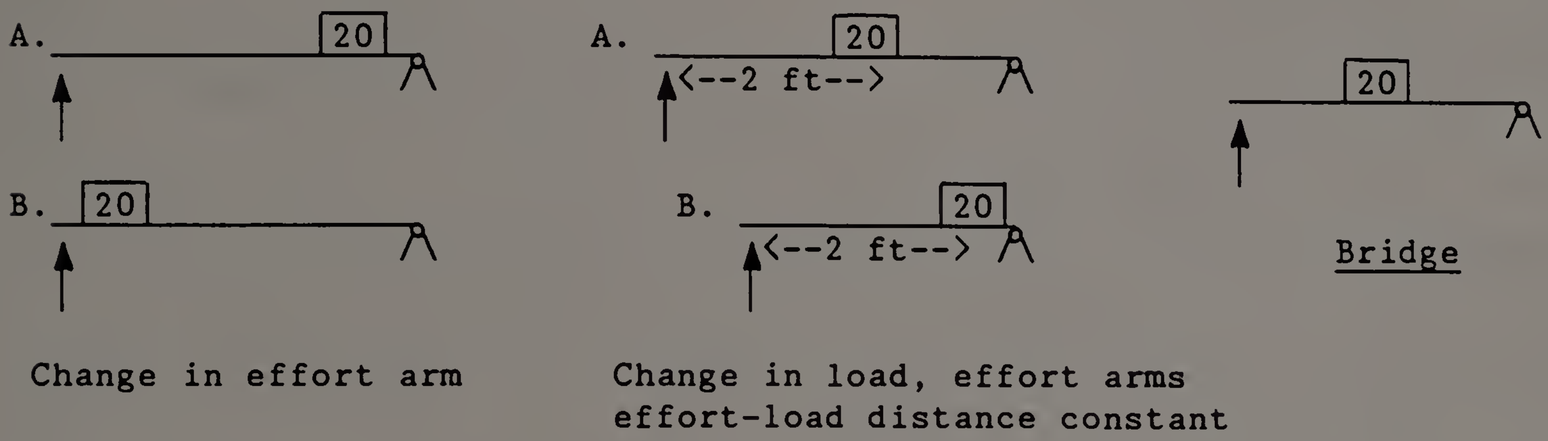
Constructing a Fulcrum-Helps Model: The Bridging Sequence

b. <u>The Extreme Cases Revisited</u>. When a child holds a strong misconception that inert objects cannot exert forces, there is reason to believe (results from experiment 1) that an equally strong and more useful intuition about the fulcrum exerting a force in the extreme case situation, may initiate the qualitative model's construction and change the misconception. However, experiment 1's results suggested that these seemingly fragile ideas need reinforcement, and the sequence was therefore changed from this point on.

Lever Terminology. The short sequence to teach lever terminology was expanded, e.g. name the fulcrum (the turning point, since the lever 'turns" around this point); name the distance from the load to the fulcrum the load turning arm, etc. The last component, the load turning arm, and similarly the effort turning arm, were two new terms. I hoped that the words would keep the distances specified (load ->turning point distance), perhaps minimizing confusion. Three exercises in simply labeling the elements of levers were added: two of those are drawings of levers (a long bar used as a crowbar and a deepsea fishing rod) and the third, a door being opened, is more difficult to categorize (the lever explanations are given in Appendix E).

Repeat Extreme Case Examples. The extreme cases were revisited to reinforce the students' ideas about the fulcrum, and in addition, to make the first attempt at differentiating the class II non-generalizable conception into the principle of levers. The children who held the inert-objects-do-not-push misconception in the bridging example were given another chance to view that problem, mainly to gauge the influence of the extreme cases on the misconception's status.

The extreme cases revisted sequence is given in Figure 4.4. The following questions are asked for the extreme case comparisons: Which person would find it easier to keep the board with the 20 lb load on it level? Does it matter what the distance from the load to the effort is?



Extreme Case Examples

Figure 4.4

Revisiting the Extreme Cases

The second comparison question is difficult, since two variables (the effort and load leverarms) are changed. However, the children appeared to have the strongest intuition about case B in this question, and the hope is that they would be able to reason from this certainty about the effect of the load leverarm rather than focusing on the effort-load separation. One may then refer back to the first extreme case question to initiate conflict about the non-generalizable conception, if the students have not noticed this spontaneously. The students with the misconception about a fulcrum that would not exert a force should now have changed this conception upon reviewing the question again.

The class II near transfer questions follow the extreme cases revisited sequence. The lesson 1 near transfer questions are used here also. The wheelbarrows question is seen to reinforce the extreme case situations and the fulcrum-helps model; while the bottle openers question was very difficult for the lesson 1 participants, and success on this question may be an early indication of the success of the lesson up to this stage. (See Appendix E for the near transfer questions in lesson 2.)

c. <u>Using Extreme Cases in Transformations</u>. Moving the class III transformation sequence from the end to the middle of the lesson, and changing the content of the sequence significantly represent the most incisive changes to lesson 1. Two issues are at stake here, i.e. the fragmentary aspect of lesson 1 and a more economical approach. The class II to III transformation evolves more logically and economically at this point - the sequence is designed to build on students' use of the fulcrum-helps model and extreme case situations from the questions above, as follows:

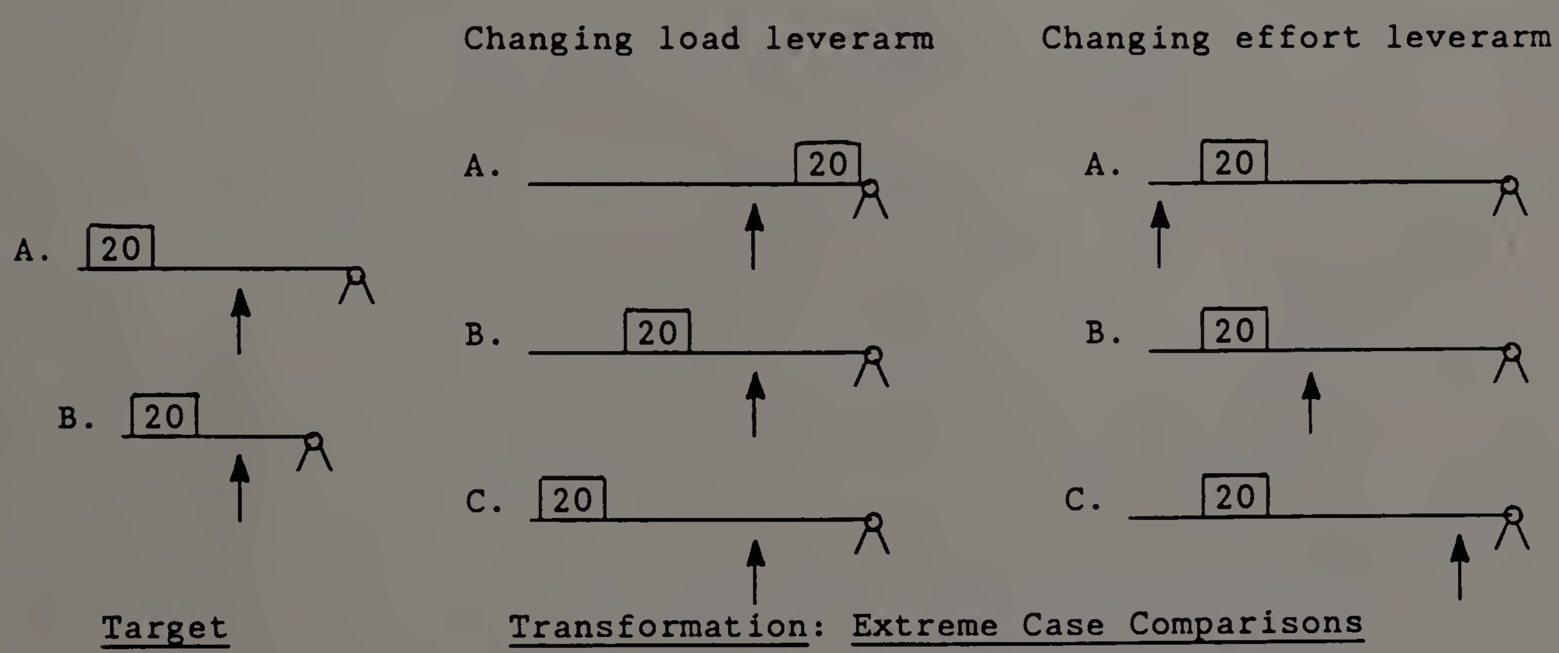


Figure 4.5

Transformation of Class II to Class III Levers

The following question is asked in the two extreme case comparison problems: Which person (looking first at A or C) will find it easier to keep the board with the 20 lb load on it level? Looking at B or C, which person will have an easier task?

Cases A and C in both of the transformation sequences are expected to be strong intuitions for the students, and they may use the fulcrum-helps model to explain their predictions about the "easier" task. The certainty about the extreme situations in A and C should then enable them to reason about the intermediate cases in B. Again, changing the load and effort leverarms should focus attention on those variables rather than on the load-effort distances.

The target situation is presented again and followed by the near transfer questions. The steam shovels comparison from lesson 1 is used, but the mechanical rakes question, that required detailed explanations before students understood its function, is replaced by the visually simpler comparison of two persons using identical brooms (a description of the near transfer questions is given in Appendix E).

d. Benchmarks for Class I Levers. With the exception of the target question, this benchmark sequence remains unchanged from lesson 1. The target question was changed so that the class II non-generalizable conception would be a misconception if applied to the situation.

The sequence's major goal, the separation of the misleading load-effort variable into the load and effort leverarm variables, may be easier to attain: the load leverarm, depending so much on the students' view of the fulcrum "helping" in the class II lever sequences, should be an independent entity at this stage.

The sequence is presented in Figure 4.6. Students are asked to estimate the "push" to keep the lever with the load on it level in each situation.

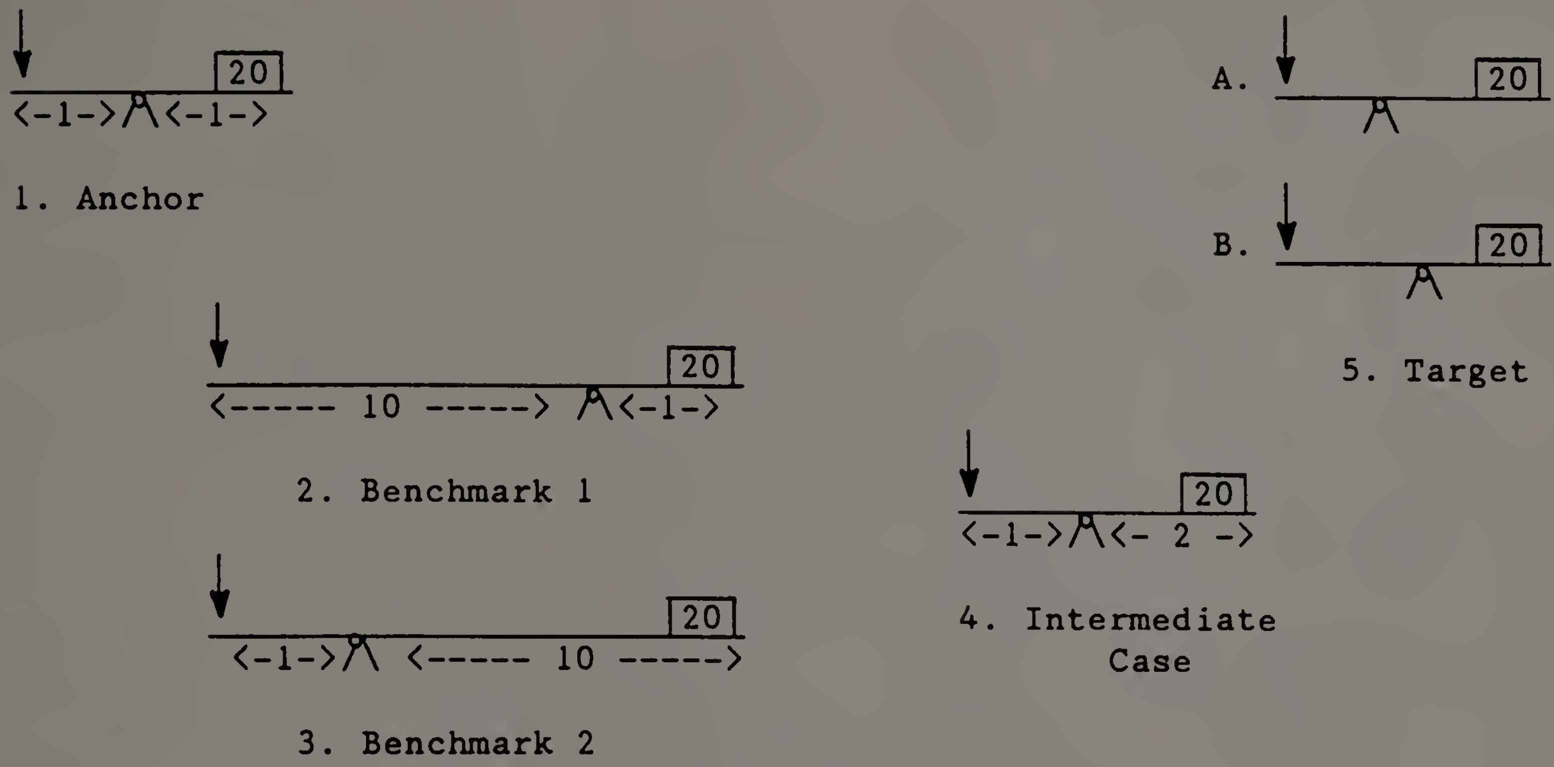


Figure 4.6

Benchmark Sequence for Class I Levers

The near transfer questions complete the lesson. One transfer question, a comparison of a pair of pliers, is different from that in lesson 1. The original luggage carts problem was too farfetched and required too much explanation before students understood the problem. The pliers are a compound lever (consisting of two class I levers) and may be more difficult than the other simple levers applications - however, it is expected that, at the end of the lesson, the students would either be able to reason about such a lever or rely on surface features in any case.

e. Opportunities to Write. The purpose of the writing opportunities is different in lesson 2. Rather than possible conflict generation situations, the statements are simply seen as summaries of what the students believe at that point in the lesson. I expected that most students would hold a fulcrum-helps model and one part of the principle of levers (a shorter load leverarm preferred) at the end of the class II teaching section and would have added the idea that the effort leverarms are important variables by the end of the lesson. The thought was that they could review all their previous answers to write their summary statements, thus providing opportunities to unite thoughts and perhaps discern the underlying principle of levers more clearly.

3. Summary

The new lesson is different in that the teaching of a single underlying principle of levers is aimed at. Students' construction of a causal, explanatory qualitative model of the forces exerted in the lever situations is seen to be the starting point. The instruction is designed to facilitate the model's construction in one of two ways: a bridging sequence or by using extreme case reasoning about class II levers. The students should then use the model and extreme cases in the other classes of levers to generate the principle of levers in a stepwise fashion.

D. Students' Knowledge Construction

The previous definition of an explanatory, qualitative physical model is broadened in this section. The aim of lesson 2 is the direction of students towards the construction of such a model, before further instruction could lead to the generation of a qualitative principle of levers.

The hypothetical model of learning is "evaluated" in this section by reconstructing the children's learning processes from their explanations. The summative evaluation of experiment 2 suggested that four students had learned a qualitative principle of levers from the lesson, while two had retained their misconceptions but indicated some far transfer abilities that suggested some understanding of the content matter. I will therefore focus on evidence about elements that make the lesson successful, as well as information about those components that were ineffective. The model construction processes of the four most successful students will be discussed first, and with regard to the following issues. First, and most important, are the models intuitively anchored, and if so, is there evidence to support the hypothetical model construction processes proposed before? Could one say that the students' qualitative levers principles emerged from the model, and is there any evidence to suggest that the principles are more than "rules"? Are there causal explanatory aspects to the models? Do these models facilitate conceptual change and the reconciliation of non-generalizable conceptions to the accepted physical views?

The model construction process in lesson 2 hinges upon extreme case reasoning. It seems reasonable to expect (from lesson 1 findings) that the extreme case examples would provide the grounding or anchoring of such a model, but why would they? Thus, why are the extreme cases so powerful; can one reason about their power in a more principled way? I shall review two papers concerned with extreme case reasoning in experts and try to find some evidence in the children's lever protocols to support or change some of the suggestions made by the authors. Finally, analyses of the children's explanations for the far transfer problems in the posttest will be presented. Transfer abilities will again be taken as a measure of a deeper understanding of the levers content.

1. Model Construction in Lesson 2

I shall assume in this discussion the definition of intuitively anchored models proposed in Chapter II. The fulcrum-helps model is seen to be grounded in the intuitive belief that the fulcrum pushes up; the model explains why, in the simple lever cases, one has to "push up less" when the load is moved towards the fulcrum. The definition of such a physical model will be augmented in the course of the discussion if needed.

The lesson is designed to facilitate the students' construction of the model in two possible ways: first via the analogical bridging sequence and second, by means of extreme case examples.

a. Constructing a Fulcrum-Helps Model: The Bridging Sequence. A summary of students' conceptions across the bridging sequence is presented in Table 4.13 as before. The abbreviations are the same as those used before, and I have added a distinction between two of the principle's "elements": principle 1 shall refer to the load leverarm and principle 2 to the effort leverarm.

Table 4.13

Successful Students' Conceptions: Bridging Sequence

Student	Target	Anchor	Bridge	Extreme Cases	Target
E2	+ non-gen	+ sharing	− m.c.	+ non-gen & f-h model	+
E3	− m.c_c.	+ sharing	− m.c.	+ f-h model	+ m.c_c.
E5	+ m.c_c.	+ sharing	+ sharing	+ m.c_c.	+ m.c_c.
E6	+ non-gen	− confusion	− m.c.	+ non-gen, model, principle1	+

Preconceptions. There were no surprises as far as the naive conceptions of the four successful students were concerned: two (E3 and E5) held control misconceptions, i.e. that a smaller effort-load separation results in more control over the load; and two (E2 and E6) gave non-generalizable explanations, thus that a greater load-effort separation would result in less effort:

> 001 E2: In case A [it is easier], because the 20 lbs is farther away from the person.

> 001 E3: For this one I think it's B, because you have to lift it up and for A it's longer [the board].

> 003 E5: Man B...and 'cause well, I thought like the other one [in pretest] you're nearer the weight you know?

> 005 E6: The man in A does not have to lift the 20 lbs directly... the man in A is lifting farther away from [the load].

Model Construction? The bridging sequence is given in Figure 4.7 below.

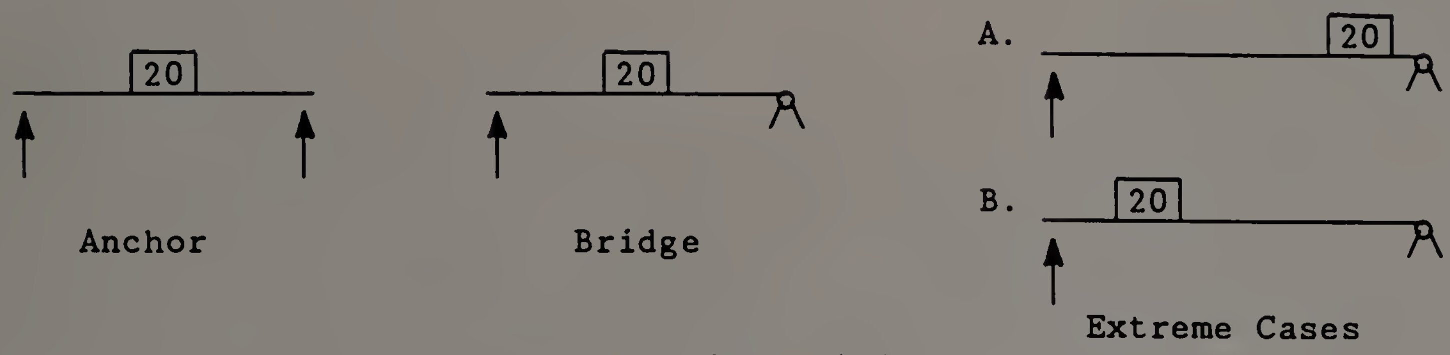


Figure 4.7

The Bridging Sequence: Class II Levers

The fulcrum-helps belief is dependent on an analogy in this short sequence: the students have to imagine the fulcrum as "the same" as the second person in the anchoring example, to be able to reason that the fulcrum pushes up, or holds, or supports part of the load's weight in the bridging example.

Three of the students were not able to construct the model from the anchor-bridge sequence alone. One student accepted the anchor-bridge analogy and became aware of a conflict between her target question explanation and that given for the extreme case comparisons, yet retained the misconception in both her final extreme case explanation and the second target explanation. She held the sharing conception for the anchoring example and accepted the analogy between the person and the "thing" as she referred to the fulcrum:

015 E5: Well/ I think // I *think* [her emphasis] it'll be 10 lbs, you know? 'Cause it's 20 lbs, but / I'm not sure because I'm not holding it up you know?
016 I: Can you tell me, how come they each have to push 10 lbs?
017 E5: Well, if they're each pushing 10 lbs then together they push 20 lbs...So I guess that...I'm not very confident.

And for the bridging example:

027 E5: Well I guess it's 10 lbs again?

028 Because I guess the thing [fulcrum] would hold the other 20 lbs up you know?
029 I: The?
030 E5: The other *10* lbs. [Marked not very confident on scale.]

Her reasoning about the extreme cases is interesting: she seems to be aware of a conflict between the intuition that some weight may be resting on the "thing", but gave as the final explanation a new variation of the control misconception:

036 E5: Hmm. [Pause 10 s.]
037 I: That one is making you think?
038 E5: Well, I don't know. I guess it's harder than the other ones.
040 I: Can you tell me what makes you think that?
041 E5: I don't know, it's just / *This* [tapped on example A] / I think [for] person A, the weight is on this thing, resting on it? But, then here [B] he is holding the weight, so. I think person B though...
048 It'll be easier for him [B] to keep the weight from moving the board.
049 I: From? To keep the weight from?
050 E5: Yeah, from holding // From putting it down / Letting the board go down / From pushing the board down.
051 I: Oh, OK. Could you tell me, in A would the 20 lbs push the board down?
052 E5: No, but [giggles] / I don't know why!

Her final explanation, from line 048 onward seems to be similar to explanations given by two other successful students (E2 and E6) to explain why the fulcrum in the bridging example could not exert a force, and it may be different manifestations of the control misconception *or* of the inert-objects-do-not-push misconception:

016 E6: They have to lift the full 20 lbs...
021 It's held on one side [B?], but if it / If the guy wasn't here [A], then the 20 lbs would bring it down like that [moved hand down in arc to show board falling].

021 E2: He'll have to push up 20 lbs, because there's // Oooh, hold it. Oh no, because the / Uhm, the hinge will not hold up // Anything, it will just let it drop, so he's going to have to hold up all of the 20 lbs.

I interpret these statements to mean that only the person has control and could prevent the load from "letting the board fall". This is really a peculiar idea, and the students were uncomfortable (similar to E5 in line 052) when probed about the same phenomenon occuring in case A.

Students E2, E3 and E6 rejected the anchor-bridge analogy; E6 was so convinced of the inert objects do not push idea, that he changed his initial correct anchor answer in an exchange that followed the excerpt from line 021 above. E3 held the more conventional misconception idea that the hinge would not "really hold it [load] up".

What is exceptional is that, for these three students who rejected the anchor-bridge analogy, the extreme case examples were apparently quite powerful in facilitating the idea that the fulcrum, exactly the same inert fulcrum from the bridging example, might "help" the person:

036 E2: [Nodded head in agreement while she read the question.] Oh, ah, in case A, because the lbs is much farther from him and so // [marked question and elaborated spontaneously]. The block is taking more of the weight. And in case B the 20 lbs is right on his hands basically and I'm sure I'm right.

030 E3: And I think that / Person A will [find it easier]. And it's different from what I said before, but I think that / Maybe the // The triangle would help a little. If the 20 lbs is there [B], nothing would help the person, it's like carrying the thing. And if anything *could* help keep the weight up, then / Then I guess it would be easier closer to the triangle...And / I'm guessing on that one.

054 E6: A. Man A [will find it easier] // The reason is because / This man [A] is lifting farther away from it [load] so the 20 lbs is focusing more on the triangle's side.
055 I: What do you mean "focusing more"?
056 E6: Well, I would / All the weight, if it was spread out on a 20 lb bar, then it would be the same, but it's a block so all the weight of the 20 lbs are / Most of the weight is going down on the triangle.
059 I; And for man B?
060 E6: For man B most of the weight is going down on him / Or his hands whatever...And I'm pretty confident on that one.

E3 was aware of her conflicting statements (for the target situation and the bridging example), and that may be the cause of her cautious the-fulcrum-might-help guess, in contrast with the notable increase in confidence evident in the explanations of E2 and E6, who were both "guessing" about the anchor and bridge examples.

In experiment 1 the same intuitive beliefs in the extreme case examples were observed. However, that sequence proceeded without further attempts to consolidate the model or to separate the non-generalizable conception into the principle of levers.

b. Model Construction via Extreme Cases. The additional extreme case comparisons in lesson 2 should provide more evidence to speak to the last two issues above, i.e. the construction or consolidation of models, and the separation of variables; as well as the possibility of conceptual change brought about by the extreme case examples.

I have defined extreme cases in Chapter II as situations where one variable is taken to a limit. In the cases used here I took care that the principle of levers would still describe the behavior of the objects. So, for example, the load leverarm is taken to be almost zero in the extreme case B in Figure 4.8, but the principle of levers could still be applied to the problem.

Little has been written about the use of extreme cases in teaching and learning, but there are at least two studies concerned with experts' reasoning that address the issue. Nersessian (1989 and 1990) analyzed the original works of physicists such as Faraday, Galileo and Maxwell. She studied periods of transition or conceptual change in the disciplines, and looked specifically at the reasoning processes, that is strategies and procedural knowledge, that scientists used during such periods. She identified a set of abstract procedures, such as the use of abstractive representations, analogies, limiting (extreme) cases and visual representations. For example, Galileo used extreme cases to reason about falling bodies by abstracting certain physical dimensions from the situations. He considered motion through increasingly less dense media until the medium was abstracted away completely. Nersessian thought that the idealized representations of such extreme case analyses often facilitated Galileo's recognition of analogies between different phenomena, or that the "idealized representations form abstract schemata common to different problems" (Nersessian, 1989:175).

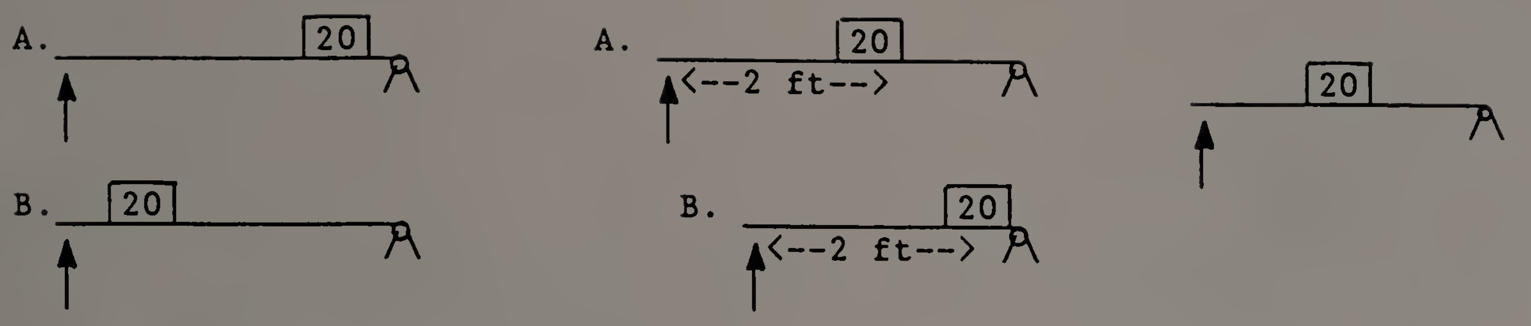
Clement (1981; in press) wrote about the role and nature of extreme cases in his work on the creative problem solving techniques of experts. He suggested that experts seem to use extreme cases to generate examples in which physical intuitions can be applied with high confidence. I interpret this to mean that, in teaching, extreme cases may be presented to students in the hope that they would "trigger" a physical intuition, as well as facilitate the use of the intuition to explain the behavior of novel phenomena. An extreme case example may initiate the use of new knowledge schemata, and allow students to think about previously misunderstood phenomena in a different way - perhaps similar to what Nersessian described as allowing the recognition of analogies between different phenomena. I infer from Clement's suggestions that the intuitive schemata initiated by extreme cases may allow the students to make inferences about situations such as the target questions that may differ from their previous beliefs - thus producing a potential conflict generation characteristic, that may also facilitate conceptual change. In addition, the intuitions triggered by the extreme cases may allow students to make comparisons between lever situations and to construct new functional relationships between variables, for example: "a smaller effort is required as the load leverarm decreases", thus introducing the load leverarm as a variable.

There is of course the obvious difference between the two expert studies cited above and the levers study. Experts generate extreme cases, in thought experiments or whilst experimenting with actual phenomena (Nersessian, 1990). In the levers lesson the extreme cases are presented to the children, and one can only propose outcomes similar to some of those suggested above. There is enough evidence to suggest that the extreme cases used in the second levers lesson were indeed powerful in triggering apparently intuitive beliefs, and it may be useful to view the students' statements with the findings and suggestions from the expert studies in mind.

I shall discuss first the role of the extreme case examples in the students' construction of the fulcrum-helps model; second, evidence for conceptual change brought about by the examples; and third, the role of the examples in the separation of a load-effort distance variable into two more appropriate variables, namely the load-fulcrum and effort-fulcrum leverarms.

Extreme Cases and a Fulcrum-Helps Model. The students had to consider two comparison situations; a load leverarm change in the first and an effort leverarm change in the second.

The comparison situations are given again in Figure 4.8 for reference.



Comparison 1 Comparison 2 Bridging Example

Figure 4.8

Extreme Cases Revisited

The conceptions used by the four students in their explanations for the revisiting of the extreme cases is presented in Table 4.14.

Table 4.14

Successful Students' Conceptions: Extreme Cases Revisited

Student	Extreme Cases 1 (change d_{Lf})	Extreme Cases 2 (change d_{Ef})	Bridge
E2	+ model; non-gen	+ model; non-gen	+ model
E3	+ [model; non-gen and principle 1]	+	+ model
E5	+ model; non-gen	+ model;principle 1	+ model
E6	− non-gen	+ model;principle 1	+ model

I noted during E2's interview, the first of the four discussed here, that there was a sense of "real" learning occuring, such as I had rarely experienced as a teacher before. That impression recurred in the other three interviews and was reinforced by the reviewing of the videotapes and the reading and analyses of the protocols. I was impressed by the ability of these children to develop arguments, first tentatively but with more confidence as their explanations appeared to make more sense to them. The growing confidence is reflected in their responses to the confidence scales, but perhaps more evident in the increasing willingness to verbalize their thoughts. These are subjective impressions, but one is left with the ultimate "why" question: their pretest and bridging sequence explanations were no better or worse than those of the other eight students interviewed. Could one therefore attribute such metamorphoses to the changes in instruction and in particular to this short sequence of extreme cases?

The ideas emerging about the fulcrum's action were consolidated and generally accompanied with high confidence in the answers to the revisited question. Three students changed their inert-fulcrum-does-not-exert-a-force misconceptions and acknowledged that they were aware of the changes; and all four proceeded to answer the near transfer questions with increasing confidence.

I am proposing that the model construction in all four students depended on their intuitions about the extreme case situations. I also infer from the protocols that the extreme examples triggered or initiated the primitive, intuitive schemata that convey the essence of the fulcrum-helps model. The protocol analyses suggest to me that there are two strong intuitions uncovered in the students: one seems to be triggered by case B of comparison 1 and the other by both case A of comparison 1 and case B in comparison 2. The following excerpts illustrate these ideas.

Comparison question 1 is explained as follows:

082 E2: This man is gonna have to do 20 lbs in case B. 'Cause it's [load] right on him. And this man [A] will probably have to do 5 to 1 lb, because it is so far away from him and the block is holding up most of the weight...

098 E3: I think here [B] the person would have to push 20 lbs...

100 Because it's [load] so close to him, that's where he has to hold it.
102 And here [A] / I think the person would have to push less, so I'll say 19 lbs.
103 Just a guess. And ah / Because again, it's [load] closer to the thing.
104 I: Closer to?
105 E3: Closer to the turning point.

200 E5: Ohhh! Well, can I change my mind about what I said before? Like if I answered this person [B] can I answer this person [A]?
202 Well, I think it'll be easier for *him* [A]...
204 Because / The more the weight is on the turning point, and he is holding less of it.
205 I: He is holding less of it?
206 E5: Yeah, he has to push up less.
208 But this guy [B] is holding the whole 20 lbs.
215 I: Uhm, and how do you feel about this?
216 E5: Ah, well more confident.

143 E6: OK. // B [wrote 20 lbs under B].
144 I: You say B about 20 lbs?
145 E6: B about 20 lbs and uh // A is about / Uh [12s pause]. Can I say anywhere from 5 to 10 lbs?
153 And I'm confident about this [case B and the load-effort distance], but I'm not confident about this [quantitative answer to case A].

I shall try to specify two separate physical intuitions in the excerpts above. The intuition that was mentioned first by E2, E3 and E6 is the "he will have to hold the whole 20 lbs" idea. One can expect that this would be an intuition informed by the children's experiences, and it interesting that the intuition is accompanied by the weight distribution non-generalizable conception described before. In other words, students are considering the effect of the weight of the load on the person's force, and the excerpts from E2 and E3's protocols are good examples of the ideas about the distributive property of the load-effort distance.

I am not sure that the second or "fulcrum holds more" intuition should be seen as distinct from the first. It is difficult to suggest what the students may be thinking, particularly since all four declared a few minutes earlier that the "thing" could not "hold any of the weight" in the bridging example. Why are they suddenly able to attribute such an action to the same object? One may suggest that the same kind of analogical extension hoped for in the anchor-bridge examples earlier, is at work here. Thus, the students may "see" the fulcrum in case A as the equivalent of the person in case B. The fulcrum is therefore holding "most of the weight" in case A. It may be that the two examples afford them the chance to do what Clement (1988:571) called "establish confidence in the validity of the analogy relation".

An argument may proceed as follows (and here I am making a liberal number of inferences): if the man in case B has to push up 20 lbs and one *knows* that, and one also knows that the person in case A will have a much easier task, then the only explanation seems to be that the person and the fulcrum are in similar situations, thus "holding the whole 20 lbs". I have no verbatim protocol evidence for this suggestion and one can also defend the first alternative; i.e. that two, distinct physical intuitions were initiated by the examples.

Comparison question 2 in Figure 4.8 was designed to initiate conflict between the students' non-generalizable conception and the obviously crucial load leverarm differences in cases A and B. In addition, the fulcrum helps idea should be reinforced again. The question is difficult, since both leverarms are changed. The expectation was that the children would rely on the fulcrum-helps model particularly to explain case B - and this did happen. So, should one accept the suggestion above that students may hold a second physical intuition about the fulcrum's action, then that intuition should be initiated here:

091 E2: [Following a discussion about the earlier definition of the effort leverarm.]
Oh no. It is different, it is different [the load leverarm in A and B], but the 2 ft distance from the load to the effort is the same amount, but the effort turning arm is different. So I'd have to say that the weight would be distributed more, in case A / It would be distributed more towards the man. It looks like he would be holding up around 12 to 13 lbs.
And the man / In case B, would make / It would be a lot easier for him, because it looks like the block is holding up about 15 to 17 or 18 lbs already for him.
Does it make a difference what the distance from the man to the load is? No, it matters where the *turning point* is [her emphasis].

111 E3: I think that this man [B] will [find it easier], because again he's / The weight is closer to the // To the turning point and therefore that might keep it up a little. Whereas this person [A] is // The distance here is so far away.
112 I: Distance from the load to the turning point you're pointing to?
113 E3: Right, from the / Ah, yeah.
114 I: And what difference does that make?
115 E3: Well, I guess if ah / If it's [load] closer to the turning point then that might / Might help keep it up...
117 ...I'm not at all sure about that.

235 E5: Ah, well, person A is // Is really far here?
236 I: Really far from?
237 E5: That's more than 2 ft [the load leverarm]?
238 I: Yeah, OK.
[A discussion of the load-effort separation, adding nothing to her explanation.]
244 E5: And on this one [B] the weight is almost on the turning point, so they don't have to / Hold very much of it?
245 I: OK, and how come they don't have to hold very much here?
246 E5: Because the turning point is holding it up.

163 E6: B. And [10s pause]. The reason for B is that it's // The // Turning point / The weight is closer to the turning point. So. The weight is closer to the turning point so. The load's closer to the turning point so the turning point supports it more.

It is still difficult to decide whether a separate "fulcrum helps" intuition was initiated by the examples: E3, E5 and E6 apparently reasoned from the "turning point help" view, while S2 first considered the effects of a longer load leverarm on the person. I think, however, that since all four students accompanied their "turning point helps" view with an explanation of the effect of the varied leverarms on the person, that there may just be one intuition, i.e. "if the load is almost on me, I'll hold most of it". This intuition is extended analogically to the fulcrum in the extreme case situation B.

However, I have no evidence to suggest why the same extension was so difficult in the anchor-bridge-share-10-lbs sequence; except that the extreme case situation may be more compelling - one certainly knows the difference when holding "most of a load" versus sharing a load with another person. Deciding which example or comparison of examples triggered an intuition could be important; the estimation problems, such as the bridging example, are all single example problems and these may be less successful since the comparison examples seem to provide two "check points" for an intuitive idea.

I infer from these statements that, at this stage of the lesson, all four students viewed the fulcrum as "helping" or "supporting", thus by inference, exerting a force; and that the load leverarm determined for them the amount of force to be exerted by the fulcrum (and the person).

c. <u>An Intuitively Anchored Model</u>? The excerpts support the proposal that a more general "force diagram" model was constructed: there is evidence in all the statements that the students were considering three "actions" (of the person, the load and the fulcrum), taken to be the naive equivalent of forces here. For example, E5 said that person A in question 2 would *push* more than B, since the weight (inferring that the weight is pushing down) is closer to him (line 204 and 206); while the turning point is *holding* more of the weight when the load is closer to it. The model (of the turning point's action) for all four students appears to be grounded in the physical intuitions of "pushing" and "holding" described above, and one can therefore conclude that the naive qualitative models of forces constructed by these four students were intuitively anchored.

There is limited support for Nersessian's notion that abstract schemata that are common to different problems, could be formed by the representations facilitated by the extreme cases. At this stage of the lesson, the students had not been exposed to many different situations, and this idea should probably be evaluated in the far transfer analyses. However, they were able to reason about a difficult problem such as case A in comparison question 2 in Figure 4.8 by applying their models.

d. <u>A Causal Mechanism</u>? I suggest that the students' qualitative models have a definite causal and explanatory element essential to students's ability to make inferences about new situations and therefore essential to a deeper understanding of the levers phenomena. This is the "fulcrum helps" idea which contrasts with the following idea in the form of a rule: A student may say, as E6 did (in line 163): "...man B will [push up less] since the load is closer to the turning point..." and leave it at that. This statement has a rule-like quality, and I am using the word rule here to mean an "if .. then" statement about observable features. However, E6 added "...so the turning point supports it [load] more.", thus giving what I construe to be a causal explanation involving a hidden causal mechanism for his prediction that person B will "push up less".

Forbus and D. Gentner (1990) described "explicit mechanisms" in qualitative physics that use processes as "agencies of causation". For example, in the statement

> "opening the throttle increases the flow rate of gas to the engine, which causes the engine to work faster" ;

"increase the flow rate of gas" is a process which allows one to explain the causal inference - the car goes faster (Forbus and D. Gentner, 1990: 673). In their sentence the rule statement would be "opening the throttle makes the car go faster", a rule that any mechanically naive driver would be familiar with but for which they would probably have no explanation or understanding.

Similarly, the "fulcrum that pushes up" or supports or holds some of the load's weight, allows one to make causal inferences about the other force in play in the lever, namely that of the person. I propose that such causal inferences depend on the fulcrum-helps mechanism, and will continue to refer to the qualitative force diagram model as the fulcrum-helps model since this specifies the model's most important element.

2. Conceptual Change Facilitated by Extreme Cases

Two misconceptions were identified in the pretest, target questions and the bridging example: the control misconception, including the "new" version - that the load will "push the board down more toward the person" if there was only one person and a fulcrum "holding" it; and the inert forces do not push up misconception, that may be influencing the thinking displayed in the version of the control misconception stated above.

a. Changing the Control Misconception. Changes in the control misconceptions of E3 and E5 were already obvious from their responses to the extreme case comparison questions above. E5 explicitly referred to this change of mind when she asked whether she could change her answers: "...can I change my mind about what I said before?" (line 200). She continued and explained that her answers are different because "...the more the weight is on the turning point, and he's holding less of it..." (line 204). This is in contrast to her answer to the target question following the bridging sequence that the person would find it easier "...because the board is shorter, so there's less he has to keep level..." (line 056). E5 was a cautious speaker and she evaluated her confidence level responses carefully, so that her indication of increased confidence in her answer in line 200 (from "guessing" for the last target question) is encouraging.

E3, an even more careful speaker, said that the shorter board would be easier to hold since "...the weight is I guess closer to you, maybe it's easier to hold..." (line 003), but she was guessing because "...there's no other way I can really think of it..." (line 005). Still guessing, and using words like "might" and "could", she thought in comparison question 1 that person B, closer to the load, would have to push more than person A, since the load "...is closer to the turning point..." (line 106) and thus, in comparison question 2, "...if it's closer to the turning point then that might // That might help keep it up..." (line 115).

b. <u>Changing the Fulcrum-Does-Not-Push Misconception</u>. E3 also held the inert-objects-do-not-push misconception in the bridging example: "...the person has to hold...the whole 20 lbs up, because if it's a hinge then it's not really holding it up..." (line 020). By the time we revisited the bridging example, she thought, obviously amused, that "...I guess it wouldn't be the full 20 lbs, 'cause it's closer to the turning point. // So I'll say 19 and a half..." (line 124). I was poised to present her here with a contradiction in her answers: she had suggested that man A in comparison question 1 would push up 19 lbs, and I was waiting for her to repeat her "full 20 lbs" or; alternatively, how she would reconcile a correct "10 lbs push" answer with the "19 lbs push" answer. It appears that she was also aware of the contradiction and slipped out of it with the 19 and a half estimation. Nevertheless, I consider her answer a change from the idea that the fulcrum could exert no force, even though the quantitative estimate was so far off the mark for the bridging example.

The explanations given by E2 for both the extreme case comparison questions indicate that she has changed her view about the bridging example, from her original position that "...the hinge will not hold up anything / It will just let it drop, so he's going to have to hold up all of the 20 lbs..." (line 023). Later she argued consistently and eloquently that the fulcrum does have a role, e.g. "...the block is holding up most of the weight..." (line 084); "...it would be a lot easier for him, because it looks like the block is holding up about 15 or 17 or 18 or 19 lbs already for him..." (line 091). For the bridging example revisited "...10 lbs, he'll have to push up 10 lbs..." (line 104), feeling "Fairly sure I'm confident." (line 106). E2 was not sure about the anchor and bridging examples "...still questioning it..." she said in line 020, but by the time we revisited the extreme case examples her confidence in the answers was consistently high.

E6 believed so strongly in the inert objects do not push idea, that he changed a correct answer for the anchoring example and admitted that "...I guess I'm just guessing about that..." (line 041). He used the fulcrum pushes up idea in both the revisited extreme case comparisons, but in the revisited bridging example reverted back to his misconception answer: "...you have to push with, hmmm, 20 lbs..." (line 188). We referred back to his estimates for the extreme cases; he thought almost a minute about this and changed his answers for case A in the extreme case comparison 1 and the bridging example: "This one [A] will have to push up with 5, with 5 or lower....and this one [load in center in bridging example] would have to push up with 10..." (line 196). He still was "...not very confident..." about the answer, but "...not quite guessing..." (lines 198 and 200).

c. Summary. One may conclude that the extreme case examples were important in most of the instances of conceptual change described above. Again, it seems that the physical intuitions triggered by these examples furnished the children with a different, more sense-making schema that replaced the old conception.

3. Separation of Variables

The load leverarm variable is appearing for the first time in the extreme case revisited explanations in the students' protocols, particularly in the examples where it is most obviously relevant. However, the question designed to initiate conflict between the non-generalizable conception and a possibly emerging principle of levers: "Does it matter what the distance from the man to the load is?" in the extreme case comparison questions, did not serve its purpose. The anticipated conflict situation was the following: given that students used the non-generalizable conception in question 1, the magnitude of the person's force in question 2, which obviously does not depend on the effort-load separation, was to cause the conflict.

All the students were aware of their different responses to this question, but were also able to explain why they answered thus, as E2 pointed out in response to comparison question 2's "does it matter what the distance ..is?" question:

091 E2: No, it matters where the *turning point* is [her emphasis].
092 I: OK, let's go back to the question here where you said it matters what the distance from the man to the load is?
093 E2: Hmm / I guess it does matter, but it matters whether / Uhm, I guess it does matter what the distance from the man to the load is, but in *this* particular one it wouldn't matter because they [effort-load separations] are the same and when it is the same it has no relevance.
094 When it's different, it does, it makes the difference.

[A discussion on whether she should write these reasons down follows.]

100 I: So let's just see, we said what is important, is the distance from {the turning point...
101 E2: To / The turning point and the distance from the man to the load. But when the load / The distance from the man to the load is the same it's not a relevant fact.
102 I: Then you're looking at?
103 E2: Then you're looking at the distance from the load to the turning point.

Other students gave similar responses, suggesting that for them, these examples were quite different phenomena, thus requiring different explanations. Their separate ideas for the situations may be an illustration of diSessa's "knowledge-in-pieces" notion (diSessa, 1985).

I saw the question as useful though, since it gave the students another opportunity to verbalize their thoughts and clarify their thinking. I also used the probes (in lines 100 and 102 above) in all the other interviews, thus getting additional responses about the load leverarm's importance (and of course about the non-generalizable conception).

I did not expect the students to reconcile the non-generalizable conception with the accepted view at this stage; and the more immediate goal, to separate the load leverarm from that conception, was attained for these four students.

4. Near Transfer

The near transfer questions are the same as those used in lesson 1 (description in Appendix E). The lesson 1 participants did very well on the wheelbarrows question that depicts situations visually and conceptually near to the extreme cases comparison 1 in Figure 4.8. However, only one person answered the second, more difficult bottle openers question correctly with an appropriate explanation.

The more successful near transfer ability evident in the protocols of the students in experiment 2 is encouraging (shown in Table 4.15). Again, the abbreviation "princ 1" refers to the load leverarm component of the principle of levers, and "princ 2" to the effort leverarm.

Table 4.15

Successful Students' Conceptions: Class II Near Transfer Questions

Student	Wheelbarrows	Bottle Openers
E2	+ non-gen	- m.c$_c$.
E3	+ princ 1; model	+ more leverage
E5	+ non-gen; model	+ principle 2
E6	+ princ 1; model	+ principle 2

Three students did very well on the bottle openers question, particularly E5 and E6. E5 held a control misconception at start of the lesson, and it is interesting to see that the temptation of control in this question, that lured nine other students in the study, was ignored.

All the students (except E2 in the bottle openers question) identified the loads, efforts, leverarms and fulcrums correctly. I was amazed at the ability of the three successful bottle opener candidates to identify the fulcrum - all three admitted that it was difficult to find, but their simulations of the motion of the opener probably helped them to determine the point.

5. Summary: Reasoning from Extreme Cases

The analyses presented above show that the physical intuitions initiated by the extreme case examples were instructionally effective in at least two important ways: one, in providing the anchoring conceptions for the construction of an apparently more general, less naive force diagram model of the levers; and two, in changing misconceptions, presumably because the less naive model enabled them to view the examples differently.

The goals of the "revisiting sequence" were also attained for the four students under discussion, in that the fleeting "fulcrum-helps" ideas initiated in the bridging sequence were consolidated; and in the apparent establishment of the desired qualitative model. There is also evidence that the load leverarm was differentiated from the non-generalizable conception.

Although it is difficult to determine the reason why the extreme case examples initiated intuitions apparently more powerful than those triggered by the anchoring-bridging sequence, these positive findings indicate that such examples can indeed be useful instructional tools.

6. Generating a Principle from a Model

The evolution of the principle of levers in the rest of the teaching sequence, as interpreted from the protocols of the four successful students, will be discussed in this section.

A major goal common to both the short class II to class III transformation sequence and the benchmarks sequence for class I levers was to complete the separation of the children's non-generalizable load-effort distance variable into load and effort leverarm variables, and it seems therefore sensible to discuss the teaching sequences under one heading.

It was anticipated that the fulcrum-helps model could support the students in their reasoning about the class III lever question, since the non-generalizable conception for this lever class presents a unique problem mentioned before. The non-generalizable conception is essentially the same as the control misconception for classes II and I levers, as is illustrated in Figure 4.9.

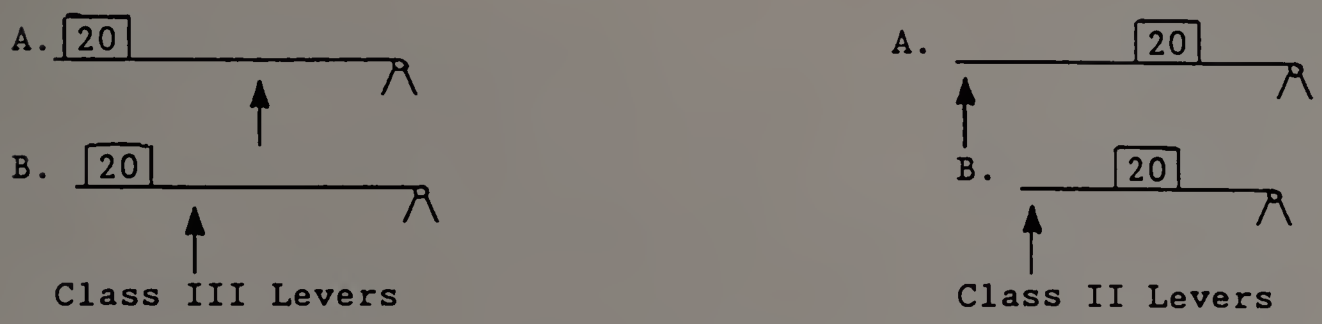


Figure 4.9

Misconception or Not: Class II and Class III Levers

Since the effort leverarm increases as the effort moves nearer to the load in the Class III lever comparison, the notion that "the nearer the load is moved to the effort, the less effort is required" gives a correct prediction and should be regarded as an acceptable explanation. The same idea applied to the class II lever comparison, decreasing the load-effort separation, implies a shorter effort leverarm, hence a misconception.

However, if one could have the students focus on the fulcrum in the two different classes, they may accept the more general qualitative idea that "the further the effort is from the fulcrum in both classes, the less effort would be required". The fulcrum-helps model may enable them to do this, since the causal explanatory part of the qualitative model is concerned with the fulcrum's function. Experiment 1 findings, although limited, indicate that students are able to distinguish the different "actions" of the fulcrums in the two classes of levers: in class II it "helps" and in class III it "hinders" or "does not help."

The last aspect of the class III transformation sequence will be discussed first, to be followed by a consideration of the success in the separation of variables in both the class III and Class I teaching sequences.

a. <u>The Fulcrum Does Not Help</u>. The situations in the transformation sequence are given in Figure 4.10 for reference. The comparison situations are all extreme case examples, and the expectation was that students' physical intuitions triggered by these examples as well as their fulcrum-helps model may make their predictions easier to explain.

The children had to compare cases A and C to start with. The assumption was that they would be able to reason about case A from the fulcrum-helps model, and that the effort leverarm may emerge as a plausible variable. The comparison between the two class III examples (B and C) should reinforce the importance of the effort-fulcrum distance (presumably students already have a notion of the load leverarm importance from the class II teaching section).

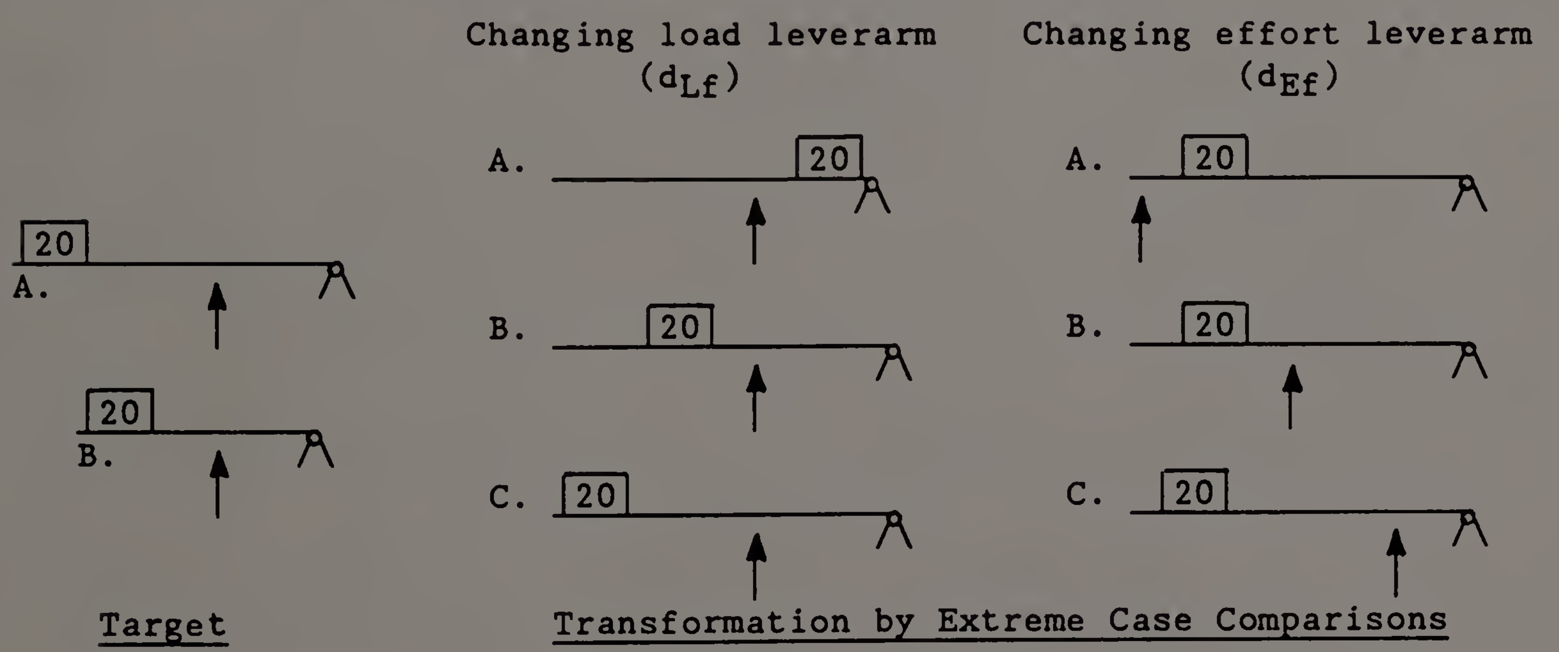


Figure 4.10

Transformation Sequence: Class II to Class III Levers

A summary of the conceptions used by the four students is presented in Table 4.16. The "non-gen" abbreviation is here referring to the class III non-generalizable conception; and "fh+" indicates "fulcrum-helps" for the class II levers and "fh-" that the student noticed the position of the fulcrum in the class III lever.

Table 4.16

Successful Students' Conceptions:
Transformation Sequence

Student	Target	Transformation 1 (change d_{Lf})	Transformation 2 (change d_{Ef})	Target
E2	+ non-gen; fh-	+ fh+; non-gen; fh-	+ fh+; fh-	+ non-gen; fh-
E3	+ non-gen; fh-	+ princ 1; non-gen; fh- ; princ 2	+ fh+; fh- non-gen	+ fh-; non-gen
E5	+ guess; fh-	+ fh+; non-gen; fh-	+ fh+; fh-; non-gen	+ non-gen
E6	+ fh-; non-gen	+ fh+; fh-; princ 1	+ non-gen	+ non-gen

The summary shows that all the students noticed the difference in the fulcrum's action for the different lever types at some stage during the period. As expected, students relied on the non-generalizable conception to make their predictions, but it is interesting that in all but four explanations the conception was preceded by an explanation of the fulcrum's role in the situation. It seems that the students did reason from their fulcrum-helps model to explain why a smaller load-effort separation gave the correct prediction. This was one aim of this instructional sequence, and supports the suggestion that the qualitative model should afford the students a deeper understanding and the ability to transfer new ideas grounded in this model.

Protocol Evidence. Student E3 was probably the most successful of the students. Her protocol indicates the emergence of the intended qualitative principle of levers, and as the excerpts show, this all results from her reasoning about the fulcrum:

Transformation 1:

195 E3: I think the A person [will find it easier].
196 I: Person A?
197 E3: Yeah. Well, it will be easier for that person, because the load is closer //
198 I: Is closer?
199 E3: To the turning point, so it would be easier to hold it up like I wrote before [referred to first written statement].
200 I: Uhhm?
201 E3: And ahh / And the effort turning arm is pretty long and for that I'm fairly confident.
202 I: OK, and let's then look between man B and C?
203 E3: I think it would be easier for B to keep it level, because it's [load] close to the person / And also closer to the turning point.

Transformation 2:

207 E3: I think man A will, because then he still has the turning point there, so and that's also helping him support the load. As with person C, there's the whole thing // The whole [indicated the length of the board]...
208 I: Length?
209 E3: Length and the weight to hold without any help from the turning point.
216 I: OK, and then again, just man B and C?
217 E3: I think again that it'd be easier for person B because again they're closer to the thing [indicated load] and it might be easier to hold.

Last target question:

222 E3: [B.] And again because the load is closer to the person and the turning point is behind, so that doesn't help it.
223 I: Doesn't help ah / What?
224 E3: That doesn't help support the weight.

E5 refrained from mentioning the load-effort distance, but was very clear about the fulcrum's different role in situations A, B and C:

233 E5: Because the effort is between the load and the turning point [class III]. And the load is distributed between the effort and the turning point [class II]. So I don't know.
234: I: Which of those two people would you prefer to be?

235 E5: Perso-o-on B, I think // I don't know why though // Oh, I don't know!

She was confident about her predictions between the class II and III situations. In class II levers "...they're not holding the whole thing, because the turning point is holding some of it..." (line 252); and "...because he is sharing it with the turning point..." (line 272); while for the person in the class III lever "...the turning point isn't really holding the weight the man is..." (line 254). The explanation for the difference between the class III situations was difficult for her, until she used an analogy and seemed satisfied with this idea:

[Responding to a question about her class III uncertainty.]

276 E5: Well, I don't know, 'cause / I don't know. 'Cause yesterday I figured it out for these ones [class II] but I didn't figure it out with this kind?
277 I: OK, good answer! And between B and C, can you say?
278 E5: B // I think.
279 I: B?
280 E5: *Yeah* [her emphasis]
281 I: You seem pretty sure about that one?
282 E5: I just decided [laughs].
283 I: Yeah?
284 E5: Because he // It's like / Oh, I know, I remember how I thought about it.
285 I: Uhhm?
286 E5: It's like / You know when you're holding a hammer, you hold it near the weight //
287 I: Yeah?
288 E5: It// It feels heavy if you hold it way back? So he's holding it near the weight.

The protocols of the other two students are much like the above. The fulcrum-helps model was used: "...[the turning point] is also holding up some of the weight..." (E2, line 137); "...the load is basically right next to the turning point...it puts it's weight right on the turning point..." (E6, line 274). And the fulcrum's changing role was noted: "...[the person] is the only part holding up the 20 lbs..."

(E2, line 136); "...the load and the person switched places. So that changes that [referred to his written statement]..." (E6, line 249).

Summary. I infer from the evidence presented above that the fulcrum-helps model was useful in an explanatory way to the students. The three comparisons in each problem served their purpose: students were certain about the class II situation, reasoned from this case and noticed the transformations intended, i.e the fragmented aspect of lesson 1 was effectively removed.

The analogy used by E5 to convince herself of the class III non-generalizable conception's acceptability was a singular instance, and although it was not expected it served its purpose.

b. An Emerging Qualitative Principle of Levers. The main goal in the class I teaching sequence was to separate the non-generalizable conceptions into the effort and load leverarms. This separation was initiated by the fulcrum-helps model for class II levers since the load leverarm became important. One did not expect any further development in this variable separation for class III levers, but the summaries of conceptions used indicated some instances where the leverarms were mentioned as important.

The Principle in Class III Levers. The principle of levers appeared on three occasions in Table 4.16 above: twice in the excerpts quoted from E3's transcript and once in E6's protocol. I did not expect students to make any progress in the variable separation issue in the transformation; the best one anticipated was an explanation for the class III non-generalizable conception. It may, however, be interesting to look at the class III near transfer questions with the variable separation in mind. Table 4.17 contains the summary of conceptions used in explanations for the transfer questions. The abbreviations are the same as those used in Table 4.16.

Table 4.17

Successful Students' Conceptions: Class III Near Transfer Questions

Student	Near Transfer 1 Steam Shovels (change d_{Lf})	Near Transfer 2 Brooms (change d_{Ef})
E2	+ fh-; non-gen	+ fh-; non-gen
E3	+ princ 1	+ princ 2
E5	+ non-gen; princ 1	+ intuition; non-gen
E6	+ princ 1	+ princ 2

The analyses of the near transfer explanations do not suggest any marked increase in the use of the load or effort leverarm variables, as opposed to the non-generalizable conception. However, it is again encouraging that students were able to make the correct predictions and use reasonable explanations throughout. One did not expect misconceptions in this class of levers, but other factors such as the increasing ease with which the children identified the lever elements (correctly), including the "floating" fulcrum in the brooms question, suggest that students were successfully transferring their new ideas to the simple machines.

The Principle in Class I Levers. The benchmark sequence's situations are given again in Figure 4.11 for reference. The students were asked to estimate the amount the person would have to push to keep the board with the 20 lb load on level.

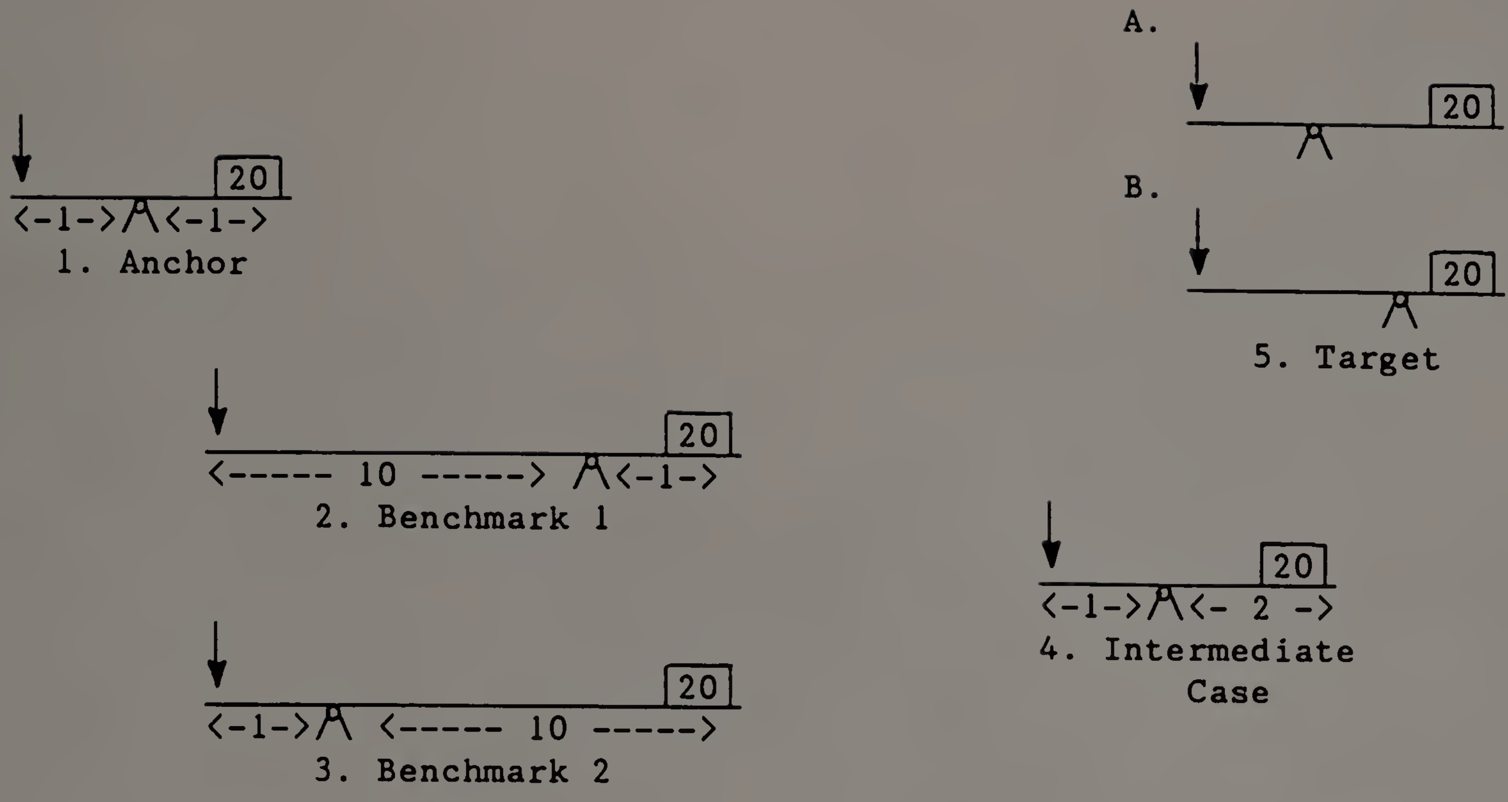


Figure 4.11

Benchmark Sequence for Class I Levers

The conceptions used in the students explanations for the benchmark sequence problems are summarized in Table 4.18.

Table 4.18

Successful Students' Conceptions: Benchmark Sequence

Student	Target	Anchor	Benchmarks 1	Benchmarks 2	Intermediate Case	Target
E2	+ princl;fh+	+ $L_f = E_f$	+ $E_f>>L_f$ principle	+ $L_f>>E_f$	+ [principle]	+
E3	+ principle	+ $L_f = E_f$	+ [used principle throughout] (analogies)	+	+	+
E5	+ principle	+ $L_f - E_f$ analogy	+ used analogy to confirm principle	+	+ [principle]	+
E6	+ principle	+ $L_f = E_f$	+ [used principle of levers through out]	+	+	+

The summary indicates that the aim of this sequence was attained: all the students eventually used only the qualitative principle of levers, and the non-generalizable conceptions seem to have disappeared. This is an important change from the experiment 1 results, where only two students were using the principle consistently at this stage in the lesson.

I had hoped that the central position of the fulcrum which separated the load and effort leverarms more distinctly, might encourage students to reason about the magnitudes of both, as these four students have done.

Although the benchmark situations triggered the intuitions expected, e.g. that a very large effort would be required in benchmark 2, one should probably refer to these examples as extreme cases rather than benchmarks. A benchmark is associated with a specific quantity; thus if one knew intuitively that a force of 200 lbs was required in benchmark 2, whilst only 2 lbs is needed in 1, these examples could be called benchmarks. The findings above suggest that the quantitative aspect was not needed, since the students correctly identified the direction of change in the variable relationships, and constructed the principle of levers as intended.

Analogies were used more frequently in the students' reasoning, perhaps because class I levers are more frequently present in a child's life. E5 struggled with the seesaw until her explanation made sense and agreed with her apparently intuitive prediction. She was quick to suggest that benchmark 1 would be "easier" than the anchor example, but had a hard time aligning her analogy to the answer; until she had an insight, described as "...I am trying to think of a seesaw..." (line 434) and gave a principle of levers explanation. A scale and a seesaw were E3's analogies to the anchor example:

275 E3: The load turning arm and the effort turning arm is the same distance...it looks to me like a seesaw, and if there was a person there [load], they'd have to weigh the same, it's like a scale. So, he'd also have to push with 20 lbs.

The students' explanations for the near transfer questions should also indicate this trend towards the use of the qualitative principle. The summary of conceptions is given in Table 4.19.

Table 4.19

Successful Students' Conceptions: Class I Near Transfer Questions

Student	Near Transfer 1 Pliers (change d_{Lf})	Near Transfer 2 Crowbars (change d_{Ef})
E2	+ principle	+ principle
E3	+ principle	+ principle
E5	+ principle	+ principle
E6	+ principle	+ principle

In addition to the understanding suggested by the correct predictions and normative explanations, there are explanations indicating a deeper understanding, for example: E3 explained that the pliers were different, but somehow the same as the simple levers.

337 E3: ...it is different //...In a lot of ways they are the same too, except you can push //
338 I: In what way are they the same?
339 E3: ...In half of this [pliers] it is the same, because if you lift up then it looks exactly like the levers...And it's different in the sense that you're pushing from both sides, except just up or [shrugged].

E5 found the pliers questions easy "...because I have used pliers to cut a wire before..." (line 565), and then added the surface feature reason used by most students in the pilot studies, a reason why I hesitated about adding this question:

567 E5: I think it's [B] because they're / It's closer to the turning point and also /
568 I: What is closer to the turning point?
569 E5: The load.
570 I: The load?
571 E5: Yeah, and also, if they are the same kind as the pliers I have, then the inside closes before the outside?

The fact that she gave the principle as her primary reason (in line 567) is at least an indication that she considered the principle (leverarm differences) a more or equally important variable.

7. <u>The Written Statements</u>

The effects expected from the opportunities to write statements in experiment 1 failed to materialize. Hence, the quality of the statements of these four students are encouraging: the growth of the principle can almost be summarized by their expressions, and I think the quotes speak for themselves (statement 1, at the end of the class II sequence, is followed by statement 2, at the very end of the lesson).

Student E2

043 E2: Think where the weight is and how the weight is going to be distributed.
260 OK. It says: Look at the length of the turning arms before you make a judgement

Student E3

181 E3: If the load is closer to the turning point then the load is supported by the turning point, then it would be easier. But then also that the effort turning arm is longer.
344 [Read statement 1.]
No, I think that I want to make the first thing I said clearer. So that, I agree that // I still agree that this, the effort turning arm is longer then it's easier and I think then also / It's about what I said there, but if the load turning arm is shorter, then it's easier.

Student E5

223 E5: OK. If the load is over or closer to the turning point, it, meaning the turning point, holds more of the weight.
590 [After reading statement 1.]
Yeah, but I want to add something.
[Writes, talked about abbreviations.]
594 OK. If the load is closer, and the effort is farther away from the turning point, it is easier.

Student E6

236 E6: The closer the load is to the turning point and the farther it is to the effort [the easier].

412 The longer the effort turning arm and the shorter the load turning arm, the easier.

8. Far Transfer

Findings from the analyses of the four successful students' explanations for the far transfer posttest questions are presented as a conclusion to this formative evaluation of the positive outcomes of lesson 2. I shall again accept A. Brown's (1990) finding, i.e. that normative explanations to the four far transfer questions require transfer on the basis of deep, structural principles. It was shown in sections 2 and 3 before that the four students were able to construct such knowledge structures: I inferred from their protocols that the final qualitative principle of levers formulated by the students in their last written statements evolved from a causal, explanatory and intuitively anchored qualitative force diagram model. I have referred to this model as the fulcrum-helps model, since the causal mechanism, i.e. that the fulcrum "helps support, or hold" the load on a lever, seemed to be initiated by the extreme case intuition in which the model is grounded.

The understanding of the four students may be evaluated on several levels: were their predictions correct; were their explanations compatible with a physical theory of levers; and, were there any indications of causal, explanatory models underlying the rule-like principle of levers used in their explanations?

a. Reasoning from a Principle and a Model. The students' explanations for the far transfer questions are summarized in Table 4.20. The far transfer questions are given in the posttest in Appendix C.

Table 4.20

Successful Students' Conceptions: Far Transfer Questions

Student	Revolving Door	Nutcracker	Shadoof	Nail Clippers
E2	+ principle	+ principle	+ principle	- surface feature
E3	+ principle	+ principle	+ principle	- principle
E5	+ principle	+ principle	+ principle	- lever recognition
E6	+ principle	+ principle	+ principle	+ principle

The Nail Clippers Problem. The nail clippers problem was the only stumbling block for the children in this posttest, and three, including E5 who responded to a probe, recognized levers in the clippers, but were unable to tie their arguments together. An expert (a person with a doctorate in theoretical physics), summarized some of the difficulties he experienced in solving the problem:

029 E: I think in an interview situation the thing that's // Could be difficult for someone, I mean I found it difficult myself, was // Unless you have an intuition about these things - perhaps you're not using a reasoning

principle at all or a physics principle or an...informal principle.

030 But *if* you're using some kind of a principle, the analysis is multi-stepped, because what you're applying the force to is quite a different object than the part of the nail clipper that actually cuts the nail.

031 So there's two levels at which you must keep track of things. And unless a person is willing to record their thinking, I think by writing as they go along, that they're apt to be confused - that they won't be able to hold the information in memory long enough to put it all together.

His suggestion in lines 030 and 031 is perhaps the best explanation for the students' difficulties - writing and "keeping" track of one's thoughts in physics is a sophisticated skill, one that these students most probably could not have developed yet. It is therefore surprising that they are able to do as well as is suggested by their arguments.

The most encouraging finding, in support of the suggestion that the fulcrum-helps mechanism in the model is essential to understanding, is illustrated by two excerpts. The students are not "jumping" into the argument with the principle, but are analyzing the instrument and using the model, before the respective load and effort leverarms are considered. This suggests to me that the students have learned more than a rule; they are actually able to reconstruct their rule-like qualitative principle, again making sense of it in a difficult problem context. The quotes are from the students' posttest protocols.

A sketch of a nail clipper is given for reference in Figure 4.12.

Figure 4.12

Nail Clippers Problem

E3 correctly identified the fulcrum in clipper A, and argued that the top lever would push the clipper part down:

026 E3: I think that it'll be easier for A, because then // The turning point is in front of the load [P]...So then the turning point would be right there and then that [indicated Q with pen] would push down the clippers //
027 ...I guess the load would be there then [Q in A], so that would be the load and this [P in A] would be the turning point...
030 And whereas here [B] // If you push here [at Z in B], then it moves it down [15s pause while she looked at drawing].
031 I: What will happen then?
032 E3: Then the turning point is here [X in B, error] and the load here [Y, error]. It's less of a lever, since there's / The turning point is behind it [load] and you push down here using the force, whatever effort //
033 I: Your effort?
034 E3: Your effort and then, ah / You don't have / There's not a turning point in between so that doesn't affect it, so that would be / Would be harder to push it down.

Her argument starts falling apart in line 032, perhaps because of the error in her identifications of the fulcrum and load in B, and it is not clear what she meant by "...less of a lever, since...the turning point is behind it [load]..." (line 032). I wanted to show, however, that it appears that she performed a thorough analysis of the levers, and reasoned from the position of the fulcrum about the effort required.

E5 gave a surface feature explanation concerning the differences in the vertical distances RO and ZO: "...It's harder when it's going from here to here [ZO] than when it is from here to here [RO]..." (line 071). She illustrated with the actual clipper (a model of clipper A), looked puzzled when the experience did not support her statement and when I probed for other possible differences, responded:

089 I: Is that the only difference between those two nail clippers [the vertical distances]?
090 E5: That *I* can see. // Well, no. // This one is over here and this one is over here.
091 I: The what is where?

092 E5: The / this [pointed to the peg connecting the levers].
093 I: The sort of connecting peg?
094 E5: Yeah. // Is that the turning point?
095 I: What do you think?
096 E5: Ah, well let me see. Well, no-o-o. // That's not, I don't think. I think that's what pu-u // Well, yeah. It is. I think it is.
097 I: OK?
098 E5: // And this one [B]. // On this one, the load, because it's pushing down here, so that's the load [correct, Y in B], that's where it's pushing the nail //
099 I: You want to write that in?
[Discussion about labeling follow.]
102 E5: And in this one [A] it's over here // And so / I don't know // Ohhh! The // [6s pause].
103 I: Did you see something else?
104 E5: Well, sort of. I don't know, it's like up here.
105 I: What is up where?
106 E5: The load //
107 I: The load?
108 E5: [Labelled correctly, X in B.] And this one [A] is down here, and it's easier to close something like, if you have tweezers like, from further back?
109 I: Uhhm?
110 E5: Well, I don't know if it is actually. But / I don't know, I can't say that! [Giggles.]
111 I: This is a tough one!
112 E5: Yeah.
113 I: Yeah. // Let's look at clipper B - do you have to push down to make it work?
114 E5: Yeah. // Well, no. Well, I don't know, if it really is there. //
115 I: If you push down, what will happen?
116 E5: No, you have to pull up. Don't you? Do you? Can you tell me?
117 I: Well, let's think about it
118 E5: {You *have* to pull up, because if you push down, this will come up [X in B], and it won't work.
119 I: You did great there!
120 E5: So you have to / It's much easier to squeeze them than to pull it up. Well, for *me* it's easier [laughs].

The argument in line 120 is of course valid; it is difficult to imagine that one may actually be able to cut a nail with an instrument such as B, but since the aim was to investigate students' ability to recognize levers, the problem is acceptable. My probe in line 113, used in all the interviews when students seemed confused, probably diverted her from pursuing the argument about "easier to close [the clippers] from further back" and her analogy to tweezers. This line of argument might have led her to the "correct" solution, that it was not "easier" from further back, because the effort leverarm is shorter in such a situation. Note again the apparent reasoning from the fulcrum: her exclamation signaling some additional insight in line 102 follows the identification of the fulcrums and loads. She identified all the components of the top levers correctly but only the load in the bottom one, and we left the problem with her statement in line 120 as the answer.

E6 was the only student who seemed to be reasoning from his qualitative lever principle. He identified the fulcrums as Q in A and X in B, which are not correct, but the conclusions from these observations cannot be faulted:

069 I: [Pause of 8s.] Can you explain the difference between the two clippers?
070 E6: Well, the effort, no // The effort turning arm is shorter in this one [A].
[Identified fulcrums, and efforts, and concluded.]
080 E6: I'll say B.
081 I: B?
082 E6: Yeah, because the effort turning arm is longer.

The Revolving Door Problem. The best example of a qualitative model underlying the levers principle in an explanation was given by student E2 for the revolving door question. She initially gave a muddled explanation:

001 E2: ...If I pushed a door open from the end, it will open... but if I pushed it open from the middle it'll open a lot faster, I mean it'll open a lot farther...

I then asked her to think about the door in terms of levers, and she responded with the most sophisticated argument in the study.

002 I: Can I ask you to think about that one in terms of levers? Do you think there's a turning point and stuff like that?

003 E2: This is the turning point right here [labelled T in drawing]. For - for Ann, this would be the load. And for Beth this would be the load and this would be the effort and this would be the effort [for Ann]

She reconstructed the problem into two levers, drew these levers to resemble the simple levers used in the lesson and then reasoned from the fulcrum-helps model about her final answer. A copy of her drawings is given in Figure 4.13. I have labelled the load, fulcrum and effort in her drawings and my additions are distinguished by *.

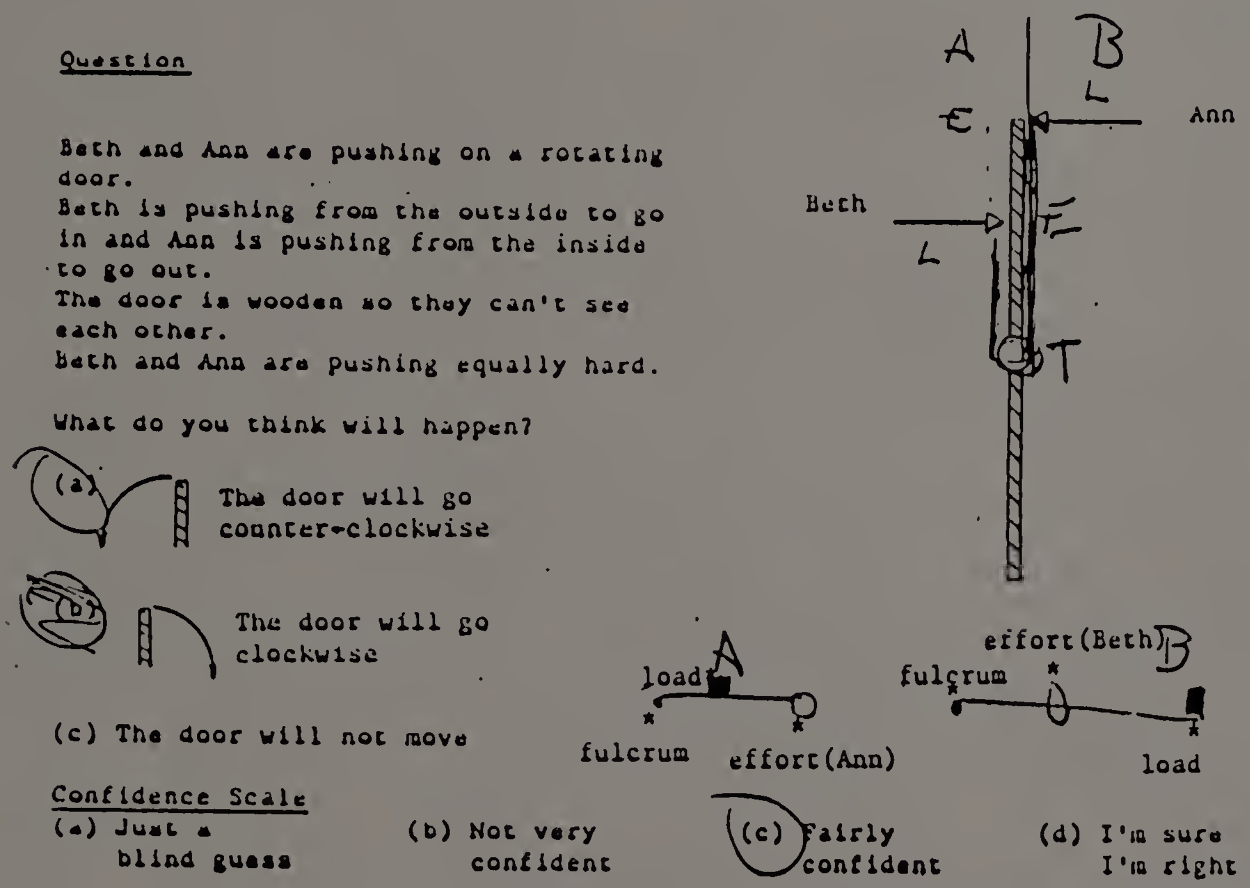


Figure 4.13

Revolving Door: Solution by E2

Her labeling on the original sketch is divided in two vertical parts: on the left is the load (L) and effort (E) for Ann (A) and on the right the same for Beth (B). She stared at this for a few seconds:

007 E2: That really makes me rethink what I've just said. Uhh / Let's see [started drawing the horizontal levers labelled A and B].
008 I: OK, so you drew Beth and Ann as load ..
009 E2: {As Ann looks and Ann / and Ann looks easier when I draw it out like that, because Ann

has the turning point taking some weight.
010 I: Uhhm?
011 E2: And / Beth has / None. So I guess I'm gonna have to redo that one, because after drawing it out, Ann would be pushing //
012 Ahhh, I get it! Beth's turning / Effort turning arm is this long and...Ann's is a lot longer. So, that will make it easier for her [drew vertical lines to show lengths of leverarms].
013 So I guess, the door would go counter-clockwise and Ann would get the door open her way.
[Marked confidence scale fairly confident.]
016 I feel more confident about that after I drew that out.

The fulcrum-helps model appears in line 009 and she generated the principle later in line 012. This example illustrates the idea of nonrule-like conceptions so eloquently that one should probably leave it at that.

b. "<u>Added Efforts</u>": <u>Examples of Lever Recognition</u>. An essential prerequisite for successful far transfer is the ability to break the complex machines in the far transfer questions down to the composite levers. The protocol analyses suggest that these four students were able to analyze the complex machines in this manner, as was illustrated in the excerpts from E2's revolving door explanation. The shadoof problem yielded some good examples too:

019 E2: B. Because the turning point is on the same side as the effort - I mean the effort is in the middle of the turning point and the load.
021 And so, since these weights [indicated the counter-weights] are probably the same, it makes no relevance, doesn't matter.
025 I: So, that irrelevant weight, what do you think is it doing there?
026 E2: The weight is to help the person / Just pull it out of the water. And it *will help* the person pull it out of the water, but when you're trying to figure out the problem, you have to //
027 *Not* think about that 'cause that's the same [her emphasis].
028 And you don't think about the turning point, 'cause the turning point's the same. You don't think about this

[pointed to the effort], what's left is you have to look at the length.
029 I: The length? / The turning arm?
030 E2: The / The uhm, the load turning arm, to figure out / Uhm, which one's would be easier.

E3 reasoned from the principle of levers, gave a good explanation and was sure about her answer. I probed, as in other cases where students did not mention the other levers themselves, and she indicated that she had noticed the other lever:

022 I: OK, when you look at that thing, how many levers do you see in that?
023 E3: Uh, I see two, because there's also the weight there [counterweight]. So then that's another / That also helps push it [bucket] up, but I saw that they were the same. I didn't think it mattered.

E5 thought very carefully about the problem before she came up with these insights:

035 E5: Uhm / I don't know, that's a hard one.
036 I: Do you think a thing like that could work?
037 E5: // Uh, yeah.
038 I: Want to try to explain to me how it works?
039 E5: Well / The water weighs a lot, so it's hard to pull it up. So the weight [counterweight] pulls up some of it?
040 I: The weights on the other side?
041 E5: Yeah. But the weight can't weigh more than the bucket, or else it also / It wouldn't go into the water?
042 I: Oh, I see.
043 E5: So, the guy has to, like pull up the extra amount of it.
044 I: So he has to push that little bit more?
045 E5: Yeah, yeah. And I think / Ohhh, that's hard, so it sort of has two weights?
046 I: There are two weights?
047 E5: So this one is pulling here [person], or pushing it.
048 I: Uhhm?
049 E5: // So, I don't know [stared at problems for 5 s].
050 I: Let's see - what is different in the two?
051 E5: The // Ohhh! / This // Which would be easier?
052 I: Yes, the question is which would be easier.
053 E5: B, I think. 'Cause the load turning arm is shorter. // If this is the load [pointed to bucket]?
054 I: OK, what do you think, is it the load?
055 E5: Yeah, and this [counterweight] is like more effort, sort of / In a way?

c. Summary. The summative analyses suggested that the experiment 2 participants were more able to transfer their acquired knowledge to the complex and conceptually far problems in the posttest than the experiment 1 participants. I think that this first indication of possible successful far transfer was supported by the qualitative analyses presented above.

A more interesting finding about students' reasoning that emerged from the formative evaluation of the lesson also seems to be supported by the qualitative analysis of the far transfer explanations. It was shown that all four students had acquired a qualitative principle of levers that appeared to be reasonably well assimilated and "automatized". It seemed that they were using only the principle in a "rule-like" fashion towards the end of the lesson in the near transfer questions for class I levers, e.g. "...the effort turning arm is the same but the load turning arm is smaller [hence B is 'easier']..." (E6 in line 386; and "...I think it will be easier with B because then the effort turning arm is longer than the load turning arm..." (E3 in line 326).

The lesson evaluation suggested that this principle evolved from a qualitative, intuitively-anchored model, and that this model and particularly its causal "fulcrum-helps" mechanism, was essential to an understanding of levers in general. Such an understanding implies that a person could make inferences about novel and more complex situations involving levers, that would not be possible for a novice or layperson without the deeper understanding. I believe that the lesson protocol analyses illustrated conclusively that the four successful students had constructed the intended model. In what I regard as the most stringent "test" of understanding in the experiment, i.e. the far transfer questions, the protocol evidence supported the emerging hypothesis that students needed the model to illustrate understanding, even though the principle was apparently securely in place. In most explanations the children simply used their new "rules", but in their explanations for the more difficult problems, they apparently reasoned from the model, rather than directly "applying" the principle.

One should probably try to find counter-examples to the suggestion above, thus instances where students are able to solve difficult, far transfer problems, and give appropriate explanations, apparently without such an explanatory model available. I shall discuss the two "unsuccessful" students next with this in mind.

9. Where the Lesson Fizzled Out

Two students did not change their original control misconceptions in their reasoning about the simple lever questions in the posttest, although the summative evaluation of the far transfer posttest questions indicated that some appropriate conceptions were used by both in the posttest. I shall identify the instances of failure within the lesson that followed successful anchoring or extreme case examples and try to present arguments as to why the lesson's aims were not attained in these two cases.

a. Preconceptions. We were immediately at a disadvantage in El's interview. She gave a "sure I'm right" confidence rating to an explanation for the target problem that obviously made a lot of sense to her. The target problem and her analogy to the arms of a man and a chimpanzee respectively, is given in Figure 4.14.

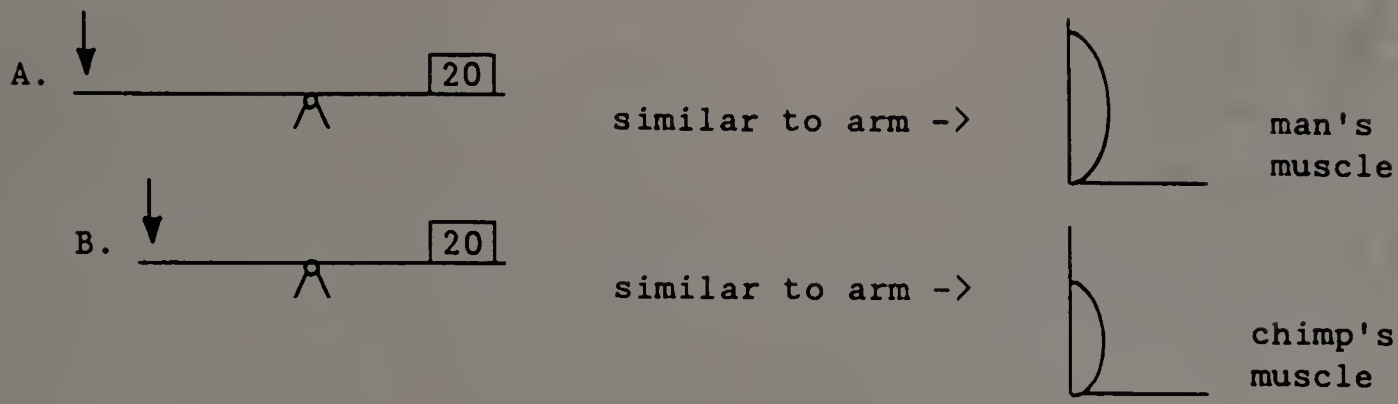


Figure 4.14

Target Problem and Analogies from E1

Her explanation follows:

001 E1: It would be easier at B.
002 I: At B?
003 E1: Uhhuh. Should I check it?
004 I: Yeah, and what makes you think that?
005 E1: Because I just did this actually in science class yesterday? Because I'm doing a report on the muscular system, and Mr. Covelli was explaining to us the difference between a man's muscle and a chimpanzee's muscle and a chimpanzee's muscle is connected here and the man's muscle is connected here, so it's easier to lift it, cause it's connected here you know (chimpanzee)? And I actually tried it with a board.

Her explanation contains the dilemma mentioned before: how to change an idea that is a misconception in two of the lever classes (I and II) but a perfectly sense-making, albeit not sophisticated, explanation for class III levers. One is therefore up against an idea that made sense, gave the correct prediction for her experiment with class III levers and was explained (correctly and no doubt very well) by one of the best science teachers I have ever seen in action. She was the fifth student interviewed, and she was followed by E3 (also in E1's science class) in one of the most successful interviews in the study. It is therefore not clear that the lesson could not have had any impact on E1's recalcitrant misconception merely because of the factors mentioned above - E3 did the same experiment, but was able to change her preconceptions. One could therefore hypothesize that the instruction may have an effect on the control misconception, however deeply believed. One assumption underlying the design is that the anchoring and extreme case examples should trigger equally strong appropriate intuitions (when compared with naive misconceptions) from which the principle of levers could evolve.

E4 held a more standard version of the control misconception, evident in his pretest and posttest explanations.

Pretest explanation:

003 E4: I don't really know, I mean I can imagine that you have this much less board to pull on.
006 Looks like B, because you're closer to this [load]. It seems like if I was doing it, this would be what I'd wanna do, it seems easier.
007 I think A, because you're closer to it [the load].

Posttest explanation:

001 E4: B, because there is less board and you're closer to it.
002 I: Closer to?
003 E4: The weight.

The only really distinctive aspect of the early part of his interview was his amusement: he could not understand why I asked him the "same question over and over", probably an early indication that his control misconception was generalized across all lever types.

b. <u>Identifiable Instances of Failure</u>. I shall present each student's progression through the lesson rather than in the separate sequences.

Summaries of the two students' explanations across the entire lesson are presented in Table 4.21. "Fulcrum-helps" is sometimes abbreviated to "fh" when more than one conception were identified in explanations.

Table 4.21

Summary of the Unsuccessful Interviews: E1 and E4

	Student E1's Conceptions	Student E4's Conceptions
Bridging		
Target	$m.c_c$.	$m.c_c$.
Anchor	sharing	sharing
Bridge	fulcrum-helps	fulcrum-helps
Extremes	fulcrum-helps	$m.c_c$.
Target	$m.c_c$.	$m.c_c$.
Extreme Cases		
Change d_{Lf}	$m.c_c$.; sharing	fulcrum-helps; non-gen
Change d_{Ef}	confusion	fulcrum-helps
Bridge	fulcrum-helps	fulcrum-helps
Transfer 1	$m.c_s$.	fulcrum-helps
Transfer 2	m.c.	surface features
Transformation		
Target	non-gen3	m.c.
Comparison 1	fulcrum-helps	fulcrum-helps
Comparison 2	non-gen3	fulcrum-helps
Target	non-gen3; load leverarm	confusion
Transfer 1	m.c.	non-gen3
Transfer 2	non-gen3	m.c.
Benchmarks		
Target	analogy	load leverarm
Anchor	symmetry	symmetry
Benchmark 1	$m.c_c$.	effort leverarm
Benchmark 2	$m.c_c$.	principle
Intermediate	$m.c_c$.	confusion
Target	load leverarm; $m.c_c$.	principle
Transfer 1	surface feature	load leverarm
Transfer 2	surface feature	$m.c_c$.

The most problematic sequence for both students was the last, the class I levers section. This sequence relies on students having an intuition about the extreme case situations and the analyses of the protocols suggest neither of these students held such intuitions. There was evidence that both were accepting the fulcrum-helps idea for the class II levers and E1 in particular still used part of the principle -

"the load is closer to the turning point argument", in the benchmarks sequence. Both struggled to apply this idea though, and did not change their minds about the class II target question - the control misconception is still evident in the class II simple levers questions posttest.

From an analysis of the responses summarized in Table 4.21 one could attribute the failure of the lesson to bring about conceptual change in these two students to two factors: first, the students apparently did not hold the strong intuitive beliefs on which the class I sequence depends to separate the non-generalizable conception into the two more appropriate leverarm variables; and second, neither of the two were able to generalize the fulcrum-helps idea to levers beyond the simple lever structures used in the lesson.

The explanations from the benchmarks sequence illustrate the first statement. (Benchmark 1 - load leverarm 1 ft and the effort leverarm 10 ft).

216 E1: Uh // Ten feet. / He's gonna have to push probably 30 lbs. Because it's farther away and the weight's down at the end and like I said about the seesaw, it's easier if both people are the same weight and it's gonna seem like this weight is more because it's farther away.
217 I: Farther away from?
218 E1: From the man and it's closer to the turning point.

237 E4: This is longer [effort leverarm].
238 I: Do you think that would make a difference? // Have you done anything like that?
239 E4: Probably have, just don't know it. Seems like it would, yeah.
240 I: Could you tell me more?
241 E4: Seems like it would make this [load] less, I don't know.

Benchmark 1 (load leverarm 10 ft and the effort leverarm 1 ft).

219 E1: Probably / I think 30 lbs again, because it's just like this [benchmark 1] but it's different, because he is sort of, like closer to the turning point, but this is farther [load], so it's gonna be harder for him because he has

less to push with.
220 I: OK, so how come that will be 30 lbs on this too?
221 E1: Because it's sort of the same. Because he is dealing with pushing down 10 ft with 1 ft here and this man is dealing with pushing down 10 ft too.
222 I: Ahh // He's pushing down 10 ft too?
223 E1: Do you understand? 'Cause if this [fulcrum] were in the middle it would be easier, so it's gonna be the same, because of the same measures and the load's just in a different place.
224 I: And that doesn't matter?
225 E1: Right, I don't think so and I'm not very confident.

249 E4: [15s pause.] I don't know.
250 I: What changed?
251 E4: These two switched.
252 I: What switched?
252 E4: The effort turning arm and the load turning arm.
253 I: So how does it look now?
253 E4: Looks harder.

It may seem that there is evidence of some understanding in all the excerpts for E4 above, but the confusion becomes apparent when one views his answers to the near transfer questions as well as the final written statement.

He answered the crowbars question correctly, but reverted back to the control conception for the pliers question, after comparing this to the simple levers used in the lesson:

The crowbars question:

273 E4: This one, B [crowbar would be easier to pull nail out].
274 I: You got that answer quickly?
275 E4: Yeah, just because the effort turning arm is longer.

The pliers question:

285 E4: ...the load is farther out from the turning point here [A] than it is here [B].
286 I: And the load is?
287 E4: The thick wire.
288 I: And does it make any difference to which one would be easier?
289 E4: Well in the other thing it would [drew a simple lever, class II, then class I].
292 E4: So this would be easier, the load is farther away from the turning point than it is in B.
293 I: Is it that way in your drawing?

294 E4: This one [A] would be easier?
295 I: Are you sure about that one?
296 E4: Not really.
297 I: What would make it easier to understand?
298 E4: Doing it in real life.

His last statement, "Doing it in real life." proved to be no help either, since he used the simple levers apparatus throughout to test his answers or to confirm hesitant predictions. He was the only student who expressed (many times) the desire to use the apparatus and since I had intended empirical feedback as a last resort, this seemed to be an ideal opportunity. However, the experiences with the apparatus did not seem to make any notable difference to his understanding, since his posttest explanations all included the control misconception. This finding is in agreement with Driver's (1989:89) observation that "...observational evidence ... is not enough by itself for pupils to reconstruct their ideas." However, the effects of empirical feedback on learning where the child has no preknowledge, should be explored in future research - it presents at least one way of constructing such knowledge.

The written statements of the two students is another indication of the limited impact the lesson had on their understanding (statement 1 is given first):

140 E1: OK. I think the lever would make it easier, because if you're trying to lift the load yourself, you're lifting the whole weight. However, with a lever helping you, you're really only lifting half the weight, OK?

282 E1: Alright, I'm going to change something, 'cause I said you're only lifting half the weight, but you're not necessarily lifting only half the weight; you could be lifting a quarter of the weight or three quarters of the weight, not just half.
283 I: OK, depending on what?
284 E1: Right, depending on where the turning point is.

And E4's statements:

122 E4: Just thinking of it in real life.

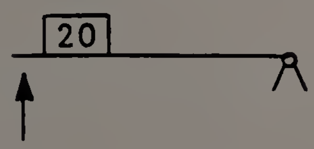
299 E4: Well, I mean for the pliers [near transfer question], referring it back to the picture I've been using in the entire lesson the other time helped. Shall I write that?

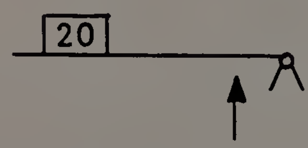
c. Positive Outcomes. Both the students were able to identify most of the fulcrums, loads, efforts and leverarms in the near transfer questions but only E1 spontaneously mentioned these elements in the far transfer questions. In the revolving door question she noticed that "...how far away from the turning point they are..." (line 007) would make a difference, but discarded the notion; while in the nutcracker problem she identified the load and effort and decided that a smaller load-effort distance would make it easier to crack the nut.

The class II bridging and extreme cases sections were marginally successful and there was evidence of emerging fulcrum-helps models for both students: "...because it's [fulcrum] holding it up, it's still doing half the work..." (E1 in line 015); "...because the weight is closer to the triangle on this one...I'd probably say that the triangle has more pressure on it..." (E4 in line 078). All of these apparent strongly intuitive beliefs capitulated before the control misconception, as E1 explained for the second class II target problems:

021 E4: ...in some, if it [load] were right on him, well - not on him but if it's right where he's holding it, I think it's harder than if it's closer to him, you know?

Her last remark suggests that more, and less extreme bridges to the target situations may be effective, to establish a "direction of change" in the variable relationships, i.e. that a smaller load leverarm always results in less effort needed etc.

In the extreme cases revisited a new problem surfaced. El apparently saw additional confirmation for her control misconception since A in 

compared with B in 

, was having a much easier task, since "...he's holding it closer to him you know..." (line 154). Although her predictions for the sequence were correct, the unintended support lent to the misconception should have been avoided, and suggests closer scrutiny of other teaching situations where a misconception explanation may yield a correct prediction.

Most of the positive outcomes were temporary and superficial, but one may infer from the success of the first two teaching sequences and the resulting rather transient, but positive outcomes, that the two "unsuccessful" experiment 2 students learned more than the most unsuccessful students in experiment 1. However, the problem of non-existing physical intuitions is a vexing one, since direct feedback from simple experimental simulations of the simple levers failed to strengthen emerging beliefs or to convince the student of unexpected findings. It is difficult to see how one may "get" a person to believe in a phenomenon when physical experiences and thought experiments apparently failed to do so. Real classroom situations and peer interactions will provide learning experiences that did not come into play in this study, and a larger, classroom study may speak to this issue.

CHAPTER V

GENERAL FINDINGS FROM THE EXPERIMENTS

The general research questions that guided the study can be answered at this stage. The nature of students' naive knowledge of levers and the effect of instruction designed around students' naive knowledge base were the two prominent issues in the studies. I shall summarize the major findings from both experiments before the discussing the children's knowledge and instructional design issues.

A. Summary: General Findings From Experiment 1

Children's Preconceptions. A "new" misconception, probably related to the symmetrical anchors in the class I and II teaching sequences, surfaced in experiment 1. The protocol analyses suggested that children's naive knowledge was more substantial and elaborated than initially believed. Both these issues will be discussed in more detail in section C below.

Summative Evaluation. Differences were found between the control group and experimental group 1 in the quantitative pre-posttest comparisons. One, the experimental group scored significant lower on the class I levers question, thus an indication that the particular teaching sequence was flawed. Two, the experimental group students used more appropriate conceptions, that is conceptions compatible with physical theory, than the control group in the far transfer posttest questions. Overall, participation in the lesson interviews did not benefit the experimental group, although two students changed their initial misconceptions, added to non-generalizable conceptions and constructed a qualitative principle of levers during the tutoring interviews.

There was evidence that a class II levers question fostered limited conceptual change in two control group students, and one of the control group students appeared to have acquired a more general, appropriate understanding of levers in the interval between the diagnostic test and the pretest interview. However, all three of these students performed poorly in the far transfer posttest questions, suggesting a limited understanding of levers.

Students' Learning Processes. Descriptions of problem areas in the instructional design identified in the summative evaluation were expanded in the lesson interview protocol analyses. Three major issues informed the design of the lesson for experiment 2.

One, the extreme case comparison problem in the class II lever bridging sequence was powerful enough to facilitate the learning intended in the sequence by itself. The example elicited a physical intuition important to the understanding of this class of levers from all six students. This is in contrast to the first part of the bridging sequence where three students were unable to analogically extend an anchoring intuition to the first bridging example. As a result, extreme cases were added to lesson 2 to both facilitate and reinforce the students' apparently transient models of class II levers.

Two, only students who had constructed a qualitative, explanatory model of the fulcrum in the class II sequence were able to generate a qualitative principle of levers by the end of the lesson. The addition of extreme case examples in lesson 2 to the class II sequence was seen to be important to the children's construction of such a model in experiment 2.

Three, the sequences dealing with class I and III levers were unsuccessful. The class I sequence facilitated confusion rather than learning and the class III transformation sequence was apparently too long and fragmented. For lesson 2, the transformation sequence followed the class II sequence, and it was shortened into two extreme case comparisons in which the class II to III transformations were obvious at a glance. The class I sequence was moved to the end of the lesson, where it was hoped that the central position of the fulcrum would encourage a final separation of naive variables into the principle's leverarms.

B. Summary: General Findings From Experiment 2

Summative Evaluation. Significant differences were found between experimental groups 1 and 2 in the pre-posttest comparisons with regard to conceptual change and far transfer. The group 2 participants fared better on all but the class II levers posttest problems. Lesson 2 was therefore successful in bringing about conceptual change as well as fostering a deeper understanding of levers, evident in the group 2 students' ability to transfer their acquired principles and models to difficult, complex and compound "real" levers and in conceptual changes apparent in the simple levers questions.

Students' Knowledge Construction. Four of the six students in experimental group 2 significantly changed their preconceptions about levers and were very successful in the near and far transfer problems. A common process of knowledge construction as a result of the instructional techniques emerged from the protocol analyses of these students' lesson interviews. This process included: one, the construction of the fulcrum-helps model in the class II levers teaching sequence; two, the separation of the children's naive, non-generalizable load-effort distance variables into two variables, i.e. one of the leverarms and a non-generalizable naive variable in the class II sequence; three, the use of the class II fulcrum-helps model to understand the function of the fulcrum in class III levers; and four, the final separation of the non-generalizable lever variables into the principle's two leverarms.

The acquisition of the qualitative principle, as well as a fulcrum-helps model, was essential for far transfer. There was evidence that students relied on the model in difficult problems to construct satisfactory solutions - thus evidence of the model that underlie their qualitative principle of levers.

Both the unsuccessful students had constructed a fulcrum-helps model, but lacked the physical intuitions on which progress in the class I lever sequence, where the principle was finally "constructed", depended. The class I sequence remains therefore the most problematic area in the design at this stage.

C. Research Question: Children's Knowledge of Levers Before and After Instruction

Analyses of the participant protocols in experiment 1 suggested that the children's naive conceptions of levers appeared to be more complex than what I had inferred from the pilot study protocols. This view was supported by the evidence from experiment 2. Two misconceptions have been identified in the experiments and students also used alternative, non-generalizable conceptions that presented new, distinct problems to the design of the instruction. In this summary statement the naive conceptions are reviewed again, and I want to conclude with a description of the "final state" of the successful students' knowledge of levers.

In the discussion "force" will be used as a collective term for the students' own terms such as "holding", "supporting" and "pushing". I am not implying that the students held an appropriate force concept, merely that there is a naive notion of "force exerted" implied in their statements.

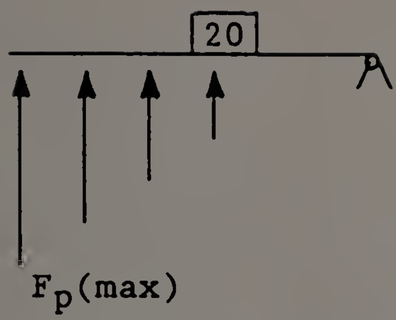
1. Naive Ideas: A Symmetry Fixation

This misconception appeared for the first time in the entire project in experiment 1 and appeared frequently enough to cause concern. The children seemed to regard symmetrical levers, where the leverarms are the same length, as the most ideal, effective levers in any situation. The symmetry misconception was particularly evident in the class I lever sequences. This belief about class I levers may be due to children's experiences with seesaws as they often indicated in the anchoring examples. Else it may be due to the fact that the symmetrical class I anchor was the only benchmark for most students, i.e. that students knew that a 20 lb load on a symmetrical seesaw was balanced by a 20 lb effort.

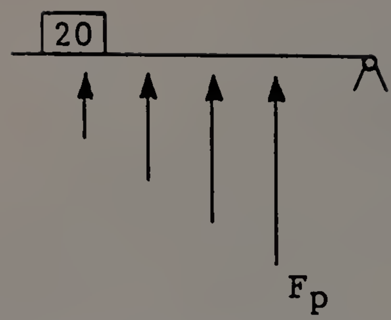
It is, however, an issue that deserves attention in further research, since the idea was compelling enough to cause one student to generalize the symmetry conception to all lever classes towards the end of the lesson. In future research, empirical feedback and group discussions may act as bridges to intuitive extreme case ideas that seemed to be repressed as a result of the competing symmetry beliefs.

2. Naive Force Diagrams

a. Conception of Control. The least appropriate of the naive conceptions is the control misconception, an apparently "egocentric" or "anthropomorphic" view, since the children were considering only the results of the person's force (F_p). A person is seen to exert a maximum possible force if that force is applied directly to an object, thus enabling the person to have more "control" over the situation. When the force is not exerted directly on the load, there is "less control"; and I infer, less force exerted on the object. I have called this a "dissipating force" notion and the diagram illustrating my representation of the idea is given in Figure 5.1 (a).



(a) Class II Levers



(b) Class III Levers

Figure 5.1

Naive Models

b. Class III Non-Generalizable Conception. One of the problems in designing the lesson was that the control notion described above could be labelled a misconception in class II and I levers and yet give the correct prediction when applied to class III levers (in Figure 5.1 (b)). There seems to be subtle differences in students' descriptions: in Figure 5.1 (a) it would be "easier since you can hold it more", but in the class III lever situations the person definitely has to "push" less. I have interpreted the latter idea as a non-generalizable conception rather than a misconception. The problem is obvious: if the children do not have a clear understanding of the inappropriateness of the control ideas in class II levers, one may reinforce the misconception ideas in teaching about class III levers, particularly if their naive, non-generalizable force ideas were to be used in the instruction.

c. Class II Non-Generalizable Conception. These ideas about the forces exerted in levers situations are more compatible with the physical theory, but are also seen to be obstructions to learning. In contrast to the single force in the conceptions above, the children attended to two forces: the person pushing up and the weight "pressing down" or pushing down on his hands. There seems to be the same notion of "dissipation" or distribution, but in this case of the *load's* weight. My representation of this naive force diagram is given in Figure 5.2.

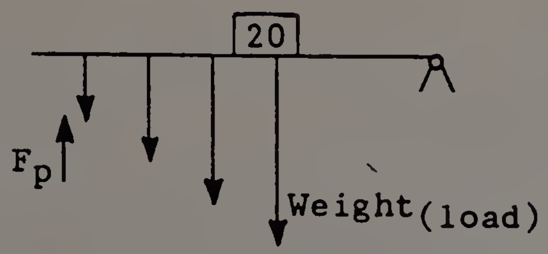


Figure 5.2

Naive Model

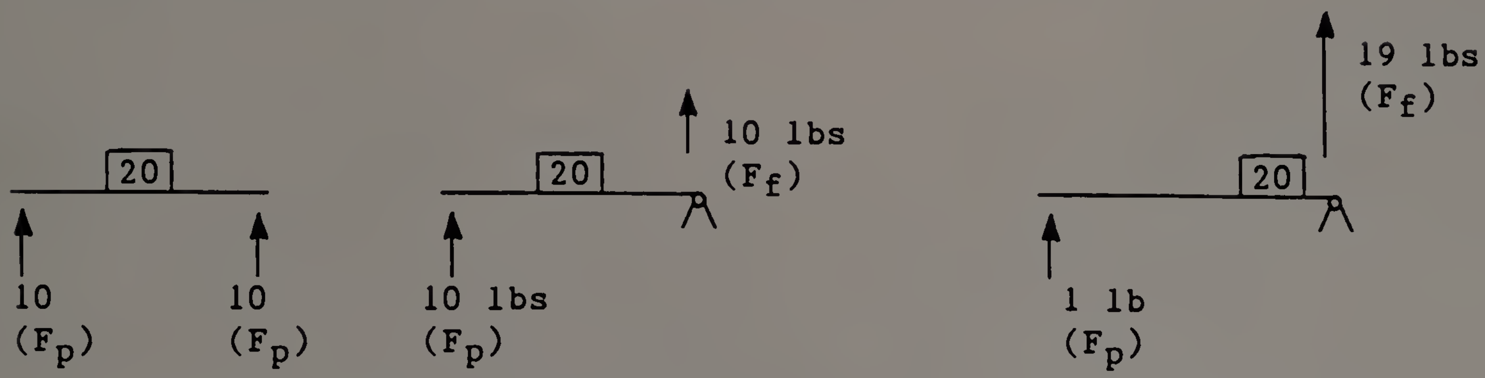
I interpreted their statements to mean: the further the person (effort) is away from the load, the less effort is required. I construed from their explanations that a "balance" of these two forces were important, thus that the person had to push up only the equivalent of the dissipated weight at any point on the lever. It seems therefore that the person is reacting to the load's weight, in contrast to the control misconception above, where only the person's possible output of force was considered.

d. Naive Force Diagrams as Barriers to Learning. The control ideas for all but class III levers are not in agreement with the levers principle. These ideas are therefore considered to be misconceptions, with all the concommitant characteristics - i.e. barriers to learning, interference with correct ideas etc. There is evidence that the misconceptions were applied across the spectrum of different examples in the experiments, thus implying that the naive ideas were broader in scope and used in an amazingly consistent fashion - an indication that they are substantial and interconnected.

I viewed the non-generalizable models as potentially useful ideas in the design of the pilot study and experiment 1 lessons. The students gave correct predictions to problems and their explanations were acceptable for some individual cases: it is certainly true that a smaller effort-load distance would result in an increased effort in the class II simple lever situations. I thought that this "abbreviated" principle of levers was quite ingenious: the one variable, i.e. the effort-load separation, summarized both the effort and load leverarm differences in the levers. The ideas are, however, cumbersome. First, the conceptions are obviously limited in scope when compared to a qualitative principle of levers. The non-generalizable conceptions require modification between lever classes; thus less efficient, and more fragmented knowledge than a qualitative principle of levers. Second, these ideas may be acceptable in qualitative explanations, but would be an obstruction to any quantitative problem solving in the domain; and thirdly, building on these non-generalizable ideas in experiment 1 probably resulted in more, rather than less fragmented knowledge.

3. A Model Compatible with Physical Theory

A missing element in all three of the naive conceptions above is a cognizance of the third force in the simple lever situations, namely that of the fulcrum. It was shown in Chapters III and IV that a qualitative, intuitive understanding of this force in class II levers was constructed by the majority of the experimental groups' students, even though some students were unable to assimilate and extend these ideas to a principle of levers (diagrams in Figure 5.3).



(a) Sharing the load (b) Extreme case

Figure 5.3

Equilibrium of Forces

The lesson in experiment 2 was grounded in two naive intuitions that students may have: one, the idea that the fulcrum (F_f), and the person (F_p) in a symmetrical simple lever situation are sharing the load of 20 lbs; and two, that the fulcrum (F_f), in an extreme case simple lever situation is obviously "pushing up" more than the person (F_p).

There is evidence in most of the students' protocols that this naive view of equilibrium was understood by the students, e.g in explanations for the extreme case in Figure 5.3 (b):

082 E2: This man will probably have to do 1 to 5 lbs...[and the
084 block will be holding up]...approximately 15 to 19 lbs.

053 S5: ...he is only holding up 3 lbs [load near fulcrum]...and this [load near person] then he would be holding up like 17 lbs...

Where students did not give a quantitative estimate the qualitative relationship $F_{fulcrum}$ increased implies F_{person} decreased, was given as an explanation, e.g.

072 S1: I think this person will find it easier // This person is far away and the block [fulcrum] is holding it [load].
073 I: The block is holding it?
074 S1: With more than the person, because it's [load] closer to it [fulcrum]. They are both holding it, but I think that

B [person] doesn't have to use so much strength.

And:

076 E4: Person B, because the weight is closer to the triangle on this one...I'd probably say the triangle has more pressure on it.

There seems to be two different ways in which students thought about these equilibrium situations. The most common statement would refer to the forces exerted by the person and the fulcrum (as stated by E2, E4, S1 and S5 above), thus a consideration of *forces on the load*. In the second type of statement, the point of view of the student changes and the result of the *load exerting a force* is implied.

Forces-on-Load Model. One is considering explanations that would suggest "the turning point holds/supports/pushes up more" or "more pressure on the turning point", implying that the person have to push up less. I am interpreting these statements as a view of the forces *on* the load (illustrated in Figure 5.3) above.

Load's Force-on-Others Model. For example, E6's explanation that "...the weight is focusing more on the turning point..." (line 054), suggests to me that he considered the load exerting force(s) down on the person and the fulcrum:

054 E6: [The person]...is lifting farther away from it [load], so the 20 lbs is focusing more on the triangle's side.
055 I: What do you mean "focusing more"?
056 E6: Well // All the weight, if it was spread out on a 20 lb bar, then it would be the same. But it's a block, so all the weight of the 20 lbs are / Most of the weight is going down on the / On the triangle...
059 I: And for man B?
060 E6: For man B most of the weight is going down on him / Or his hands, whatever.

These ideas are illustrated in Figure 5.4 (a) and (b) below.

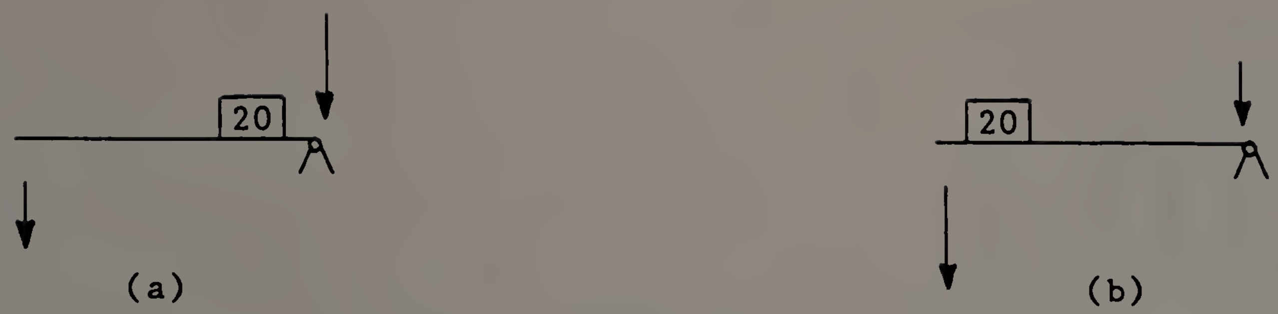


Figure 5.4

Load's Distribution

Other students explained that the load's weight would be "distributed" more towards the fulcrum (or the person), apparently another dissipating force idea. One may accept that the students view the person and the fulcrum's forces as reactions to the portion of the load's weight focused on, or distributed to, the fulcrum and the person.

For the second type of statement, the point of view of the student changes - the result of the *load exerting a force* is implied. The two models are both acceptable as outcomes of the lesson, since the aim was to include the fulcrum in their reasoning.

Summary. Students came to the lesson interviews with either a control misconception or non-generalizable conceptions. In those cases where the lesson's effects were apparent in students' increased understanding of levers, a model such as one of the two naive, but normative models above was constructed. This model then facilitated the separation of the load-effort distance variable in students' non-generalizable conceptions into the two leverarms important in the principle of levers.

The transfer questions in the two experiments were included only as measures of students' understanding of levers acquired as a result of the lessons; in other words, I did not set out to uncover the conditions which would reveal transfer or not. Even so, it was clear that all participants who had constructed the model and subsequently the principle, were successful in the far transfer problems. An interesting aspect of efficient transfer was that some students referred back to their fulcrum-helps models in the most difficult transfer problems.

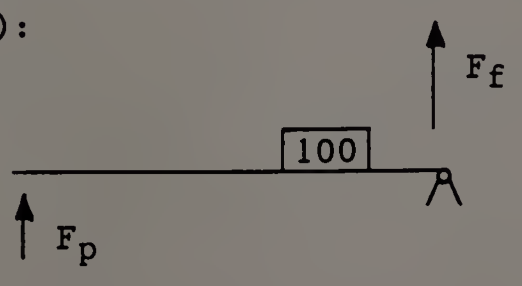
In contrast, the students who did not construct a qualitative levers principle failed to transfer, but attended instead to prominent surface features in the problems.

D. Useful Characteristics of a Limited Model

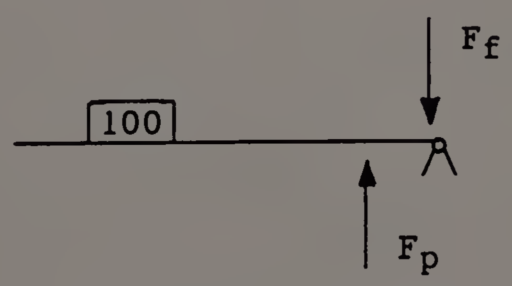
The fulcrum-helps ideas discussed before may seem limited; they are after all directly applicable in the class II lever situations only. However, they have some interesting characteristics, discussed below.

1. Robust Models

The fulcrum-helps idea about the fulcrum and its action allows one an overview of a lever situation to start reasoning from; for example when considering the two cases in Figure 5.5, it is clear that the fulcrum (F_f) "helps" the person (F_p) in (a), but "hinders" the person in (b):



(a) Class II lever



(b) Class III lever

Figure 5.5

Robust Models

The model seems more like a small set of concepts that requires minimal modifications to explain different levers situations. For example, in a typical explanation for the differences between case (a) and (b) in Figure 5.5, the non-generalizable conception may be mentioned; the forces exerted by the load, the person and the fulcrum could be considered; and the load-fulcrum distance appears as a separate variable in explaining the behavior of class II levers.

> 029 S6: [It's easier for]...person A, 'cause the block [fulcrum] is supporting more, because it's [load] closer to the block on this one...It's [load] closer to the man in example B, so the person would have to hold up more.

The robustness of the model becomes evident in explanations for other lever classes, say the class III lever. By just arguing that the fulcrum does not help, an explanation with a modified set of conceptions, e.g. the class III non-generalizable conception and the effort leverarm, could be generated from the same model. It seems therefore that the model represents a system of objects and relationships that allows considerable flexibility.

2. Generative Models

One of the most useful characteristics (instructionally) of the models is evident in the emergence of one of the leverarms as an important variable (for example in the quotation above), apparently as a result of the students' comprehension of the forces at play in the situation. One may propose that these student models are generative; that is, although the fulcrum-helps ideas are limited in scope, they provide ways to reason about the other two lever classes. Students were therefore able to think about class III levers as levers where the fulcrum does not help, and some used the model in their reasoning about class I levers. I have shown that the latter is probably not an appropriate way of reasoning, i.e. not acceptable to physicists, but this counter-intuitive property of class I levers may be useful in later quantitative problem solving as a conflict generating teaching strategy.

3. Causal Mechanisms

Another interesting aspect of the model is the causal mechanism that underlies it. I have discussed this earlier as an essential element of the model - the causal, explanatory mechanism allowed part of the principle of levers to emerge and, most important, seemed to bring a "sense-making" notion into the children's explanations. The mechanism also represents the part of the model that was described as the intuitive anchor, which suggests that the ideas that emerge from the model are not rule-like. There were a few examples in students' explanations for the far transfer problems that supported the view that the explanatory, causal element of the model was potentially very powerful - when students encountered difficult situations they returned to this intuitive, explanatory model to make sense of novel and difficult cases.

E. A New Instructional Technique

The intuitive appeal that extreme case situations have for learners was, for me, the most exciting finding in this study. The use of expert reasoning strategies, e.g. analogical reasoning, as instructional "tools" has been investigated extensively. However, children's ability to reason from extreme cases appear to be a largely undeveloped domain.

Some characteristics of reasoning from extreme cases have been suggested by the findings in the experiments:

Implicit, intuitive knowledge structures were initiated as a result of the limiting case examples presented to the children. For some students the extreme cases triggered ideas directly in conflict with other existing conceptions, while for other students they appeared to strengthen existing intuitions. For example, in the former case, a child holding an inert-objects-don't-push misconception could believe that the same inert object exerts a force in an extreme case; while, for the latter, a tentative belief that the inert fulcrum may share a load in a symmetrical class II situation, is strengthened by the intuition for the extreme case.

The possibility that students would activate new schemata, or at least rethink their original ideas was increased in both the cases outlined above.

Increasing the Scope of Students' Knowledge. The intuitively understood, qualitative idea that was reinforced by the extreme case was then extended analogically to other problems, allowing the use of more appropriate ideas and at the same time, broadening the scope of students' knowledge.

New functional relationships between variables also emerged as a result of the children's reasoning about the extreme case situations. For example, the extreme case in which the load was almost on the fulcrum generated the load leverarm relation to effort for many students.

Although the presentation of extreme cases fostered learning, there was only one instance of a student spontaneously using an extreme case situation to explain a decrease in effort. This was by S2 in the target question to lesson sequence 1 in experiment 1:

> 058 S2: ...if you were holding it [board at the end] when the board went all the way to the end of the library you'd barely have to hold it at all.
> [We were sitting in one corner of the library.]

The children did use analogies (spontaneously) in this study, but often inappropriately (from the point of view of a physicist) and mostly on the basis of surface features. In contrast, there was no evidence of reasoning on the basis of surface features when dealing with extreme cases. However, the domain is unexplored and one would have to look exclusively at children's ability to spontaneously generate extreme cases, as well as their thinking in response to presented problems in order to suggest an understanding of the reasoning processes involved.

This study suggests that, as an instructional move, extreme cases are appealing. Nersessian (1990) suggested doing analyses of the writing of eminent physicists to obtain instructional examples, such as Galileo's reasoning about falling objects. The idea is not to "teach" a strategy, but to foster the use of such strategies by presentation of the examples, as A. Brown (1988) did in her studies that investigated the promotion of the use of analogies by young children.

CHAPTER VI

EPILOGUE

The study yielded many small surprises and observations; but also, in my opinion, some valuable new insights in the following different contexts: explanatory, causal models and learning and the ability of young children to reason in sophisticated and creative ways about the physical phenomena in the study. These issues will be presented as concluding remarks and with ideas about further research that have been suggested during the study.

A. Models for Robust Understanding

I am not suggesting that "having a model" would automatically solve all problems in learning about levers. It was clear from the study that some students' intuition about the fulcrum was not sufficient by itself to overcome misconceptions or even prevent the construction of inappropriate ideas. Yet, the evidence from this study suggests that the qualitative principle of levers was only evident in the protocols of students who were able to reason from the fulcrum-helps model.

The creativity apparent in a typical levers principle construction process is remarkable. The students came to the interviews with alternative, incomplete models of levers, with two, sometimes three variables rather than the four stated in the principle, and with various misconceptions. From this state they were able to add to their knowledge, separate the one variable into two more acceptable to the physical theory, change conceptions and perhaps most impressive, transfer their qualitative principles to difficult, novel problems. There is evidence in this study of conceptual change leading to new and robust understanding and, in the case of at least one student, his new qualitative understanding allowed him to construct the quantitative principles in this domain.

One direction for future research would involve further development of the lesson to eventually "teach" the quantitative levers principle. It should be interesting to compare students' quantitative reasoning about balance beams if they learn the principle via the fulcrum-helps model, with that of children in balance beam experiments (without models) such as those that were conducted by Siegler (1978, 1982) and Hardiman (1983).

B. Creative and Sophisticated Reasoning by Children

I came to this study believing in the abilities of children to make sense of their world, and was pleasantly surprised to have my expectations exceeded. I still believe that the clinical interview context was not an environment most conducive to motivation and inspired thinking. Nevertheless, all the children except one foreign student who had very little English, amazed me with their application and creativity.

There are numerous examples of their creativity and abilities as "naive scientists": the generation of an extreme case cited in Chapter V; several examples of their spontaneous use of analogies to explain situations to the silent and uncomprehending interviewer; the sophisticated analysis of student E2 of a far transfer problem that was reminiscent of an expert interviewed on the same problems; the spontaneous construction of a conception of the center of mass by student E6; and perhaps most impressive, S6's successful struggle to formulate the quantitative principle of levers. They accomplished the learning primarily on their own, with only the study's problems as the primary directive, and in an environment devoid of any external motivations, teaching aids or novel inventions such as "interactive learning environments".

I believe that this study adds to the research currently accumulating about "children as scientists" (Driver, 1983) and the accompanying implications for instruction.

The idea is not to encourage the variation of "discovery learning" where the child is left alone with materials, but rather to provoke children into using their latent abilities. To be sure, finding examples that may "provoke" or initiate the kind of reasoning and subsequent learning desired is not an easy task, as the lever study clearly illustrated. What seems to be needed is a model of curriculum development and research that incorporates and values the child's potential and view of the world.

APPENDIX A

LETTERS OF CONSENT

PARENTAL PERMISSION FOR SCIENCE LEARNING STUDY.

My name is Aletta Zietsman and I am a graduate student in the School of Education at the University of Massachusetts. I am associated with the Scientific Reasoning Research Institute at the University which is conducting a study of factors that may influence the learning of science. The study should help us to design better science courses. I would like to request your permission to have your daughter or son participate in the study. The data from the study might be included as part of my doctoral thesis and possibly included in other reports of the study.

Each student will be asked to solve real life science problems, concerning the use of levers. Students' names will not be used in reports of the study and the results will not affect the students' grades. Further, participation or lack of participation will not have any effect on students' grades. Participation is voluntary and consent can be withdrawn at any time.

Please sign and return this form indicating your wish about your daughter or son's participation in the study. Please call me if you have any questions.

Aletta Zietsman
Scientific Reasoning Research Institute
University of Massachusetts, Amherst.
(tel #: 545-2077 or leave message at 545-0988)

Please check one of the following and sign below:

____I give permission for ____________________________to participate in the Physics Learning Study.

____I do not give permission for________________________to participate in the Physics Learning Study.

____
Parent or Guardian

______________________________ ___________1990
Student Date

PARENTAL PERMISSION SLIP FOR STUDENT PARTICIPATION IN A SCIENCE LEARNING STUDY

A science education project at the University of Massachusetts is conducting a study of factors that influence the learning of science. The study should help us to design better science courses. We would like to request your permission to have your daughter or son participate in the study.

The study will involve two 30-40 minute interviews after school, and the students will be paid $5 per interview. Each student will be asked to solve some science problems concerning the use of levers in simple machines aloud in an interview. The interviews will be audio or video taped. This allows us to keep an accurate record and to study the learning processes that occur. Students' names will not be used in the reports of the study and the results will not affect the students' grades. I shall use the interview data in my dissertation and for contributions to journals and conferences. I may also use the video tapes for demonstrations at conferences. All identities will be protected at all times.

<u>Participation is voluntary and consent can be withdrawn at any time</u>.

Please sign and return this form if you wish to give your permission for participation in the study. Students usually find the problems interesting to solve. Please call one of us if you have any questions.

Aletta Zietsman or Prof. John Clement
SRRI, 314 Hasbrouck Laboratories
University of Massachusetts, Amherst MA 01003
(tel#: 549-0988)

I give permission for ______________________________ to participate in the Science Learning Study.

____
Parent or Guardian

______________________ ________1990
Student Date

I give permission for the interview tape to be used at a seminar or a conference on science learning. (<u>Names of participants are not used</u>.)

____I give permission for data to be used in this way.

____I do not give permission for data to be used in this way.

____
Parent or Guardian

____
Student

APPENDIX B

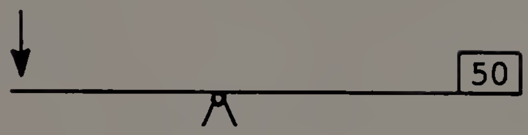
DIAGNOSTIC TEST

Question 1

Two 50 lb loads are held level on the two different boards as shown. The boards are light but strong.
Where would it be easier to hold the board level?

A.

B.



__(a) In case A
__(b) In case B
__(c) The same force would be needed to hold each board level

Confidence Scale

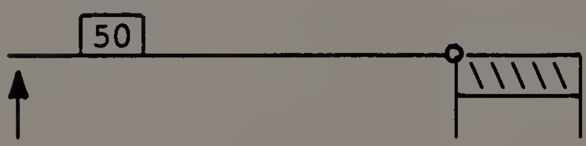
(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

Question 2

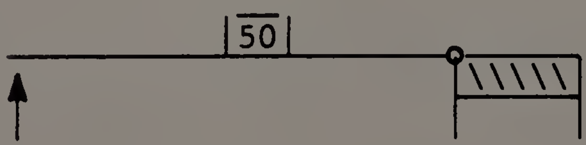
Two very light but strong boards are fixed with hinges to the edges of tables as shown.

Where would it be easier to hold the board level?

A.



B.



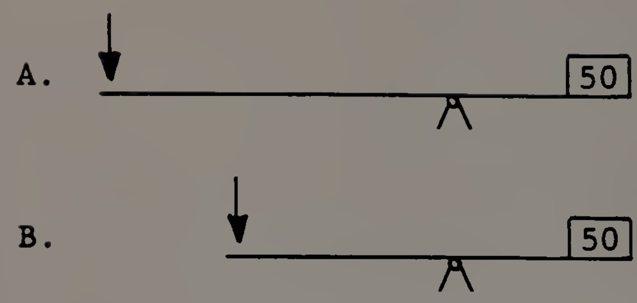
__(a) In case A
__(b) In case B
__(c) The same force would be needed to hold each board level

Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

Question 3

Where would it be easier to push down to just lift the 50 lb load on the end of the light, sturdy board?



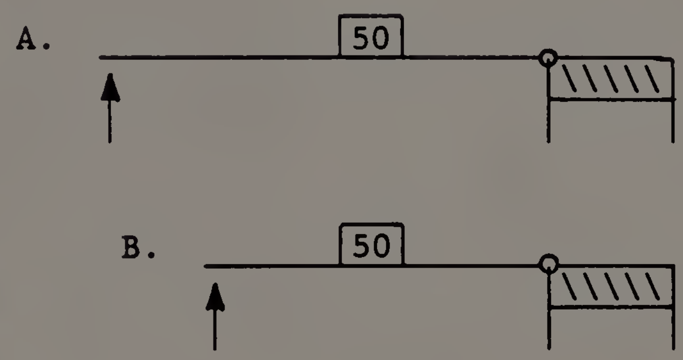
__(a) In case A
__(b) In case B
__(c) You would have to push down with the same force at A and B.

Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

Question 4

Two very light, but strong boards are fixed to the edge of a table with a hinge as shown. Board B is shorter than board A, but the loads are the same distance away from the tables. Where would it be easier to keep the board level by pushing up as shown?



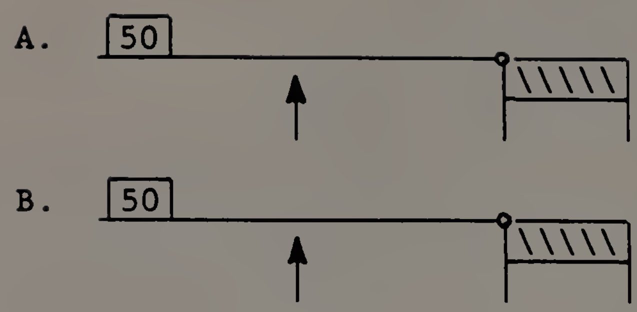
__(a) In case A
__(b) In case B
__(c) The same force would be needed to hold each board level.

Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

Question 5

Two light, strong boards are fastened to tables with hinges as shown in the drawing.



Where would it be easier to hold the board level?

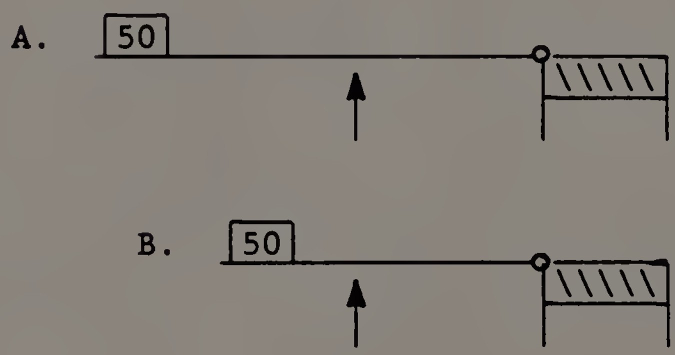
__(a) In case A
__(b) In case B
__(c) The same force would be needed to hold each board level

Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

Question 6

Two light, but strong boards are fastened to tables with hinges as shown.
Board B is shorter than board A, but the person is pushing up on the boards at equal distances from the tables.
Where would it be easier to hold the load on the board level?



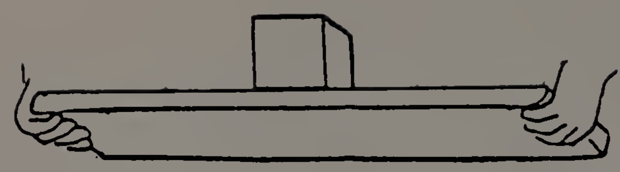
__(a) In case A
__(b) In case B
__(c) The same force would be needed to hold each board level

Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

Question 7

You are holding a 50 lb load on the center of a very light, strong board that rests on your hands as shown.



To hold the load level, each hand has to push up with a force of:

A. About 0 lbs
B. About 25 lbs
C. About 50 lbs
D. About 100 lbs

Confidence Scale

(a) Just a blind guess   (b) Not very confident   (c) Fairly confident   (d) I'm sure I'm right

Question 8

With what force would you have to push down to keep the 50 lb load on the end of the very light, strong board level?

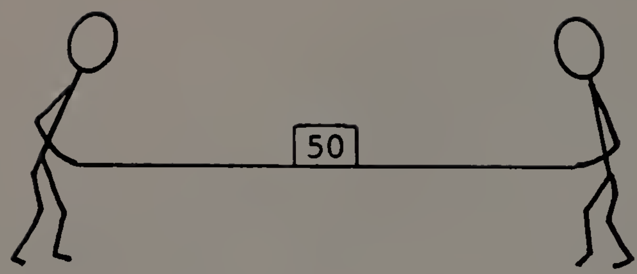
A. About 0 lbs
B. About 25 lbs
C. About 50 lbs
D. About 100 lbs

Confidence Scale

(a) Just a blind guess   (b) Not very confident   (c) Fairly confident   (d) I'm sure I'm right

Question 9

Two people are holding a 50 lb load level on the center of a very light but strong board. With what force does each person have to push up to hold the load level?



A. About 0 lbs
B. About 25 lbs
C. About 50 lbs
D. About 100 lbs

Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

APPENDIX C

PRE- AND POSTTEST

Pretest

Question 1

Where would it be easier to push down to just lift the 50 lb load on the end of the light, sturdy board?

A.

B.

__(a) In case A
__(b) In case B
__(c) You would have to push down with the same force at A and B.

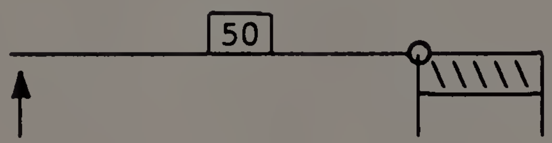
Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

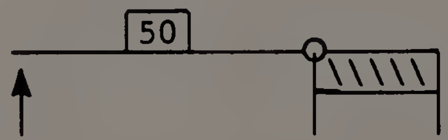
Question 2

Two very light, but strong boards are fixed to the edge of a table with a hinge as shown.
Board B is shorter than board A, but the loads are the same distance away from the tables.
Where would it be easier to keep the board level by pushing up as shown?

A.



B.



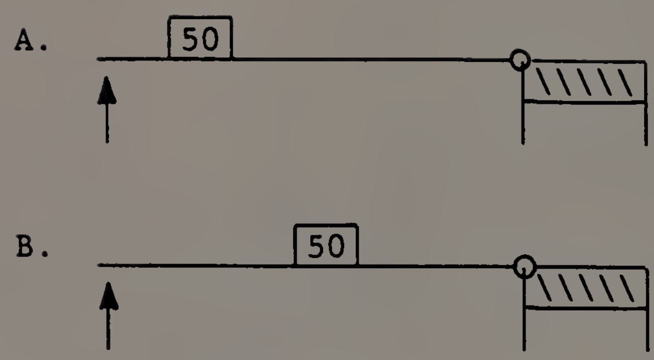
__(a) In case A
__(b) In case B
__(c) The same force would be needed to hold each board level.

Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

Question 3

Two very light but strong boards are fixed with hinges to the edges of tables as shown.



Where would it be easier to hold the board level?

__(a) In case A
__(b) In case B
__(c) The same force would be needed to hold each board level

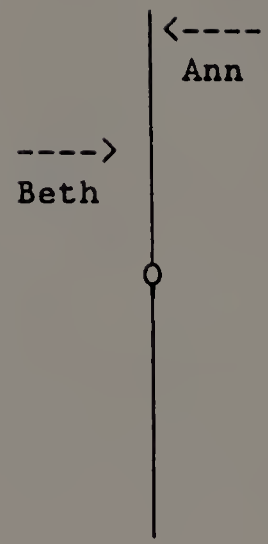
Confidence Scale

(a) Just a blind guess
(b) Not very confident
(c) Fairly confident
(d) I'm sure I'm right

Posttest

Question 1

Beth and Ann are pushing on a rotating door.
Beth is pushing from the outside to go in
Ann is pushing from the inside to go out.
The door is wooden so they can't see each other



Beth and Ann are pushing equally hard.

What do you think will happen?

A. The door will go clockwise

B. The door will go counter-clockwise

C. The door will not move.

Confidence Scale

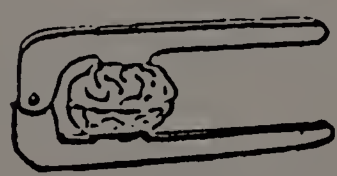
(a) Just a blind guess | (b) Not very confident | (c) Fairly confident | (d) I'm sure I'm right

Question 2

Where will it be easier to crack the nut?

A.

B.



A. In case A

B. In case B

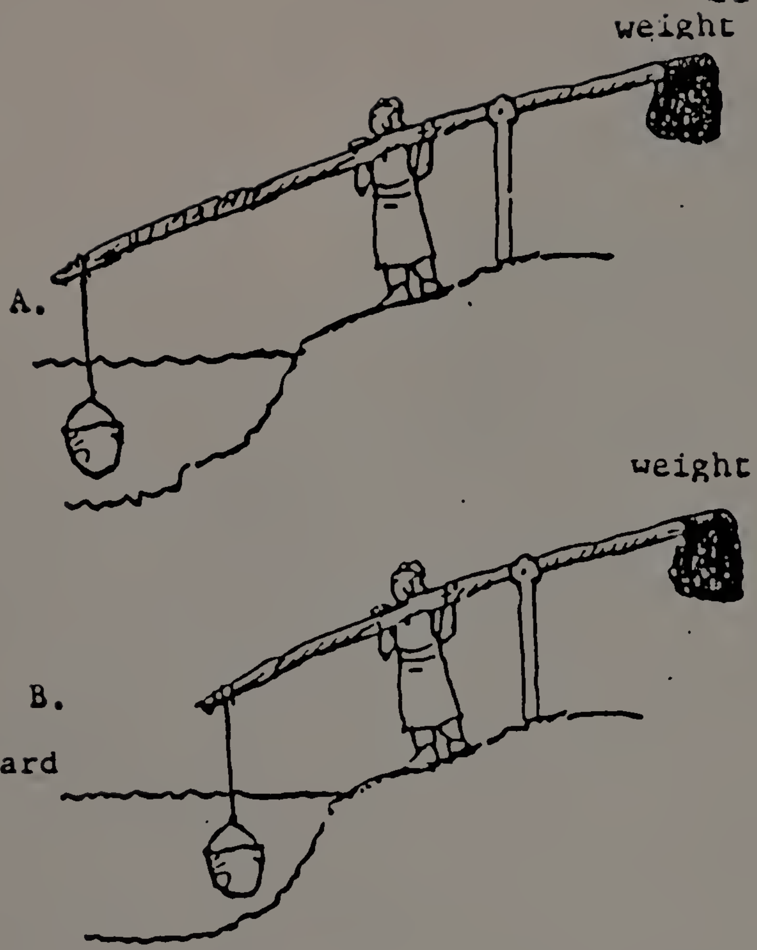
C. You will need the same force to crack the nut at both A and B.

Confidence Scale

(a) Just a blind guess | (b) Not very confident | (c) Fairly confident | (d) I'm sure I'm right

Question 3

The construction in the drawing is called a shadoof and was used to haul water in ancient civilizations.



Which person will find it easier to haul the bucket full of water out?

A. Person A

B. Person B

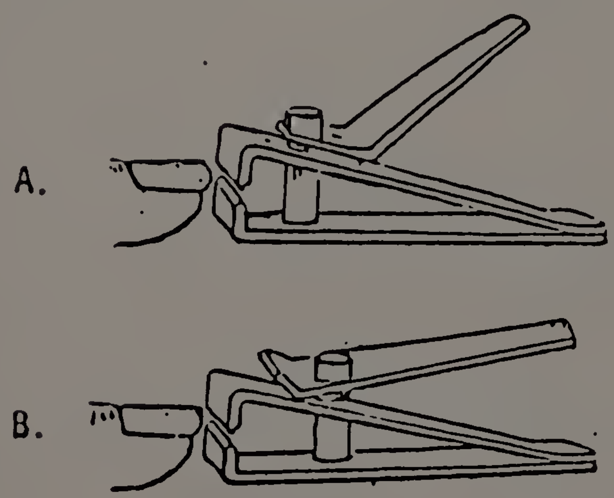
C. They will have to pull equally hard

Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

Question 4

The different parts of a nail clipper are drawn in the picture.



Could you explain which of the two clippers could exert more cutting force (on a tough nail)?

A. A

B. B

C. The same cutting force will be exerted by A and B.

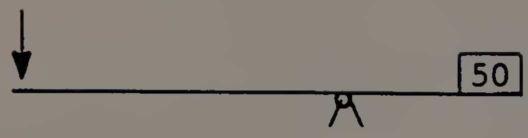
Confidence Scale

(a) Just a blind guess (b) Not very confident (c) Fairly confident (d) I'm sure I'm right

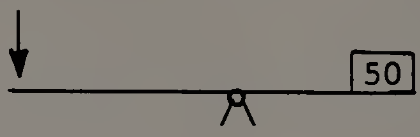
Question 5

Where would it be easier to push down to just lift the 50 lb load on the end of the light, sturdy board?

A.



B.



A. In case A

B. In case B

C. You would have to push down with the same force at A and B.

Confidence Scale

(a) Just a blind guess

(b) Not very confident

(c) Fairly confident

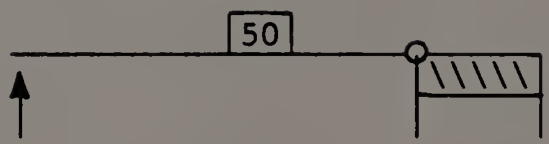
(d) I'm sure I'm right

Question 6

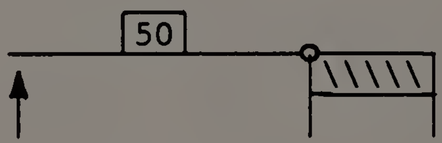
Two very light, but strong boards are fixed to the edge of a table with a hinge as shown.

Board B is shorter than board A, but the loads are the same distance away from the tables. Where would it be easier to keep the board level by pushing up as shown?

A.



B.



A. In case A

B. In case B

C. The same force would be needed to hold each board level.

Confidence Scale

(a) Just a blind guess

(b) Not very confident

(c) Fairly confident

(d) I'm sure I'm right

Question 7

Two very light but strong boards are fixed with hinges to the edges of tables as shown.

Where would it be easier to hold the board level?

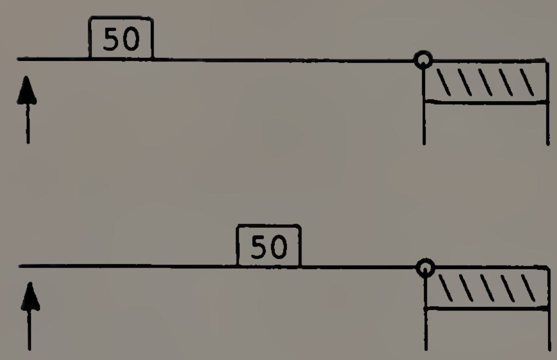
A. In case A

B. In case B

C. The same force would be needed to hold each board level

A.

B.



Confidence Scale

(a) Just a blind guess

(b) Not very confident

(c) Fairly confident

(d) I'm sure I'm right

APPENDIX D

LESSON 1: EXPLANATION AND TRANSFER QUESTIONS

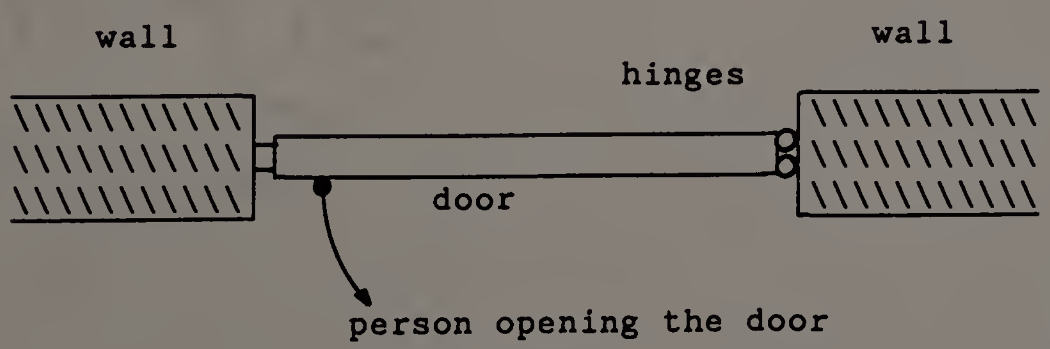
Explanation

The drawings in the problems before are all of very simple levers.
We have to name different parts of the LEVER, to be able to talk about it in the same way:
The block or support in the lever is called the TURNING POINT (or FULCRUM).

A person exerts a FORCE on the one end of the lever.

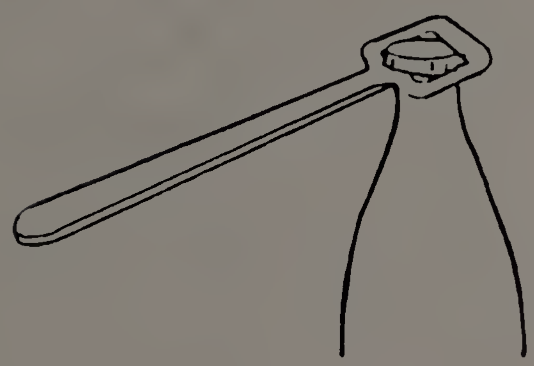
A lever may be used to lift, or crush or grab a certain LOAD.

In the drawing, could you label the: FORCE, LOAD and TURNING POINT?



Near Transfer Questions

With which of the two bottle openers will it be easier to lift the bottle's cap off?



A.

B.

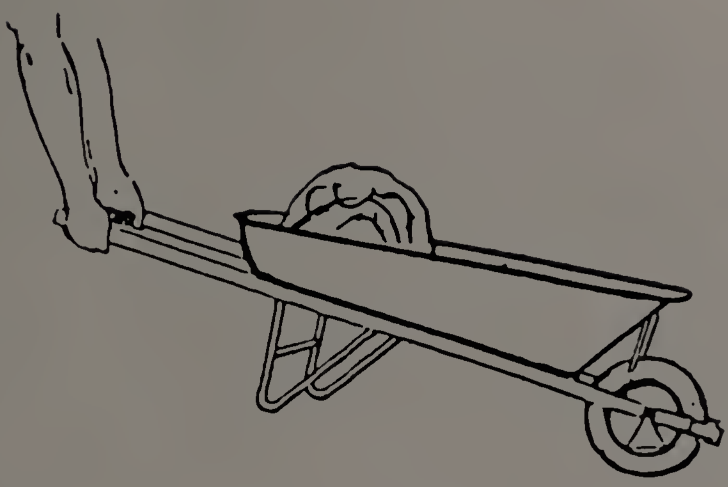
Is the bottle opener a LEVER? __________

Can you show the
(1) TURNING POINT,
(2) LOAD
(3) and where the FORCE is applied?

Confidence Scale

A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

Which of the two wheelbarrows will be easier to hold as shown in the drawing?



A.

B.

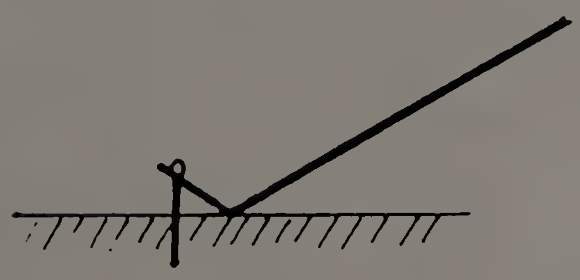
Is the wheelbarrow a LEVER? __________

Can you show the
(1) TURNING POINT,
(2) LOAD
(3) and where the FORCE is applied?

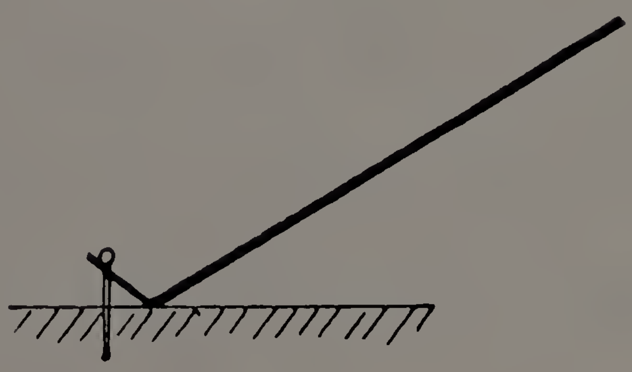
<u>Confidence Scale</u>

A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

In the drawing are two crowbars used as nail extractors.
With which crowbar will it be easier to pull the nail out of the wooden floor?



A.



B.

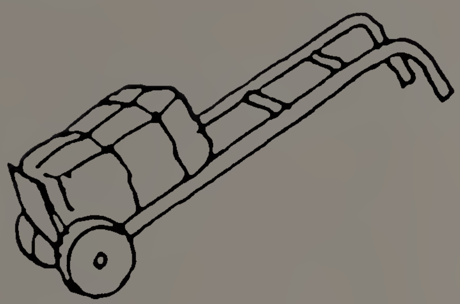
Is a crowbar a LEVER? __________

Can you show the
(1) TURNING POINT,
(2) LOAD
(3) and where the FORCE is applied?

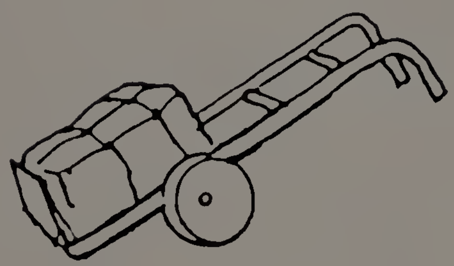
<u>Confidence Scale</u>

A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

In the drawing are two luggage carts. Where would you have to push down with the least force to hold the heavy trunk?



A.



B.

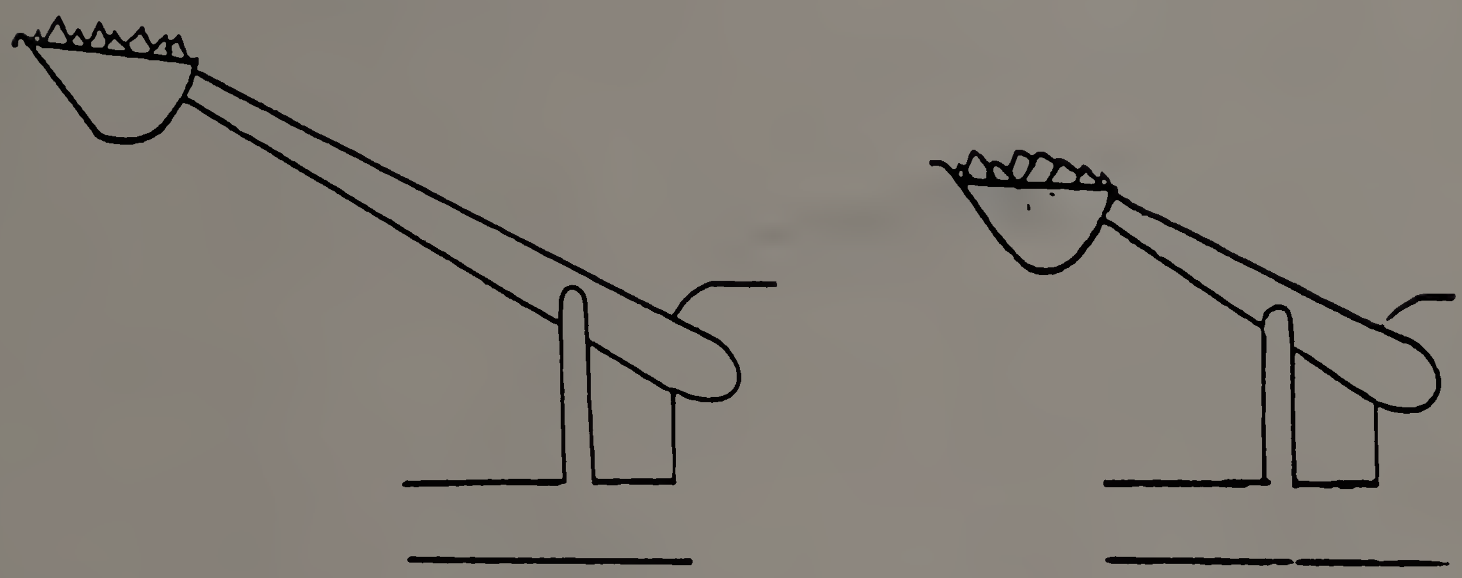
Is a luggage cart a LEVER? __________

Can you show the
(1) TURNING POINT,
(2) LOAD
(3) and where the FORCE is applied?

<u>Confidence Scale</u>

A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

In the drawing are two steam shovel arms, pushed up by pistons below as shown.
Which steam shovel's piston will have to push with the least force to hold the 100 lbs up?



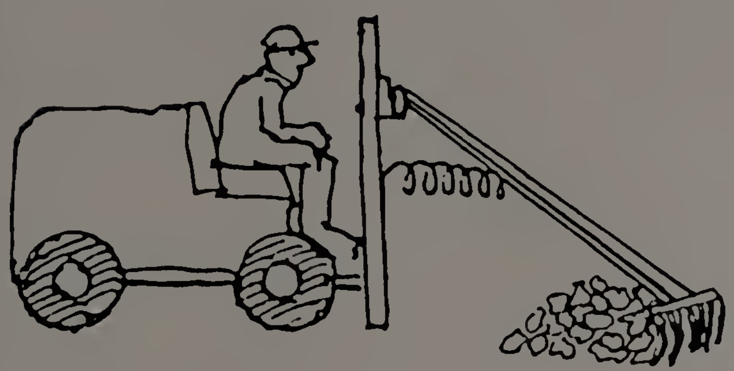
A. B.

Is the steam shovel arm a LEVER? ______

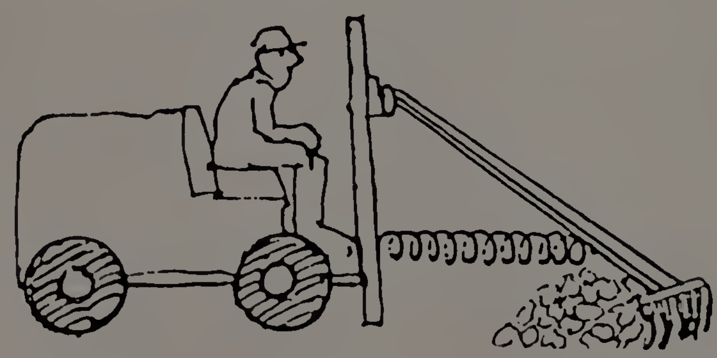
Can you show the
(1) TURNING POINT,
(2) LOAD
(3) and where the FORCE is applied?

<u>Confidence Scale</u>
A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

In the drawing are two mechanical rakes, used to rake wet leaves. The leaves are heavy, and each rake has to pull 150 lbs of leaves toward the truck.
Which rake will have to pull with the least force to gather the leaves?



A.



B.

Are these rakes LEVERS? __________
Can you show the
(1) TURNING POINT,
(2) LOAD
(3) and where the FORCE is applied?

<u>Confidence Scale</u>

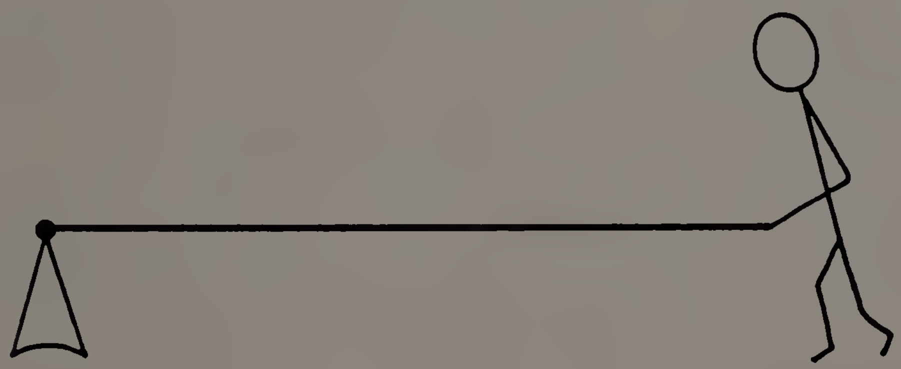
A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

APPENDIX E

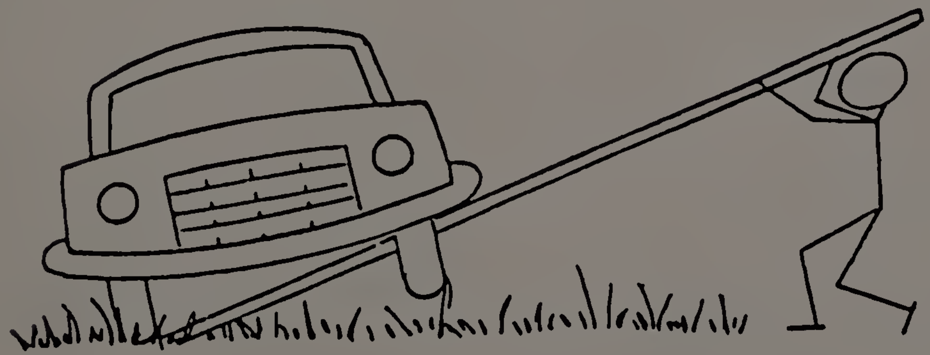
LESSON 2: EXPLANATION AND TRANSFER QUESTIONS

Explanation

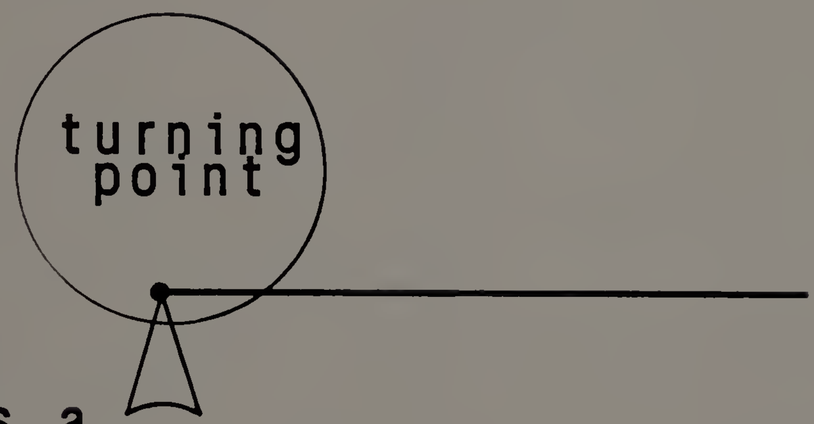
The drawings in the problems before are all of very simple LEVERS. We can say that these levers consist of a board that could TURN around a point when a FORCE is applied to the board, for example:



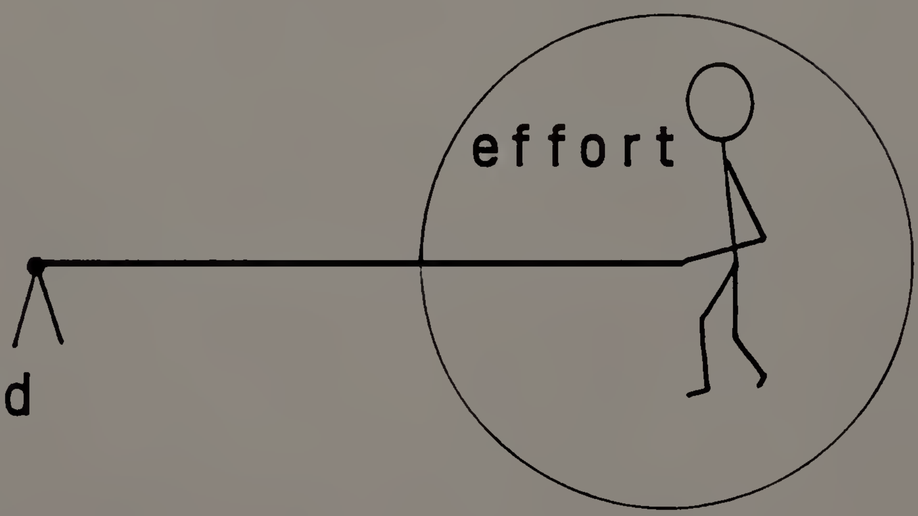
Or, in the same way, a lever is used to help the man lift a car and hold it up for a wheel change.



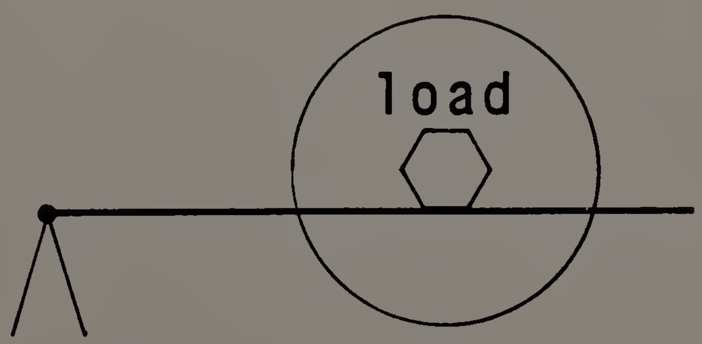
Any LEVER has the following parts:
The point that the lever turns around is called the TURNING POINT (or FULCRUM)



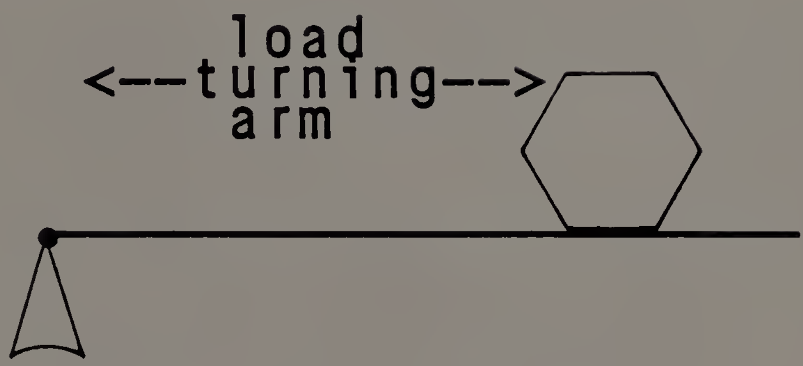
A person exerts a FORCE on one end of the lever, and we call this the EFFORT



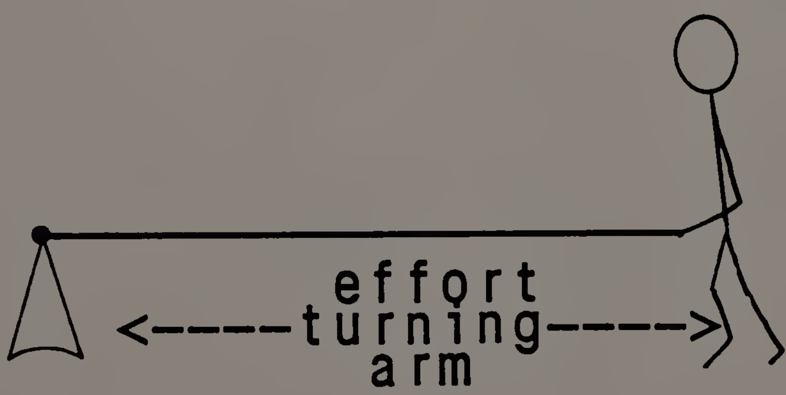
A lever may be used to lift, or crush, or pull or grab a certain LOAD.



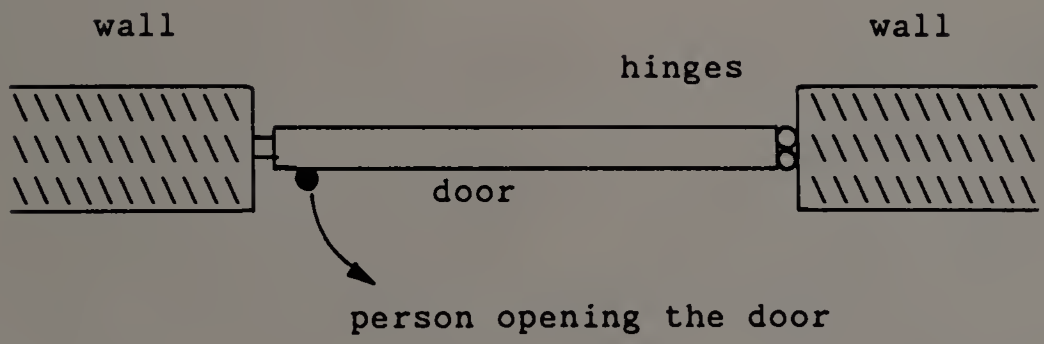
The distance from the LOAD to the TURNING POINT we'll call the LOAD'S TURNING ARM



The distance from the EFFORT to the TURNING POINT we'll call the EFFORT'S TURNING ARM



The person below is seen from ABOVE as she is trying to open the door.



Is the door a lever?_ _ _ _

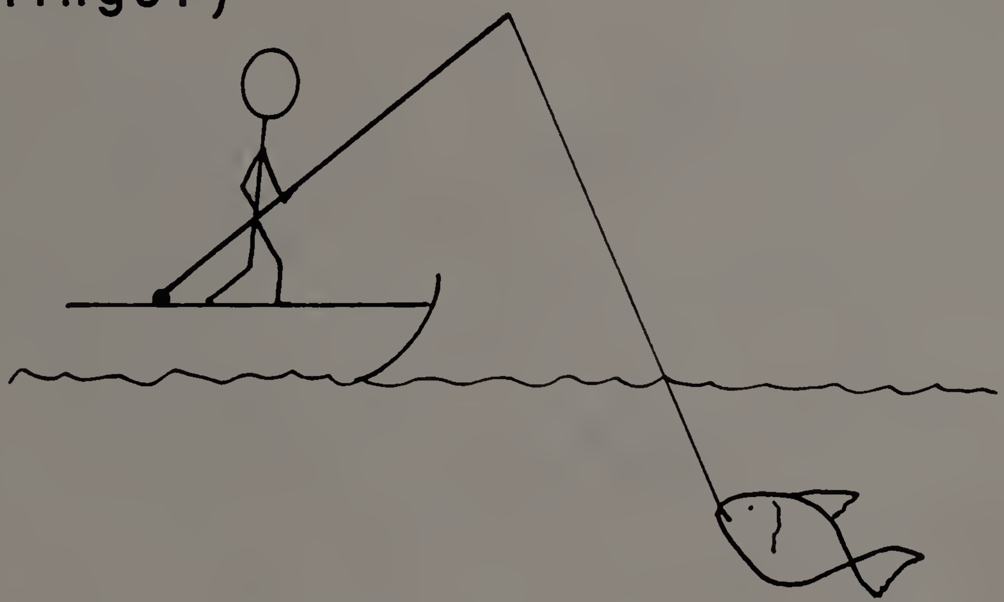
Can you show

(a) the turning point

(b) the load and load turning arm

(c) the effort and effort turning arm.

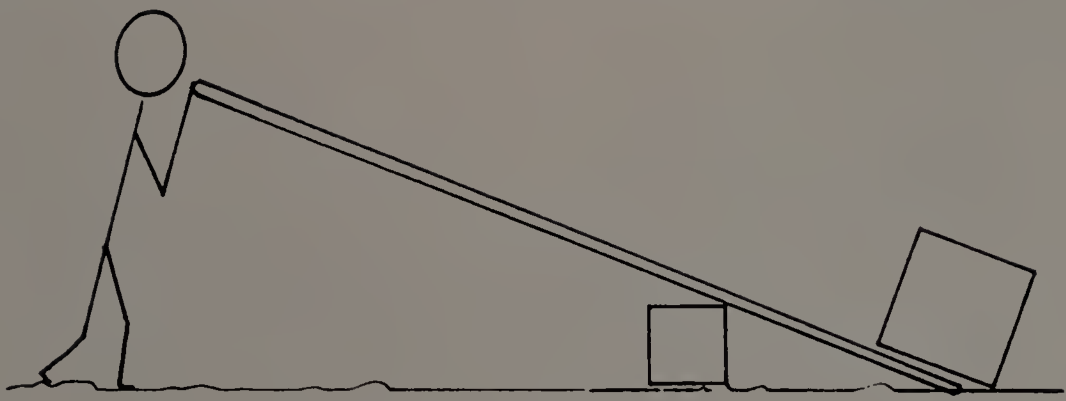
The person below has caught a fish and is using the deepsea fishing rod to pull it out of the water. (The fishing rod is fixed to the floor of the boat with a hinge.)



Is he using a lever?_ _ _ _

Can you show
(a) the turning point
(b) the load and load turning arm
(c) the effort and effort turning arm.

The person below is lifting a heavy block.



Is he using a lever?_ _ _ _

Can you show
(a) the turning point
(b) the load and load turning arm
(c) the effort and effort turning arm.

Near Transfer Questions

With which of the two bottle openers will it be easier to lift the bottle's cap off?

A.

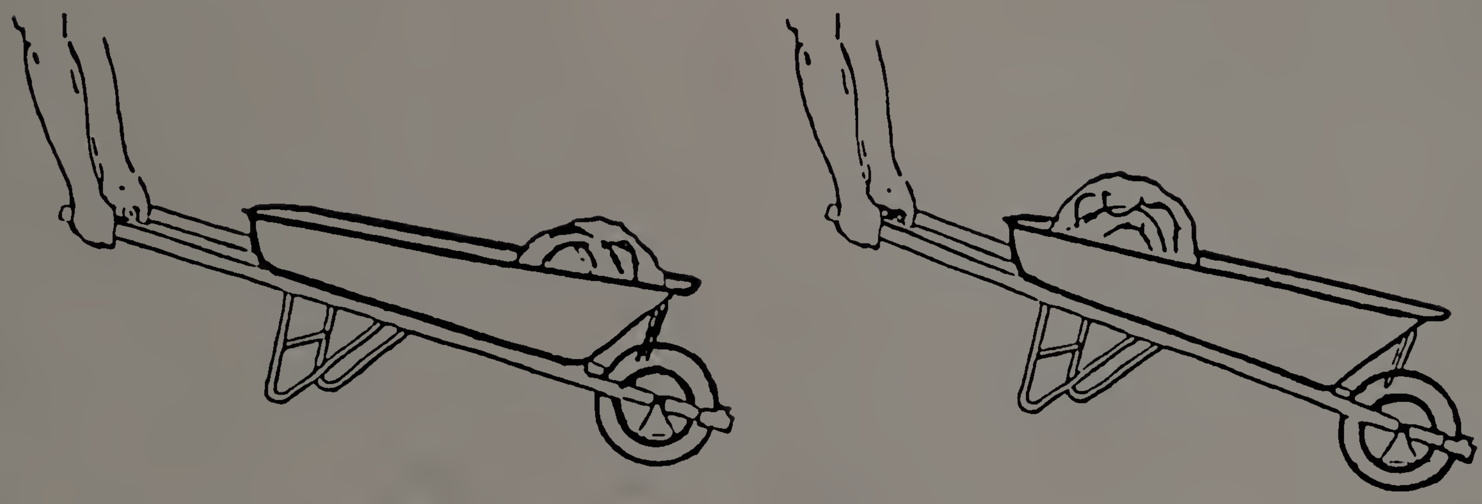
B.

Is the bottle opener a LEVER? __________

Can you show the
(1) TURNING POINT,
(2) LOAD and LOAD TURNING ARM
(3) EFFORT and EFFORT TURNING ARM?

Confidence Scale
A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

Which of the two wheelbarrows will be easier to hold as shown in the drawing?



A. B.

Is the wheelbarrow a LEVER? __________

Can you show the
(1) TURNING POINT,
(2) LOAD and LOAD TURNING ARM
(3) EFFORT and EFFORT TURNING ARM?

<u>Confidence Scale</u>

A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

In the drawing are two crowbars used as nail extractors.
With which crowbar will it be easier to pull the nail out of the wooden floor?

A. B.

Is a crowbar a LEVER? __________

Can you show the
(1) TURNING POINT,
(2) LOAD and LOAD TURNING ARM
(3) EFFORT and EFFORT TURNING ARM?

<u>Confidence Scale</u>
A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

In the drawing are two pairs of pliers. With which pair will it be easier to cut a thick wire?

A. B.

Is a pair of pliers a LEVER? __________

Can you show the
(1) TURNING POINT,
(2) LOAD and LOAD TURNING ARM
(3) EFFORT and EFFORT TURNING ARM?

<u>Confidence Scale</u>
A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

Buck and Chuck are sweeping the floor. They are using exactly the same kind of broom. Who will have to exert the least force to sweep?

A.

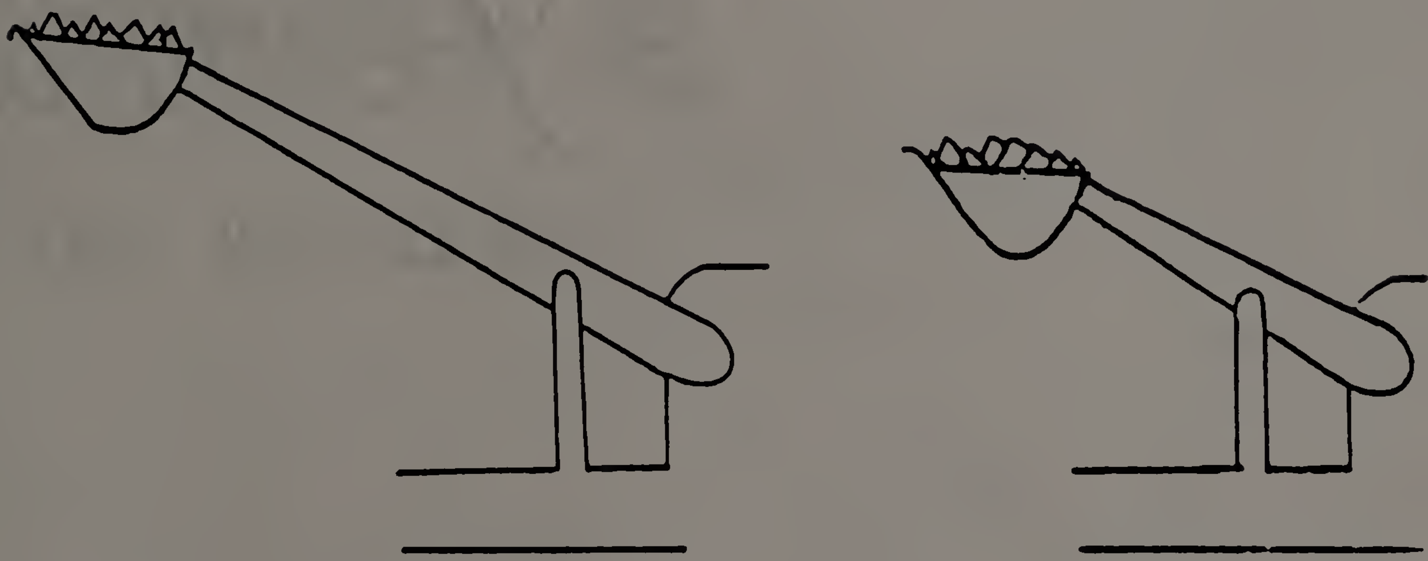
B.

Are the brooms LEVERS? __________

Can you show the
(1) TURNING POINT,
(2) LOAD and LOAD TURNING ARM
(3) EFFORT and EFFORT TURNING ARM?

<u>Confidence Scale</u>
A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

In the drawing are two steam shovel arms, pushed up by pistons below as shown.
Which steam shovel's piston will have to push with the least force to hold the 100 lbs up?



A. B.

Is the steam shovel arm a LEVER? ______

Can you show the
(1) TURNING POINT,
(2) LOAD and LOAD TURNING ARM
(3) EFFORT and EFFORT TURNING ARM?

Confidence Scale
A. I am just guessing
B. I am not very confident
C. I am fairly confident
D. I'm sure I'm right

BIBLIOGRAPHY

Ausubel, D. (1968). Educational psychology: A cognitive view. New York, NY: Holt.

Bassok, M. and Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 153-166.

Brown, A. L. (1987). Strategies in the very young. Paper presented at the annual meeting of the Society for Research in Child Development, Baltimore Maryland.

Brown, A. L. (1990). Domain-specific principles affecting learning and transfer in children. Cognitive Science, 14, 107-133.

Brown, A. L., Kane, M. J. and Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. Cognitive Development, 1, 103-121.

Brown, A. L. and Kane, M. J. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. Cognitive Psychology, 20, 493-523.

Brown, D. (1987). Using analogies and examples to help students overcome misconceptions in physics: A comparison of two teaching strategies. Unpublished Ed.D. dissertation, University of Massachusetts.

Brown, D. (1990). Facilitating conceptual change using mechanistic conceptual models. Paper presented at the annual conference of the American Educational Research Association, Boston MA, April.

Brown, D. and Clement, J. (1987). Overcoming misconceptions in mechanics: A comparison of two example-based teaching strategies. Paper presented at the annual conference of the American Educational Research Association, Washington DC, April.

Brown, D. and Clement, J. (1989). Overcoming misconceptions via analogical reasoning: factors influencing understanding in a teaching experiment. Instructional Science, 18, 237-261.

Carey, S. (1985). Conceptual change in children. Cambridge, MA: MIT Press.

Champagne, A., Klopfer, L. E. and Gunstone, R. F. (1982). Cognitive research and the design of science instruction. Educational Psychologist, 17, 31-53.

Clement, J. (1982). Students' preconceptions in introductory mechanics. American Journal of Physics, 50, 66-71.

Clement, J. (1986). Misconceptions in mechanics and an attempt to remediate them: The use of analogies and anchoring conceptions. Paper presented at a conference sponsored by the National Science Foundation, "The Psychology of Physics Problem Solving: Theory and Practice", Bank Street College, New York.

Clement, J. (1987) with the assistance of Camp, C., Brown, D., Kudukey, J., Minstrell, J., Palmer, D., Schultz, K., Shimabukuro, J., Steinberg, M. and Veneman, V. Overcoming students' misconceptions in physics. In J. Novak and H. Helm (Ed.), Proceedings of the Second International Conference on Misconceptions and Educational Strategies In Science and Mathematics, Ithaca, NY: Cornell University Press.

Clement, J. (1988). Observed methods for generating analogies in scientific problem solving. Cognitive Science, 12, 563-586.

Clement, J. (1989a). Use of physical intuition in expert problem solving. Paper presented at the Workshop on Implicit and Explicit Knowledge, University of Tel-Aviv, October 1988.

Clement, J. (1989b). Learning via model construction and criticism: Protocol evidence on sources of creativity in science. In J. Glover, R. Ronning and C. Reynolds (Eds.), Handbook of Creativity: Assessment, Theory and Research. New York, NY: Plenum.

Clement, J. (in press). Non-formal reasoning in experts and in science students: The use of analogies, extreme cases and physical intuition. In J. Voss, D. Perkins and J. Siegel (Eds.), Informal reasoning and education. Hillsdale, NJ: Lawrence Erlbaum.

Clement, J., Brown, D. and Zietsman, A. (1989). Not all preconceptions are misconceptions: Finding anchoring conceptions for grounding instruction on students' intuitions. International Journal of Science Education, 11, 554-565.

diSessa, A. (1981). Unlearning Aristotelian Physics: A study of knowledge based learning. Cognitive Science, 6, 37-75.

diSessa, A. (1985). Knowledge-in-pieces. Address to the Fifteenth Annual Symposium of the Jean Piaget Society "Constructivism in the Computer Age", Philadelphia, June 1985.

Dreistadt, R. (1968). An analysis of the use of analogies and metaphors in science. The Journal of Psychology, 68, 97-116.

Driver, R. (1983). The pupil as scientist? Milton Keynes: The Open University Press.

Driver, R. (1987). Changing conceptions. Paper presented at the International Seminar: Adolescent Development and School Science, King's College, London, September 1987.

Driver, R. (1989). The construction of scientific knowledge in school classrooms. In R. Millar (Ed.), Doing science: Images of science in science education. London: The Falmer Press.

Driver, R. and 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.

Duit, R. (1981). Students' notions about the energy concept - before and after physics instruction. Paper presented at the International Workshop on problems concerning students' representation of physics and chemistry knowledge, Padagogische Hochschule Ludwigsburg.

Dupin, J. J. and Joshua, S. (1989). Analogies and 'modeling analogies' in teaching: Some examples in electricity. Science Education, 73, 207- 224.

Edwards, D. and Mercer, N. (1987). Common knowledge. New York, NY: Methuen.

Engel, E. and Driver, R. (1982). Children's interpretations of scientific phenomena - analysis of descriptive data. Paper presented at the British Educational Research Association's annual conference, University of St. Andrews, September.

Erickson, G. L. (1979). Children's conceptions of heat and temperature. Science Education, 63, 221-230.

Forbus, K.D. and Gentner, D. (1990). Causal reasoning about quantities. In D. S. Weld and J. de Kleer (Eds.), Qualitative reasoning about physical systems. San Mateo, CA: Morgan Kaufmann Publishers, Inc.

Fredette, N. and Lochhead, J. (1980). Student conceptions of simple circuits. The Physics Teacher, 18, 194-198.

Freer, D. (1982). English in Soweto: A slender lifeline to a wider world. Journal of Education for Teaching, 8(3), 223-241.

Furlong, V. and Edwards, A. (1978). The language of teaching. London: Heinemann.

Gentner, D. and Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner and A. L. Stevens (Eds.), Mental models. Hillsdale, NJ: Erlbaum.

Gick, M. and Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Gick, M. and Holyoak, K. J. (1987). The cognitive basis of knowledge transfer. In S. M. Cormier and J. D. Hagman (Eds.), Transfer of learning. San Diego, CA: Academic Press, Inc.

Gilbert, J. K. and Watts, D. M. (1983). Concepts, misconceptions and alternative conceptions: Changing perspectives in science education. Studies in Science Education, 10, 61-98.

Gilbert, S. W. (1989). An evaluation of the use of analogy, simile and metaphor in science texts. Journal of Research in Science Teaching, 19, 315-327.

Gilot, F. and Lake, C. (1965). Life with Picasso. New York, NY: McGraw-Hill.

Glaser, B. and Strauss, A. (1967). The discovery of grounded theory. Chicago, IL: Aldine.

Goldberg, P. (1983). The intuitive edge. Los Angeles, CA: J.P. Tarcher.

Guidoni, P. (1985). On natural thinking. European Journal of Science Education, 7, 133-140.

Gunstone, R. F. and White, R. T. (1980). A matter of gravity. Paper presented to the Australian Science Education Research Association, Melbourne.

Hamilton, D., Jenkins, D., King, C., MacDonald, B. and Partlett, M. (Eds.), (1977). Beyond the numbers game. London: Macmillan.

Hardiman, P. T. (1983). The importance of analog knowledge in understanding the mean. Unpublished M.S. thesis, University of Massachusetts, Amherst.

Hardiman, P. T., Pollatsek, A. and Well, A. D. (1986). Learning to understand the balance beam. Cognition and Instruction, 3, 63-86.

Hegarty, M., Just, M. A. and Morrison, I. (1988). Mental models of mechanical systems: Individual differences in qualitative and quantitative reasoning. Cognitive Psychology, 20, 191-236.

Helm, H. (1980). Misconceptions in physics amongst South African students. Physics Education, 16, 92-105.

Herscovics, N. and Chalouh, I. (1984). Using literal symbols to represent hidden quantities. Paper presented at the PME-NA meeting, Madison WI.

Hewson, M. G. (1986). The acquisition of scientific knowledge: Analysis and representation of students' conceptions of density. Science Education, 70, 159-170.

Hewson, M. G. and Hewson, P. W. (1983). Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. Journal of Research in Science Teaching, 20, 731-543.

Hewson, P. W. (1984). Microcomputers, conceptual change and the design of science instruction: Examples from kinematics and dynamics. South African Journal of Science, 80, 15-20.

Hewson, P. W. (1985). Epistemological commitments in the learning of science: Examples from dynamics. European Journal of Science Education, 7, 163-172.

Hewson, P. W. and Hewson, M. G. (1988). An appropriate conception of teaching science: A view from studies of science learning. Science Education, 72, 597-614.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. In G. H. Bower (Ed.), The psychology of learning and motivation. Orlando, FA: Academic Press, Inc.

Johansson, B., Marton, F. and Svensson, L. (1985). An approach to describing learning as change between qualitatively different conceptions. In L.T. West and A. L. Pines (Eds.), Cognitive structure and conceptual change. Orlando, FA: Academic Press, Inc.

Kvale, S. (1988). The 1000-page question. Phenomenology + Pedagogy, 6, 90-106.

Locke, L. F., Spirduso, W. W. and Silverman, S. J. (1987). Proposals that work. Newbury Park, CA: Sage Publications.

Macauly, D. (1988). The way things work. Boston, MA: Houghton Mifflin.

Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. Annual of Mathematical Statistics, 18, 50-60.

Mayer, R. E. (1975). Different problem-solving competencies established in learning computer programming with and without meaningful models. Journal of Educational Psychology, 67, 725-734.

Mayer, R. E. (1989). Models for understanding. Review of Educational Research, 59, 43-64.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner and A. L. Stevens (Eds.), Mental models. Hillsdale, NJ: Lawrence Erlbaum.

McCloskey, M., Washburn, A. and Felch, L. (1983). Intuitive physics: The straight-down belief and its origin. Journal of Experimental Psychology: Learning, Memory and Cognition, 9, 636-649.

Minstrell, J. (1982). Explaining the 'at-rest' condition of an object. The Physics Teacher, 20, 10-14.

Muller, K. M. (1986). The practice of school science. Unpublished M.Ed. thesis, University of the Witwatersrand, Johannesburg.

Murray, T., Schultz, K., Brown, D. and Clement, J. (in press). An analogy-based computer tutor for remediating physics misconceptions. Artificial Intelligence and Education.

Nersessian, N. J. (1984). Faraday to Einstein: Constructing meaning in scientific theories. Dordrecht: Martinus Nijhoff.

Nersessian, N. J. (1989). Conceptual change in science and science education. Synthese, 80, 163-183.

Nersessian, N. J. (1990). The logic of discovery. Paper presented at the annual conference of the American Educational Research Association, Boston MA, April.

Novak, G. S. (1977). Representations of knowledge in a program for solving physics problems. International Joint Conference on Artificial Intelligence, 5, 286-291.

Nussbaum, J. and Novick, S. (1982). Alternative frameworks, conceptual conflict and accommodation: Toward a principled teaching strategy. Instructional Science, 11, 183-200.

Oppenheimer, R. (1956). Analogy in science. American Psychologist, 11, 127-135.

Osborne, R. J. (1981). Students' ideas about electrical current. New Zealand Science Teacher, 29, 12-19.

Pfundt, H. and Duit, R. (1985). Students' alternative frameworks and science education. IPN Reports in Brief, Institute for Science Education, University of Kiel, Kiel.

Redhead, M. (1980). Models in physics. British Journal of Philosophy of Science, 31, 145-163.

Royer, J. M. (1979). Theories of the transfer of learning. Educational Psychologist, 14, 53-69.

Royer, J.M. (1987). Designing instruction to produce understanding: An approach based on cognitive theory. In S. M. Cormier and J. Hagman (Eds.), Transfer of learning. San Diego, CA: Academic Press, Inc.

Royer, J. M. and Cable, G. W. (1976). Illustrations, analogies and facilitative transfer in prose learning. Journal of Educational Psychology, 68, 205-209.

Siegler, R. S. (1978). The origins of scientific reasoning. In R. S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, NJ: Lawrence Erlbaum.

Siegler, R. S. (1982). The development of intelligence. In R. J.Sternberg (Ed.), Handbook of human intelligence. Cambridge: Cambridge University Press.

Smith, C., Snir, J., Unger, C. and Grosslight, L. (1990). Facilitating conceptual differentiations using conceptual models. Paper presented at the annual conference of the American Educational Research Association, Boston MA, April.

Solomon, J. (1980). How children learn about energy, or does the first law come first? School Science Review, 63, 415-422.

Stenhouse, L. (1975). An introduction to curriculum research and development. London: Heinemann.

Sutherland, S. (1989). The international dictionary of psychology. New York, NY: Continuum.

Trowbridge, D. E. and McDermott, L. C. (1980). An investigation of students' understanding of the concept velocity in one dimension. American Journal of Physics, 48, 1020-1028.

Trowbridge, D. E. and McDermott, L. C. (1981). An investigation of students' understanding of the concept acceleration in one dimension. American Journal of Physics, 49, 242.

Viennot, L. (1979). Spontaneous reasoning in elementary dynamics. European Journal of Science Education, 1, 205-221.

von Glasersfeld, E. (1989). Cognition, construction of knowledge and teaching. Synthese, 80, 121-140.

Watts, D. M. (1983). A study of schoolchildren's alternative frameworks of the concept of force. European Journal of Science Education, 5, 217-230.

White, B. Y. (1983). Sources of difficulty in understanding Newtonian dynamics. Cognitive Science, 7, 41-45.

White, B. Y. (1990). Criteria for successful intermediate models: Three models. Paper presented at the annual conference of the American Educational Research Association, Boston MA, April.

Wilcoxon, F. and Wilcox, R. (1964). Some rapid approximate statistical procedures. Pearl River, NY: Lederle Laboratories.

Woods, D., Bruner, J. S. and Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17, 89- 100.

Yukawa, H. (1973). Creativity and intuition. New York, NY: Kodansha.

Zietsman, A. I. and Hewson, P. W. (1986). Effect of instruction using microcomputer simulations and conceptual change strategies on science learning. Journal of Research in Science Teaching, 23, 27- 39.