CHAPTER TWO

Learning Conceptions and Outcomes

LARS-OWE DAHLGREN

The Quantitative Conception of Knowledge

One of the longest-running programmes on Swedish television is a series called *Double or Quits*, which was modelled on similar quiz programmes in the United States and Britain. Below are some examples from one of the programmes:

> Which nations were involved in the battle of Lizza in 1866? (from Maritime History).
>
> In Chopin's manuscript of the *Preludes* and in the original German edition, he dedicates them to Joseph Kessler. The first French and English editions, however, are dedicated to another contemporary of Chopin's. Whom? (from Chopin and his Music).

Questions of this kind are typical of those put to laymen or experts on many radio or television programmes. Since those who take part in the Double or Quits programmes are experts in their chosen subjects, however, the questions asked are ones which the man in the street could not be expected to answer. Yet regardless of their level of difficulty, the questions are all similar in structure in that they demand a brief answer which takes the form of the name of a person or a place, a year when something occurred, a technical term, and so on. Seldom if ever are there questions asking, for example, why something happened.

Conceptions of knowledge form a very important component of what we call the cultural basis of a society. In its purest and most tangible form knowledge is observable in the educational system. The point in presenting the excerpts from the TV-programme, however, was to illustrate that the dominating conception of knowledge is also visible elsewhere. We find signs of an identical conception in informal discussions with adults who lack personal experience of upper-secondary or higher education. When asked, for instance, what they think university students of mathematics or history are engaged in, some will answer in a way that may make professional mathematicians or historians smile, but which nevertheless reveals a conception that is probably very widespread among people in general. Thus it is not uncommon for people to imagine that university students of mathematics are working on immensely difficult calculations, that they are subtracting or multiplying enormously large numbers or unbelievably small fractions. Students of history or professional historians are likewise described as persons who know "a hell of a lot of history", that is, they know not only the year of an important historical event, but also the precise date. Further, a sophisticated historian also has to know not only the prominent historical figures, but also their relatives and the year, date and place of their birth.

As well as appearing ingenuous, these answers tell us something about the way experience of schooling influences our way of apprehending knowledge. It is, however, also self-evident that if one lacks any insight into the qualitative change that the content of studies undergoes at more advanced levels one makes a linear—and horizontal—extrapolation from what is known into the unknown.

A comparison between the questions put in the *Double or Quits* programme and the answers given by people with only a basic education reveals that they are strongly related to each other. None of them indicates a qualitative change in knowledge from trivial to advanced levels. The *Double or Quits* questions are basically of the same kind at the beginning and at the end of the game even though they have become progressively more difficult. The difference that can be observed is that the questions become more and more peripheral to the phenomenon in question, e.g. knowing the name of a person to whom Chopin dedicated a particular composition must be regarded as being of minor interest compared to understanding the structure of the music.

Difficult questions in these contexts are also narrower than "easy" ones, in that they deal with very specific details of an event or minor parts of a phenomenon. This difference between what is trivial and what is advanced is to a great extent preserved when we move into the world of the educational system.

The measurement of knowledge has as long a history as the educational system as a whole. Over the years a number of ways of approaching this problem have been tried, involving both the more technical aspects of educational measurement as well as attempts at more thoroughgoing re-evaluations. Yet if we compare the present state of the art with the past, irrespective of what level of the educational system we refer to, none of the basic characteristics of test items has changed in any dramatic way. There are also very obvious parallels between the demands put on students and on contestants in quiz programmes. These are probably at their most visible in questions representing so-called objective tests, which came into frequent use from the early sixties onwards. Some examples taken from various subjects are given below:

(A) The capital of Albania is:
1. Belgrade
2. Tirana
3. Lisbon
4. Lagos

(B) Relate the following South America countries to the product which is their most important export:

1. Venezuela	a. Copper
2. Chile	b. Coffee
3. Brazil	c. Oil

(C) Complete the sentence below by filling in the missing information:
The Swedish King Gustav II Adolf was killed in the battle of ..... in a long war between Sweden and ..... which ended with the ..... peace treaty in the year .....

Many teachers will probably recognize their own way of constructing examination test items in these examples. They will also be aware of the reasons why questions are presented in that form, and to a large extent these reasons are simply pragmatic. Test items should be easy to construct, to answer, and to mark.

A less obvious reason for this form of question is that it is symptomatic of a *conception of knowledge* which has a long tradition in education as well as in quiz programmes. This conception, which was introduced in Chapter 1, can be characterized as *quantitative* and *reproductive*. The degree of difficulty sought is achieved by formulating questions which refer to low-frequency, peripheral and narrow information. Generally speaking, neither understanding nor analytic ability is required of the respondent. That would create problems of judgement for the teacher or the compere of the quiz programme. It is much easier if answers are recognisably right or wrong.

This widely held and culturally deep-rooted view of knowledge is found in the study of Perry (1970) which was described in Chapter 1. Perry found freshmen students generally to have a *dualistic* conception of knowledge indicated by the expectation that higher education would provide an opportunity to learn to discriminate between true and false, between right and wrong. Many of the students had later abandoned this conception in favour of a *relativistic* one. The students had recognised that, to a large extent, phenomena are described and explained in different ways even in academic textbooks or by different teachers. The solution to this pluralistic world of competing explanation lies in a *personal commitment* whereby students take individually distinctive interpretative stances in deciding how to make sense of central phenomena in their field of study.

Our earlier discussion about the quantitative conception of knowledge suggests that the lower stages of the educational system may be, to a large extent, responsible for reinforcing a dualistic conception of learning.

Traditional Psychological Experiments

As we have already seen in Chapter 1, experiments in the psychology of learning have relied extensively on learning materials which have a low degree of meaningfulness. Since the underlying aim has been to arrive at a description of the process of learning in general, this choice has been justified on methodological as well as theoretical grounds. Hence there is seldom any description of the outcome of learning other than in purely quantitative terms, thus reinforcing that reproductive conception of knowledge. We can see this clearly in Hilgard and Bower's *Theories of Learning*, an authoritative textbook which was first published in 1948 and had reached its fifth edition by 1981. A careful examination of the subject index yields few references to *knowledge*. The most significant entry directs us to the following passage:

> A strong emphasis within Gagné's analysis is upon the structure of knowledge, an important supplement to principles of learning whenever a practical instructional task is under consideration.

What is explicitly stressed here is the process of learning. Precisely what the subjects are asked to learn is seen as a problem to be considered elsewhere within the separate domain of instruction.

There is an additional reason why so little is said about the outcome of learning in most literature in the field. In accordance with the research tradition which evolved in the natural sciences, it has become the dominant paradigm of the social sciences to reduce the descriptions of complex phenomena to a minimum number of dimensions. 'Intelligence' or 'learning capacity' is one such dimension that is considered to be of great importance in describing human functioning. For reasons primarily of experimental design, however, such a dimension has to be content neutral, which means that the content of a learning task has the status of a series of examples which are of little interest in themselves.

Against the background of this view of learning it is also easier to understand why certain materials came to be widely used in empirical studies of learning. Nonsense syllables or, more recently, narrative or descriptive texts specially written for the experiments are essentially homogeneous. Each segment of the material is of equal value, and so the likelihood that any one segment will be recalled in a subsequent retention test is no greater than that of any other. Take the following example, taken from Thorndike (1977):

> Circle Island is located in the middle of the Atlantic Ocean, north of Ronald Island. The main occupations on the island are farming and ranching. Circle Island has good soil, but few rivers and hence a shortage of water. The island is run democratically. All issues are decided by a majority vote of the islanders. The governing body is a senate, whose job is to carry out the will of the majority. Recently, an island scientist discovered a cheap method of converting salt water into fresh water. As a result, the island farmers wanted to build a canal across the island, so that they could use water from the canal to cultivate the island's central region. Therefore, the farmers formed a "Pro-canal Association" and persuaded a few senators to join. The Pro-canal Association put the construction idea to the vote. All the islanders voted. The majority voted in favour of construction. The senate, however, decided that the farmers' proposed canal was ecologically unsound. The senators agreed to build a smaller canal that was two feet wide and one foot deep. After starting construction on the smaller canal, the islanders discovered that no water would flow into it. Thus the project was abandoned. The farmers were angry because of the failure of the canal project. Civil war appeared inevitable.

The performance of a subject in a learning experiment using this text would be judged in terms of the sum of the various questions which could be derived from the text, such as: Where is Circle Island situated? What are the main occupations? Why was the canal built? and so on. An alternative way of testing retention would be to ask students to recount the story and mark the number of correct statements included. In both cases the result is a measure of the degree to which the precise wording of the text is remembered. Thus the degree of isomorphy between the stimulus (the text) and the response (its retention) has been the chief interest of learning researchers.

If a similar text were to be used in an educational setting, the measurement of the learning outcome would probably be of the same kind. Even if a task such as "Write a short essay about Circle Island" were assigned, the judgement would probably be based on a scrutiny of how many items from the text had been included in the essay. Consider, however, the following excerpts from an undergraduate textbook (Samuelson, 1973 p. 14):

> If all farmers work hard and nature co-operates in producing a bumper crop, total farm income may fall, and probably will.
>
> Attempts by individuals to save more during a depression may lessen the total of the community's saving.

These two sentences are taken from one of the most widely used university textbooks in Economics. If a group of students were asked to explain why the sentences are correct, even though they appear to be false, the probability of a correct answer would be highly related to whether the students had understood the principles of Economics that could be applied. It is, however, still the case that a typical test question based on these statements would be of the form "Name the principle in Samuelson's first chapter which is exemplified in these two statements". A question of that kind would not enable a teacher to judge which students had really understood the meaning of the examples.

A Qualitative Conception of Learning

It is obvious from this comparison of different kinds of texts – and of different purposes in reading a text – that prose learning is not an homogeneous phenomenon. Psychological research, in its attempt to investigate learning processes in a 'pure' form, has restricted its definition of learning. By using materials with little or no inherent meaning, such experiments describe and explain only how students set about learning when the task has been drained of meaning. Yet most human learning depends on meaning and it is directed towards it. To learn is to strive for meaning, and to have learned something is to have grasped its meaning. In spite of this dominant interest in learning defined as a quantitative phenomenon, since the time of Bartlett (1932) there has also been a concern with learning defined in qualitative terms. Bartlett investigated the ways in which students recounted a story they had read. The differences in the form of these responses led Bartlett to abandon the conception of memory as a reproductive storage mechanism, where every impression with all its specific characteristics is stored in a defined, neural region. Instead Bartlett's conception of the memory depends on the reconstruction of meaning in terms of *schemata* which represent personal reinterpretations of the learning material.

The qualitative approach to research on learning which is reported in this book represents a development of Bartlett's conception. It rejects the description of knowledge as discrete pieces of knowledge passed passively from teacher to learner, and tested in terms of whether or not the student can reproduce verbatim those elements. Instead of concerning itself with "how much is learned", it seeks to investigate "what is learned". Necessarily this qualitative type of research is concerned with the learning of realistically complex passages which contain a

description or an explanation of a phenomenon. If students are given such a text and asked to read it carefully in order to be able to answer questions about its content, it is possible to investigate "what is learned" in a naturalistic setting – an experimental situation in which both content and instructions are closely similar to what students normally experience in higher education.

The next step in the research process depends on generating data about how the subjects have understood the content of the text. The need for intensive and deep information places limitations on the choice of methods. The general research strategy has been to use semi-structured or thematic interviews which are tape-recorded. Identical introductory questions on each topic are followed by questions aimed at eliciting answers in more depth. Depending on the structure and comprehensiveness of an initial answer the interviewer may have to ask for clarification, elaboration or examples. The interviewer must, however, avoid giving any clues about the desired direction which the process should lead. The tape recordings are then typed up and the resulting protocols – once they have been checked by the researcher – constitute the data on which analysis is carried out.

The aim of the analysis is to yield *descriptive categories* of the qualitative variation found in the empirical data. The process involves the reduction of unimportant dissimilarities e.g. terminology or other superficial characteristics, and the integration and generalisation of important similarities i.e. a specification of the core elements which make up the content and structure of a given category. Some examples of this kind of analysis will be presented below.

Many of the studies carried out in Gothenburg during the first half of the 1970s took the form of text reading experiments. Thus in one investigation (Marton, 1975b: Marton *et al*., 1977) forty students of education were asked to read an article from a Swedish newspaper. The article (which was written by Urban Dahllöf, a Swedish professor of education) was a contribution to a debate about a reform in the Swedish system of higher education. The article can be summarised as follows.

By re-analysing the empirical data used in an investigation initiated by the National Board of Universities and Colleges, Dahllöf arrives at a conclusion which differs from that drawn in the original study. In that study the pass rate of students was found to be very low in the faculties of liberal arts and social science. The pass rate was however considerably higher in more vocationally oriented fields such as medicine, civil engineering, etc. It was therefore concluded that the pass rate could be improved if a number of fixed combinations of subject areas was introduced, in order to make schemes of study in the "free" faculties similar to their more vocational counterparts.

In his re-analysis of the data, Dahllöf makes the assumption that many students who enter the system of higher education do so without the intention of graduating, but only to study a particular subject over a number of terms. Dahllöf excludes from the empirical material students older than twenty-five on the assumption that, at that age, they have probably already gone through some kind of post-secondary education and want to complete that education with a few terms of university studies. Although this group of students are officially defined as drop-outs, that definition does not match their own intentions. Further more, Dahllöf

splits the data into sub-groups according to university, sex, subject area, and grade point average from upper secondary school. He thus finds that there are large differences between the different sub-groups. Some have a very low pass rate and some have a pass rate which is similar to that found in the medical or engineering faculties. Dahllöf draws the conclusion that if the purpose of the reform is to raise the pass rate in the faculties of humanities and social science, selective rather than general measures should be taken. The grounds on which he therefore challenges the wisdom of the reform are that a closer look at the empirical data shows that the situation is satisfactory as far as many groups of students are concerned, and very problematic in the case of others.

In the learning experiment students were invited, individually, to read Dahllöf's article carefully at their own pace. They were asked to read it in their usual way, but they were told that they would be asked questions about it afterwards. They were then interviewed and asked questions initially about the general meaning of the article – "Try to summarise the article in one or two sentences. In other words what is the author's intention?" Other questions related to specific aspects of the article and to the processes of learning. Here we are concerned only with the analysis of the extent to which the main point of the article could be recounted.

By applying a rigorous qualitative analysis, the students' responses can be grouped into a number of categories, according to the basic underlying structure expressed. This means that the protocols have to be studied with the intention of understanding what the students are expressing, irrespective of what words or examples they may use, which may show a considerable variation even between answers belonging to the same category. Starting with a comparatively large number of categories the researcher will gradually refine these, arriving at a smaller set of categories that may finally be difficult or impossible to collapse further. In the case of the Dahllöf article, the empirical analysis of students' answers yielded four categories of outcome:

A. Selective measures should be taken.
B. Differential measures should be taken.
C. Measures should be taken.
D. There are differences between different groups of students.

What then differentiates these categories one from another? Clearly there is a hierarchical relationship between A, B and C with regard to their degree of specificity, in that selective measures (A) are a special case of differential measures (B) while the same relation is applicable also for B in relation to C. Category D, on the other hand deviates from the others by expressing only an aspect of the empirical data. Categories A and B both involve the use of evidence in support of *conclusions*, while categories C and D represent *descriptions*. The C-answers may appear conclusion-oriented but the very general conclusion that 'measures should be taken' is not rooted in the empirical data, but is rather a kind of addition to the reported main point about the differences in pass rates. In other studies (Dahlgren, 1975; Säljö, 1975; Marton, 1976a; Svensson, 1976) categories of outcome have been reported which occupy a level below that of description. Instead, there is a reliance merely on *mentioning* elements remembered from the text.

The range of categories of response found in this study (and in other similar investigations) can be described as the *outcome space* for the text concerned. The outcome space provides a kind of analytic map of variations in what has been learned from a given learning task. It is therefore an empirical concept which is not the product of logical or deductive analysis, but instead results from intensive examination of empirical data. Equally important, as used here, the outcome space is content-specific: the set of descriptive categories arrived at has not been determined *a priori*, but depends on the specific content of the learning material. Indeed it should be stressed that, as Entwistle (1976) has observed, "the verb to learn takes the accusative (case)." There is no learning without a content, and thus no phenomenon of learning *per se*.

Structural Aspects of Outcomes of Learning

This does not mean that differences in outcome are wholly content-based. Although the categories which summarise each level of outcome may also preserve (as in the case of the Dahllöf article) a description of the content, more general structural differences can frequently be identified. For example, as we have just seen, outcomes can be categorised as conclusion oriented, descriptive or mentioning, and such differences can also be said to represent distinct levels of outcome. Similarly sets of outcome categories can sometimes be shown to represent hierarchies, where outcomes are related one to another in terms of their degree of specificity, inclusiveness or completeness.

Our next two examples are both to varying extents concerned with the structural properties of differences in outcome. The first is a study by Wenestam (1980). Like many of the Gothenburg studies, it is a text-related analysis of the content of learning. Instead of making use of a single text, however, Wenestam selected four texts which share a common structure; each describes a particular principle which is then illustrated by an example. The texts vary in length from two to six pages, but in each case, the account of the example takes up a substantial proportion of the passage.

One of Wenestam's texts dealt with the scientific work of the physician Ignaz Semmelweis, who is the discoverer of micro-organisms as the origin of the epidemic diseases. Semmelweis's discovery and the thought and experimental work that proceeded from it, is used as an example of the scientific way of hypothesis testing by means of the experimental method. One of Wenestam's questions was:

> Try to summarise the text in a few sentences. In other words, what did the author want to say?

Four categories of answers to this question were identified:

A. The main point of the text (the testing of hypothesis by comparing two conditions where only one factor, the assumed cause, differs) and its relation to the example (the work of Dr. Semmelweis and the mode of action in his investigations) has been understood.
B. The main point of the text has been understood but not its relation to the example.

C. The main point of the text has not been understood but some other main point has been described in a rather general way (e.g. it is about the causes of a phenomenon: or a method for the solution of a problem).
D. The focus is on one or more of the concrete examples (e.g. it describes a doctor at a hospital in Vienna who worked to find the cause of the high mortality rate in childbed fever among women in labour; or the necessity of maintaining a high standard of hygiene).

This result, that is the gradual weakening of the importance of the principle and the successive upgrading of the example from categories A–D, was an important aspect of the variation in the answers on all four texts used by Wenestam. It exemplifies a tendency which we have termed *horizontalisation.* In texts such as these, the intention is to convey a principle. Examples have a subordinate function, which is to illustrate the principle outlined. In horizontalisation, however, this hierarchy is not preserved; no distinction is made between the status of the principle and the status of the example.

It seems probable that horizontalisation is not confined to learning involving a specific kind of text, but is to be found more widely in formal education. Teachers undoubtedly both hope and believe that the examples or metaphors they use to illuminate a given principle will prove less enduring than the principle itself, but how often this actually occurs is open to doubt. The striking concrete example may turn out to be more memorable than the imperfectly understood abstract principle it was meant to illustrate. Yet though horizontalisation seems to reflect a structural difference of a given kind in the quality of learning outcomes, it should be stressed once again that such differences have to be looked for in relation to specific content and depend, moreover, on empirical analyses of outcomes. There have been attempts to establish general taxonomies (Bloom, 1956; Gagné, 1977) through which the content of different learning tasks can be analysed, but such taxonomies are of little relevance here, for they represent logical analyses of the content and processes of learning. They do not derive from studies of the different outcomes arrived at for a given subject-matter.

In this respect, an Australian study by Biggs and Collis (1982) is an evident exception. Their SOLO taxonomy (in which SOLO is an acronym for the Structure of the Observed Learning Outcome) is an attempt at empirical classification of levels of outcome in a form which has wide applicability. The theoretical basis of Biggs and Collis' taxonomy derives in part from the stages in cognitive development described by Piaget and in part from theories of information processing. Breaking away from Piaget's use of stages to describe the developmental level of an individual child, Biggs and Collis seek to describe the range of answers given to a specific question—in our terms the 'outcome space'. They assume that such levels have a general reality, irrespective of content and question form, and describe five categories as follows, with increasing levels of sophistication.

1. *Pre-structural.* In relationship to the prerequisites given in the question, the answers are denying, tautological, and transductive – bound to specifics.

2. *Uni-structural.* The answers contain "generalisations" only in terms of one aspect.
3. *Multi-structural.* The answers reveal generalisations only in terms of a few limited and independent aspects.
4. *Relational.* Characterised by induction, and generalisations within a given or experienced context using related aspects.
5. *Extended abstract.* Deduction and induction. Generalisations to situations not experienced or given in the prerequisites of a question.

Biggs and Collis provide several examples from different school subjects showing how the SOLO taxonomy may be applied in analyses of learning outcomes. One of these examples is an analysis of the answers given to a question which asked why the side of a mountain that faces the coast is usually wetter than the side facing the interior. The following responses illustrate the five categories described above.

1. Because it rains more on the coastal side.
2. Because the sea breezes hit the coastal side first.
3. Because the sea breezes contain water vapour and they first strike the coastal side and so it rains on them and after that there's no rain to fall on the other side.
4. Because the prevailing winds are from the sea and they pick up moisture and as they meet the mountain they're forced up and get colder, the moisture condenses, forming rain. By the time the winds cross the mountain they are dry.
5. This is likely to be true only if the prevailing winds are from the sea. When this is so, the water vapour evaporated from the sea is carried to the mountain slopes, where it rises and cools. Cooling causes the water vapour to condense and deposit. Not only is the wind now dryer, it is then carried up the mountain further, is compressed, now warm, and thus is relatively less saturated than before: the effect is similar to the warm climates experienced on the Eastern slopes of the Rockies in Canada in winter. However, all this makes assumptions about the prevailing wind and temperature conditions: if these were altered, then the energy exchanges would differ, resulting in quite a different outcome.

(edited from Collis and Biggs, 1982, pp.4-5)

The authors conclude that,

These responses are increasingly better in quality. The first is not incorrect, but it tells you nothing about the quality of learning: it could have been given by a student who hadn't learned anything from the lesson. The second presents one relevant fact, the third several; neither gives an adequate explanation. The fourth response gives an interconnected and logical explanation, but as the fifth response makes clear, it could be an incorrect overgeneralisation. The fifth response considers all aspects, including some not given in the original lesson. (It introduces general abstract principles that cover both this situation and others; and it considers alternative possibilities to that implied in this question) (ibid, p.5 adapted).

The SOLO taxonomy represents a general structural analysis of the outcome of learning, as a complement to content-oriented analyses of the kind undertaken in the experiment which used the Dahllöf article. However, the great strength of such a taxonomy – its generality of application – is also its weakness. Differences in outcome which are bound up with the specific content of a particular learning task may remain unaccounted for.

In some of our analyses (e.g. Dahlgren and Pramling, 1982) structural differences in outcome similar to those represented in the SOLO taxonomy can be observed, and yet differences dependent on the specific content are repeatedly found. And subsequent analyses (Pollitt *et al*., 1985) have shown how the form of an examination question affects the outcome space 'available' to the student. Thus although structural similarities may be useful up to a point, they are likely to be more informative in their instructional implications, if they are combined with content-specific characteristics.

Outcomes as Conceptions

The content-specific analysis of outcome is important in another fundamental aspect. In some analyses, the categories of outcome arrived at can be considered as representing qualitatively distinct conceptions of a phenomenon. In other words, each constitutes a particular way of viewing and thinking about an aspect of the surrounding world. This is best illustrated by a study (Dahlgren, 1978) which ranged beyond the confines of a text-based learning experience. In this study, university students of economics were asked the apparently simple question:

Why does a bun cost about one (Swedish) crown?

In this case two categories of outcome accounted for the qualitative variation in the students' answers:

A. The price is dependent on the relationship between the supply of and demand for buns.
B. The price is equal to the (true) value of the bun.

Answers in category A represent a conception of price as system dependent, in that the price of a commodity is unknown until it is subject to a bargaining situation between producers and consumers in the market. Neither the costs of production nor customers' willingness to pay a certain price can alone determine the price. In the long run the price is determined at the point where customers and producers agree that goods or services will be bought and sold.

The category B answers on the other hand reveal a more object oriented conception of price, for these answers state that the price depicts the production costs and reasonable profits on the various constituents, whether they be products or services. In a sense this also means that B-answers give expression to a product-oriented conception of price. The B conception is one which is commonly found in everyday situations. It is often used, for example, by salesmen of luxury goods

like oriental rugs or paintings. As a customer you may hear for example that "this beautiful rug used to cost £1000 but we're only going to ask for £800". What is actually said, or what should have been said if we strive for a more correct description from an economic viewpoint, is that "we tried to sell this rug for £1000. It turned out, however, that that was a wrong price on the market we operate in, so now we are making a new attempt at the level of £800".

To sum up, these two categories of outcome are not just variations in what has been learned from textbooks, but represent two distinct and contrasting conceptions of a real-life phenomenon. The earlier examples can be viewed in the same way. In the case of the article by Dahllöf, for example, the variations in outcomes constitute different conceptions of Dahllöf's analysis of the shortcomings of an impending reform measure. What distinguishes the example about the price of a bun from the Dahllöf example, however, is that the phenomenon concerned occupies a relatively wider and more prominent position in everyday life – and thus more obviously draws on our experience and understanding of the surrounding world that is not confined to a particular text or set of learning materials. But in each case, the outcome does not amount to the retention or non-retention of a disembodied fact which has no meaning beyond itself. Instead, the phenomenon is invested with a specific meaning that both reflects and colours how the phenomenon is thought about.

From this same perspective, we can go further and define learning itself as a change in conception. In other words, when learning has occurred, there is a shift from one conception to another which is qualitatively distinct. Thus a student who had held conception B prior to an economics course and who is subsequently shown to display conception A has achieved more than the acquisition of an understanding of the laws of supply and demand. For the student, the phenomenon of price is now looked at in a fundamentally new way. Thus learning, within this perspective, is not a discrete and self-contained entity but one which has the potential of enabling individuals to consider afresh some part or aspect of the world around them.

The Effects of Education on Conceptions

But to what extent do learning experiences in formal education result in changes in conception? In reviewing the findings of the study of the introductory economics course (Dahlgren, 1978), we had concluded on a far from optimistic note. The main change we had observed was in the students' use of the terminology of economics. There was little evidence of qualitative changes in the students' conceptions of phenomena which had had a central place in the content of the course. Clearly, if these particular findings were representative of the effects of education in general, a reappraisal of the form and content of curricula seemed to be called for (Dahlgren, 1978, p. 18). And, indeed, similar findings were obtained in an investigation involving mechanical engineering students by Johansson *et al.* (1981). By choosing the seemingly trivial but very fundamental physical concept of force, they demonstrated that although it was taken for granted that the students

held the Newtonian conception of force (i.e. that a force is only involved in physical events where there is a change in velocity or direction), some of the students were in fact found to hold a different conception.

One of the questions put to the students was:

> A car is driven along a motorway in a straight line at a high constant speed. What forces act on the car?

An analysis of the answers yielded two categories of conceptions of a body moving at a constant velocity. A body in this kind of motion was apprehended either as

A. Having a constant velocity, due to the equilibrium of forces

(When he drives at a constant speed all the forces counterbalance each other); or

B. Moving, due to a "motive inequilibrium" of forces.

(And then a force that is directed forwards which has to be greater than those... forces directed in the opposite direction, otherwise it wouldn't move forwards).

Of the 22 students who were asked questions about bodies moving at a constant velocity, a total of 7 gave B answers at the first interview (prior to the course in mechanics) and 6 at the second interview (after the course had finished). Although the outcome on other questions (e.g. the case of decelerated motion, illustrated by an ice-hockey puck gliding straight forwards on smooth ice) was more positive, it seems nevertheless remarkable that a significant proportion of the students could preserve an Aristotelian conception of force.

The effects of formal education on conceptions have also been investigated by Hasselgren (1982), in a longitudinal study. His study focused on the structural level of the subjects' conceptions. A group of pre-school student-teachers were asked to describe what they saw in video-tape sequences of children at play. The sessions were repeated three times; at the start of the course, in the middle of the second term, and at the beginning of the third and final term. A group of physiotherapy students constituted a control group. In interviews following the video-tape sequences, the subjects were questioned about what they had seen. The transcripts of the taped interviews were analysed and a set of four categories of outcome were identified (Hasselgren, 1982, pp. 50-52):

A. *An abstracting description.* In relating the content of the video recordings, what is shown on the screen is not taken for granted, but instead is considered as a concrete illustration of a principle or abstract idea which might be applied to the material.
B. *A chronological description.* The activities of the group of children are understood as a chain of events, following a temporal sequence.
C. *A partialistic description.* The account given deals with a part rather than the whole of the video-tape, often by focusing only on the actions of one of the children.

D. *A fragmentary description.* The account is impressionistic and diff use, lacking an identifiable perspective and only mentioning what is immediately observable. The children, their play, and the setting in which they are playing, are given equal importance.

In Hasselgren's analysis, these four categories are considered as forming a developmental sequence, in which there is progression from either the fragmentary or partialistic description to the chronological description, and hence to the abstracting description. His analysis shows a substantial difference in the patterns of regression, stability and development for the experimental and control groups as shown in Table 2.1. There were only five instances of regression, all confined to the control group, and a very much higher rate of instances of development amongst students in the experimental group. Hasselgren therefore concludes that the formal educational experiences undergone by the experimental group have had an impact on their way of apprehending a phenomenon that is central to pre-school teacher education.

TABLE 2.1
Distribution of changes representing regression, stability or development
(from Hasselgren, 1981, p. 63)

	Category			
Group	*Regression*	*Stability*	*Development*	*N*
Experimental	—	37	19	56
Control	5	23	3	31

An attempt at summarising research evidence on the effects of education, within the qualitative perspective adopted here, leads to the following observations:

- Education does have an impact as far as the acquisition of subject-specific terminology or the mastery of problem-solving algorithms are concerned, and such outcomes may be the most permanent of any effects which can be identified.
- Conceptual changes are undoubtedly more difficult to trace. Such changes do take place but are probably relatively rare, fragile and context-dependent occurrences. (Dahlgren, 1978; Brumby, 1979; Johansson *et al.*, 1981).
- Nonetheless, at a macro-level of analysis, education has demonstrable effects in terms of structural properties of the ways in which phenomena are apprehended. (Perry, 1970; Hasselgren, 1982).

The Qualitative Analysis of Learning

Having provided some examples which illuminate the kind of results about learning that a qualitative analysis can yield, we may make an attempt to integrate the conception of learning and knowledge that springs out of that perspective.

The first point to emphasise may seem obvious, but is sometimes ignored: learning is a many-sided phenomenon. Just as there are many different things to learn about, so too are there different processes of learning and different outcomes of learning. In this chapter we have tried to contrast two main categories of learning. On the one hand there is learning from materials that lack an internal order which might permit us to talk about meaningfulness. In such cases the learning process involves pure memorising either by dint of constant repetition or by imposing some kind of meaningfulness, often through the use of mnemonic strategies.

But a substantial proportion of learning depends on understanding material which does have an internal structure that can be grasped. In these cases the process of learning should aim at finding this structure in as deep a sense as possible. This is a qualitatively different kind of learning which will result in a different outcome. The nature of this outcome is that it represents a conception of a phenomenon in the surrounding world. A conception can in principle mean those very superficial characteristics of a phenomenon such as size, shape or colour. Here that conception is taken rather to denote the nature of an object or an event.

To "understand" or "accept" the colour or the size of an object is a process of a totally different kind than to understand its nature. In the latter case, what is pivotal to understanding is the grasp of the relationships between a phenomenon and its context. External or concrete characteristics of a phenomenon do not alone provide a basis for understanding. In this respect everything is always a part of something larger or more inclusive (i.e. it has a meaning beyond itself) and it is this which makes up what we might call the context of understanding. Meaningfulness is thus not an inherent property of nature or culture. It is imposed by human consciousness, which is itself evolving continually. Learning, then, should be regarded as that aspect of human life through which the environment – or man himself – appears with a higher degree of meaningfulness than before. From this perspective – as in some others too (cf. Popper, 1972) – knowledge is nothing but a series of occasional, provisional steps towards what is often described as an unreachable complete knowledge about reality. Similarly a conception, as Marton (1978) describes it,

> often denotes the implicit (tacit) – that which does not need to be expressed or cannot be expressed because it has never been the object of reflection. (p. 20)

This chapter has shown how it is possible to describe what is learned in terms of sets of categories which can often be differentiated in terms of their structural properties. Such structural differences would seem to hold open the possibility of devising empirically derived taxonomies, such as SOLO, which would allow the quality of a wide range of learning outcomes to be systematically analysed. Yet

our research has drawn attention to variations in outcome which cannot fully be understood except in relation to the content of learning. Analyses of learning outcomes in relation to content enable us to describe variations in the conceptions students hold about important parts of their course. These analyses also suggest that, at present, formal education is not as successful as it might be in helping students to develop more sophisticated conceptions. When the questions asked of students are at base quantitative or fail to penetrate beyond what can be more or less unreflectively retained in the memory, students' misapprehensions are disguised within spuriously satisfactory answers or cloaked in technical jargon. More searching questions, though framed in a direct and straightforward way, show up fundamental misunderstandings. Thus a study of qualitative differences in outcome has a vitally important role to play in helping to determine – and ultimately improve – the quality of student learning.