LEARNING AND TEACHING

Donald A. Norman, et al

California University

Prepared for:
Office of Naval Research
Advanced Research Projects Agency

March 1974
LEARNING AND TEACHING
DONALD A. NORMAN
DONALD R. GENTNER
ALBERT L. STEVENS

ARPA Order Number: 2284
Task Authority Identification Number: NR 154-360
Program Code Number: 3D20
Effective Date of Contract: 1 February, 1973
Contract Expiration Date: 31 January, 1975
Amount of Contract: $221,012

Organization: University of California, San Diego
Principal Investigator: Donald A. Norman
[Phone (714) 453-2000 x 2018]
Scientific Officer: Dr. Marshall Farr, Director
Personnel and Training Research Programs
Office of Naval Research

The views and conclusions contained in this document are those of the authors
and should not be interpreted as necessarily representing the official
policies, either expressed or implied, of the Advanced Research Projects
Agency, the Office of Naval Research, or the U.S. Government.

This research was supported by the Advanced Research Projects Agency and the
Office of Naval Research of the Department of Defense and was monitored by
Learning and Teaching

Donald A. Norman
Donald R. Gentner
Albert L. Stevens

Center for Human Information Processing
and the
Department of Psychology
University of California, San Diego
La Jolla, California 92037

A considerably revised version of this report will appear as
Chapter 15 of
D.A. Norman, D.E. Rumelhart, and the LNR Research Group,
"Explorations in Cognition"
San Francisco: W.H. Freeman and Co. 1975
In this report we examine the process of learning large, complex subject matters, asking about the ways by which teacher and student communicate the necessary knowledge structures to each other, and asking how a student comes to select an appropriate paradigm for solving a problem and to revise inappropriate ones. We examine protocols from several different learning situations, and have made a first attempt to put together the pieces into a cognitive theory of learning and teaching.
Learning and Teaching

Donald A. Norman
Donald R. Gentner
Albert L. Stevens

Learning and Teaching as a Process of Communication

Learning and teaching can be viewed as a process of communication. The teacher has the task of conveying a particular knowledge structure to the student. The learner has the task of deducing just what structure is intended by the teacher, as well as the additional task of adding the new information to that already existing in such a way that it can be referred to and used at a later time. Many of the problems of learning and teaching can be understood as problems in this communication process. Learning, however, is unlike most simple communications in that the structures that are to be acquired are complex, and it is not always clear just how they are to fit together. Moreover the differences in the knowledge shared among the participants in a learning situation are often considerably greater than in simple discourse.

In this chapter we examine just what is involved in learning and teaching when there is a reasonably complex, yet well defined topic matter. The topic should be something that takes weeks or months to learn to a reasonable level of understanding. For the purposes of this discussion, the topic could be almost anything of suitable size and complexity. In fact, we examine two topics that meet these requirements for definiteness and complexity:
An elementary computer programming language

Cooking -- specifically an understanding of the French sauce family, as covered in pages 54 through 93 of Child, Bertholles, and Beck (1961).

When a teacher wishes to start, he must have some idea of the structure of the material that is to be covered as well as some idea of the knowledge that his student has of the subject matter. Then, the task is to give instruction on the difference between these two structures. But to do this requires knowing just how the learner will interpret the information given to him. The network of information is rather complex, and there are large numbers of interrelated concepts. Just how they are related to each other, and just how they should be approached is not a simple issue. We return to this point later. Somehow, though, the teacher must use his model of both the subject matter and the student to determine how to present the information. He must interact with the student in doing this so that the student can improve his own model. In fact, a good part of the teaching task is perhaps best seen as giving the student sufficient knowledge of the nature of the task and the nature of his own processes that he can acquire the subject matter for himself.

The prior knowledge of the student obviously makes a large difference. Consider the difference between teaching someone who already has a good general background in a subject matter and someone who has none. The difference is most obvious in simply comparing how much must be told to the two students. Thus the recipe for mayonnaise for an advanced cook takes exactly 6 sentences (67 words) and a list of ingredients. Essentially the very same recipe for a beginning cook is around 13 times longer: 3 pages of text or about 900 words and a list of ingredients. (The advanced example comes from "Masterpieces of French Cuisine" by Amunategui, 1971, and the introductory
example from "Mastering the Art of French Cooking" by Child, Bertholle, and Beck, 1961.) Similarly, a skilled programmer can learn a computer language such as Basic or Flow essentially by looking at the commands, so that it takes a very short session to learn the language, even with no prior knowledge of its structure. A beginning programmer, however, can struggle for 6 weeks to reach the same level of skill. These examples illustrate that the difficulties in a subject matter are not necessarily intrinsic to the topic itself, but often simply reflect the rather large number of prerequisites that must first be acquired.

The learner has a problem somewhat different from that of the teacher. It is the task of the learner to deduce the structure of the material that is to be learned and to determine just how it is to be added to his already existing knowledge structures. He cannot do this until he fully understands what he is supposed to learn. In fact, understanding is the key to the whole task, for once the material has been understood, then often the act of learning follows trivially. There are pitfalls for the unwary student. A structure that appears to fit may in fact be wrong. Too often a learner acquires a new concept by false analogy to an old one or by too simple a structure. In these cases, he may construct a knowledge structure appropriate to the material he has been exposed to so far, but which is incorrect in general and so will cause difficulty later.

Often, the learner cannot easily even determine just what is expected. The teacher has a lot of material to convey, some of it less obvious than others as to the role it plays in the overall knowledge structure. As the learner experiences each concept, he must somehow determine how it fits into the overall structure that he is trying to create. He faces a task well
known in the concept formation literature in which a multi-dimensional structure is presented to the subject on each trial and the subject must extract the relevant dimensions from the irrelevant ones. A difficulty in many learning situations is that the participants do not recognize the aspects that make the task one of concept formation. As a result, there is minimal feedback to the student. Worse, undue concentration on the irrelevant dimensions of a task can temporarily appear to lead to progress, thus leading to the selection of erroneous structures.
A recursive definition of understanding is that we understand a concept when:

If a concept is not analyzable into underlying concepts, then
we understand it whenever we have acquired its structure.

If a concept is analyzable, we understand the concept when
1. we understand each of its components
2. we understand how each component fits into the general
   structural framework for the concept
3. we understand how the concept fits within our knowledge
   of previously acquired structures.

The essential point is this: we understand when a knowledge framework or
schema exists for the material that is to be acquired. The focus of our
concern is on how the appropriate knowledge schema is acquired.

In thinking about understanding it is important to realize that it is
possible to process information at different depths. Thus, one can fail to
understand at many levels. One could understand at the surface level, but
fail at the semantic. One could understand at the semantic level, but fail
conceptually. One could understand conceptually, but fail to integrate the
new material within the old. We can illustrate this with an example.

The following story comes from a paper by Bransford and McCraeell (1972:
the same example also appears in Bransford & Johnson, 197[3]). A group of
subjects listened to a story. Afterwards, they were asked to judge its
comprehension and to recall as much as they could. Here is the story.

The procedure is actually quite simple. First you arrange
things into different groups. Of course one pile may be suf-
cient depending on how much there is to do. If you have to
go somewhere else due to lack of facilities that is the
next step, otherwise you are pretty well set. It is import-
ant not to overdo things. That is, it is better to do few
things at once than too many. In the short run this may not
seem important but complications can easily arise. A mis-
take can be expensive as well. At first the whole proce-
dure will seem complicated. Soon however, it will become
just another facet of life. It is difficult to foresee any
end to the necessity for this task in the immediate future,
but then one never can tell. After the procedure is completed
one arranges the materials into different groups again. Then
they can be put into their appropriate places. Eventually
they will be used once more and the whole cycle will then have
to be repeated. However, that is a part of life.

(1973; Manuscript pages 26 and 27)

The story, as it stands, makes little sense. The problem is that there
is no overall structure to the story. What makes this an effective demonstra-
tion of the power of a structural schema on comprehension is that we need
only add two words to make the story intelligible. The two words are the
title, Washing Clothes. Now, suddenly, all the phrases fit together in the
story and it becomes both comprehensible and recallable. Note that in the
actual experiment it was important that the title be given before the story.
The subjects who heard the title before hearing the story gave it a mean
comprehension score of 4.50 (on a 7 point scale), whereas subjects who heard
the title only after hearing the story rated its comprehension as 2.12. (The
comprehension score given by subjects who were not given the title was 2.29.)
In a similar fashion, the mean recall of "items" in the story was 5.83 (out of a possible 18) when the title was told before hearing the story, but the mean recall score with the title given last was 2.65. (Subjects who did not hear the title yielded a recall score of 2.82.)

The problem in comprehending the "washing clothes" story comes from the absence of a schema that would bind together the elements of the story. Without such a schema, then each sentence can be parsed and used to construct a simple semantic net, and new sentences can refer to the concepts set up by previous ones. We have the concept of a pile, of facilities, and of doing things. But there is no overall structure: the purpose and result of all the activity remains unspecified. The title, however, evokes a rich set of past experiences. These provide a structural framework for the story. From the schemata that have resulted from our personal experiences with the washing of clothes, it is possible to generate a schema that ties the story together. Note that Bransford and McCarrell illustrate that the conceptual structure must exist at the time the story is heard: it is too late to provide it afterwards. Thus, the formation of the semantic structure for any particular incident is affected by the overriding organizational schema. It is not sufficient that the schema provided by the title gives a concrete reference for the concepts of the story, for if that were so, there would be equal comprehension scores when the story title was presented first and later. The organizational framework for organizing and relating concepts together must be active at the time the information is received.

A Structural Network Analysis of Learning

When one examines the structure of memory, one sees that it is an interconnected network. If we wish to insert new information, then somehow
the new information must get related to the old. Examine Figure 1.

In Figure 1A we show a segment of knowledge from memory. Suppose we were to acquire two new concepts and some relations between them, say that \( R \) relates \( C1 \) to \( C2 \) (Figure 1B). This knowledge must be interconnected within the previous knowledge, else retrieval and understanding will be difficult. Thus, in Figure 1C we show the two newly acquired nodes to be well integrated. To use the obvious analogy from the structure, in Figure 1C we say that the newly learned components are well supported, whereas those of 1B lack support.

The normal events of the world are easy to remember. Nonsense syllables are difficult, unless some suitable mnemonic is found, in which case they at times can become easy. Cooking can be difficult unless some prior background knowledge exists within which to classify each new recipe and technique. Programming is difficult, unless sufficient knowledge is already known that a new programming language can be classified as a simple case of something already known.

The point can be illustrated with another experimental task performed by Bransford and Johnson, one that is ideally suited to illustrate the point made in Figure 1. Subjects heard the following story:

1. Title: Watching a peace march from the fortieth floor.
2. The view was breathtaking. From the window one could see the crowd below. Everything looked extremely small from such a distance, but the colorful costumes could still be seen. Everyone seemed to be moving
in one direction in an orderly fashion and there seemed to be little children as well as adults. The landing was gentle, and luckily the atmosphere was such that no special suits had to be worn. At first there was a great deal of activity. Later, when the speeches started, the crowd quieted down. The man with the television camera took many shots of the setting and the crowd. Everyone was very friendly and seemed to be glad when the music started.

(Bransford & McCarrell, 1972, p. 27)

The issue, obviously, has to do with the recall of the sentence:

(3) The landing was gentle, and luckily the atmosphere was such that no special suits had to be worn.

We have one overall schema for the story, with the title setting the stage. From the title alone we deduce that there will be a crowd of people, viewed from within a tall building, through windows (or on a viewpoint) above. There is a strong presupposition that we are watching from a building. Hence, all the definite references in the story (except those of sentence 3) have a concrete realization. It is easy to draw the semantic structure that represents the story.

Now consider how we solve the problem of reference. Starting with the title, several concepts are introduced and added to the general context. In the terms of the analysis introduced in Chapter 3, these concepts are "foregrounded." Thus, the phrase "view from the fortieth floor" brings to the fore such concepts as a building, a city, a tall height, and a window or a roof. This allows us to understand the references in the first two sentences.
to "the view," "the window," and to "the crowd." In fact, all the concepts are easily related to the story's structural schema until we reach the critical point: the fifth sentence of (2), shown as (3). Sentence 3 introduces three new concepts: "the landing," "the atmosphere," and "special suits." Of these three concepts only, perhaps, the second, "the atmosphere," can find a definite reference within the schema. As a result, given that we can find no place in our memory structures for the concepts of sentence 3, we are forced simply to make up arbitrary nodes for them, nodes that have essentially no contact with the rest of the story.

The semantic structure that results from hearing the "Peace March" story looks very much like the structure shown in Figure 1B. Most of the sentences form a cohesive, well integrated structure, as depicted by the shaded area of the Figure. But sentence 3, the anomalous sentence, forms a structure much like that shown as the isolated relation within Figure 1B. Although well formed, the structure makes no contact with the rest.

In the experiment performed by Bransford and McCarrell, most of the sentences of story (2) were recalled well, with the exception of sentence 3. Sentence 3 was not well recalled even when presented with a cue outline of the form

(4) Luckily the landing and the atmosphere.

When the same story was read to subjects, but with a different title:

(5) Title: A space trip to an inhabited planet

subjects showed much greater recall of sentence 3. That this should be so is obvious by considering how the title change affects the determination of the reference for the concepts in the story. Essentially, the major change is for that of sentence 3. Now the structure is transformed into something much more like that shown in Figure 1C.
Cooking

How to make Mayonnaise

The study of cooking provides a useful example of the difficulty of learning complex subjects. In cooking -- especially classical French cooking -- any given recipe contains a large number of components, each relatively simple by itself. Thus it is with mayonnaise: the ingredients and the form of combination are almost trivial. But, to a non-cook, the combination is not at all an obvious one. First, no heat is applied. Second, the basic components do not match most people's conceptions of what it should take to make that smooth, white, creamy substance we know as mayonnaise. In fact, asking naive subjects just what they expect mayonnaise to be made of helps expose their basic conceptual schemata of cooking ingredients. Mayonnaise is an especially good thing to ask people about because its structure is so non-intuitive that the respondent must derive an answer; usually there is no specific knowledge that can help. For your amusement, in the accompanying box we present protocols collected from naive subjects:

---

Insert Box Here

---

The mayonnaise protocols illustrate that previously developed schemata can be applied to new problems, not necessarily in appropriate fashion, of course. The determinants of which schemata get applied are the features which come from what is known: in this case, the features of the end result. The problem is that the relevant properties of egg and oil and their interactions (that they can form a white creamy substance, are not
DAN: How do you make mayonnaise?

CN: How you make mayonnaise is you look at a cookbook.

DAN: OK, but without looking at a cookbook, can you guess what it is that's inside of mayonnaise?

CN: Uh.

DAN: How you would make it?

CN: Uh Butter -- uh let me think (5-sec pause) hmm (10-sec pause) whipped cream very very very fine--ly whipped so it's smooth.

That's probably how you make it, just with whipped cream, very very very fine and smooth.

DAN: Anything else?

CN: You might add a little taste to it.

DAN: Taste of what?

CN: (10-sec pause) Sort of a vanilla taste.

DAN: Suppose I said that mayonnaise is made from egg yolk -- and oil.

What would you say?

CN: I would say it's very very -- wrong.

DAN: Why?

CN: You can't just make mayonnaise out of egg yolks and water -- I mean and oil.

DAN: Why not?

CN: Because of taste and smoothness and stuff like that.

[Protocol of the experimenter, DAN, and CN, an 8-year old female]
DAN: How would you make something like mayonnaise?
GB: Mayonnaise? How do you make mayonnaise? You can't make mayonnaise, it has to be bought in jars.

Mayonnaise. Um. You mix whipped cream with, umm, some mustard.

[Protocol of the experimenter, DAN and GB, an adult male psychology professor]
part of any stored schema, so the correct answer cannot be derived. As the examples illustrate, mayonnaise has properties which make it look more like certain dairy products than the result of mixing the yolks of eggs with oil.*

---

* Footnote: To make mayonnaise, one puts 2 egg yolks in a mixing bowl with 1 teaspoon of vinegar or lemon juice. Seasonings may be added (dry mustard, salt, white pepper). Then, while beating vigorously with a wire whip, oil is added, drop by drop, until the mixture is thick and creamy (about 1/3 cup of oil). More oil is then added in a slow steady stream (all the while beating vigorously). When 1 to 1 1/2 cup of oil has been added, mix in a second teaspoon of vinegar or lemon juice.
Tutorial Interaction*

The amount of prerequisite knowledge required by French cooking is very high, yet reasonably accessible through observation or interviews.

As an example, consider the family of French sauces. They are organized in a systematic fashion, with white and brown sauces, butter sauces, the mayonnaise family, and oil and vinegar sauces. Within these classes, certain basic components occur over and over again.

**In sauces:**

- **Roux,** a mixture of butter and flour
- **Stock,** clear or brown liquid
- **Bechamel,** heated milk
- **Veloute,** a basic white sauce
- **Enrichment,** with butter, cream, or egg yolks

**In Mayonnaise -- Hollandaise type sauces**

The method of adding butter or oil to a mixture of egg yolk and vinegar or lemon juice.

Now, none of these compounds or techniques is fixed, but rather every one can be modified without changing the principles. With sauces, therefore, it is possible to test for such things as the influence of prior experience on the learning of these sauces and the generalizations and contrasts acquired by the subjects.

* We thank Julie Lustig, Sandford Schane, and Fred Wightman for serving as Tutors and Tutees.
Suppose you set out to teach someone the family of French White Sauces. The question is how should you teach it? Just how does one get the entire grid of interrelated concepts across. We have tried some exploratory studies in tutoring, studying how beginning, advanced, and expert cooks teach and learn these concepts when acting as tutor (or tutee).

**Advanced Tutor -- Beginning Student**

When an advanced cook tutors a beginner, there is a tendency to lecture at first, describing the overall family of sauces, but then, when the overall description has been completed, usually the beginner's lack of understanding causes a fumbling, exploratory interaction, to ensue in which each tries to understand what the other is doing.

Here is an example of a segment of such a tutorial. The advanced tutor (T) had finished a 15 or 20 minute lecture on the white sauces. He concluded by saying "And I think that's all there is to say." The student (S) had been following, making appropriate comments along the way. But now, unexpectedly, the student asks a question which indicates that she does not understand the overall pattern of sauces. The tutor is disturbed, and there follows a period in which the tutor tries to straighten out the concepts. This is a portion of that conversation. It starts with the student attempting to summarize her understanding of the sauces.

*S:* We start out with two different white sauces, veloute and Bechamel -- and Bechamel is with -- milk did you say?

*T:* Right.
S: Veloute is with egg yolk and cream so it's richer.
T: Uh -- no. Bechamel is with milk, veloute is with stock. That's the basic difference.
S: Oh that's right. We made the veloute into Parisienne.
T: Either the Bechamel.
S: OK.
T: Or the veloute can be a Parisienne.
S: What American cooks mean when they say a white sauce is -- Bechamel. They mean a roux with --
T: Milk.
S: Milk.
T: And that -- in incidentally there's another sub subcategory and that is simply the basic Bechamel sauce enriched with just a little bit of cream. So that we've
S: Mm--hm.
T: got the the roux plus milk, and then you can make it a little richer just simply by adding cream, then it becomes a sauce Creme. If you do the same thing to a veloute sauce, it becomes a sauce Supreme.
S: Mm.
T: Those are those are typically end points. If you simply desire a very simple sauce not enriched with butter and egg yolks as the Parisienne is, if you simply want the basic sauce without the the egg yolks enrichment, you add the cream, and if you start with the Bechamel you get a sauce Creme, if you start with the Veloute you get a sauce Supreme.
S: (long pause) OK. (long pause) If you take -- a -- sauce made with stock instead of milk and cream, and add egg yolks I -- is that ever done?

T: That's a Parisienne.

S: The Parisienne doesn't have to start with the veloute.

T: The veloute is --

S: Has -- cream or milk.

T: No -- no -- the veloute is the roux and the stock.

S: I keep -- I keep mixing that up. The veloute is stock -- and then cream --

T: The Bechamel is the milk -- base. (pause) And from both of those you can get to Parisienne by adding egg yolks and cream.

S: OK.

T: (pause) Now there are names for the various types of sauces which can be classified under the same category Parisienne and that may be a confusing -- a confusing category because we started out originally with Bechamel and veloute which are fundamentally different white sauces.

S: I'm confused because Bechamel is what I originally learned as a white sauce and a white sauce is a large class with all these different kinds of sauces.

T: That's right.

S: And that's why I'm confused -- and I understand what you told me. I just have to think it through to remember whether the Bechamel or veloute has the stock -- and it's the veloute that has the stock.
In this tutorial there are clear signs that the student is attempting to integrate the new descriptions of the sauces into her previous idea of a white sauce. That is, a good deal of the difficult, was that she entered the learning situation with a partial knowledge of sauces (the standard American white sauce) that conflicted with the entire family of sauces she was being tutored on. In this case, the previous knowledge was especially disruptive to the learning experience because it was so closely related: at first, it appeared to be quite compatible with what she already knew. It was only when the discrepancy between her previous conception of white sauce and the French conception was explicitly mentioned that she began to make sense out of her lesson. The interesting thing about this tutorial session was that for the first 30 minutes, neither tutor nor student realized that there were any problems. It came as somewhat of a shock to both participants to realize that there were vast confusions, and the entire session lasted for 45 minutes beyond the point where the tutor had initially ended the session, saying, "And I think that's all there is to say."

**Expert tutor -- Advanced Student**

When an expert tutors an advanced student, then the tutorial takes on quite a different quality. The interactions become more like a normal conversation than a tutorial. After teacher and student each come to appreciate the range of knowledge understood by the other, then they can query back and forth, much more in the manner of conversation than of teaching.
The discovery of common points may occupy a good deal of the session, however. Life is much easier for the tutor if he can assume the student knows nothing. When this is not the case, problems can arise because the possible range of knowledge is so wide, that there are many possible places for confusion. In a broad topic such as "classical French cooking," it could easily take days (weeks) to discover the exact range of knowledge of the student.

In this next example, between a tutor (T), an expert, and an advanced student (S), much of the dialogue deals with this period of learning about the other, as the tutor continually revised his assessment of his student. (In this session, the student, S, is actually the same person who was the tutor, T, in the previous protocol.)

At first, the tutor and student spent some time getting adjusted to the situation and to each other. Initially, the tutor underestimated the level of skill of his student, but slowly picked up hints that his estimation was off. For example, the student once, unexpectedly, used the word "roux", a word that the tutor thought he shouldn't know:

S: Mm-hm. A roux.

T: (after a pause) You already know something.

As the session went on, the mismatch between the two continually arose.

T: -- I mean, you may also be curious about why you have to do things in a certain way and why --

S: (simultaneously with the end of the above) Of course, yeah, yeah. I'm particularly interested in the latter. It seems when I make a white sauce they do turn out, and you know, they're really good and all that, so I presume I'm not going too far -- astray

T: So you're interested in some theory other than just engineering.
The tutor decided it was the student's turn to talk, thereby to provide information about his level of knowledge.

T: OK. Why don't you tell me how you would make a white sauce.

This finally got things going. The student gave a long monologue on his procedure, from which the tutor was able to pick up several points that required more discussion. We follow the tutorial at the second such point.

T: Um -- now there's another interesting point that you mentioned and that is that you said that you must boil the liquid first -- or you said I would have boiling liquid ready, whether it was fish stock or chicken stock.

S: OK -- yeah

T: or something or milk

S: That comes -- that comes

T: and then you would pour it in

S: That comes right from the cookbook. I mean I just -- I just remembered that. Alright? Now, I don't know why boiling should be important or whether it's desirable, particularly in the case of milk it seems -- it's kind of an odd thing to think about but -- but yeah, go ahead, what were you going to say about it?

T: That's an interesting thing, because I've experimented with that and what I've found is the following. That if you -- obviously if it calls for a cup of uh -- liquid

S: Mm-hm

T: and you take -- let's say you're making a white sauce and you're using milk --
S: Mem-hm

T: Then you take the milk right from the refrigerator. OK if you pour a cup of milk right into this this roux

S: Mem-hm

T: then you're going to get exactly the same thing happening that you don't want happen -- that is it's going to lump up

S: Hmm. Temperature is clearly very important.

T: So -- right, the temperature's important, but -- it's not -- the critical thing, the critical thing is the quantity of milk that you're pouring in, OK? So I never heat up my milk, I don't have the time. The only time I heat up is if I'm using, say I just poached fish and obviously the liquid is there and it's hot, I'll pour

S: Right

T: it all in and make the sauce

S: Right

T: but if it's calling for milk, I'll take it out of the refrigerator, and pour in a small amount at a time -- OK?

S: And then it works?

T: and allow the pan to keep coming to a -- a boil

S: I see -- you've got the, you've got the heat on so -- in fact you're you're heating the liquid in the pan (simultaneously with following)

T: (simultaneously with above) OK? and you pour in more and more and more, right, so it seems there's some critical --

T: there's some critical cold temperature at which you wreck the thing, OK? but you can add cold liquid in slowly, slowly, slowly and keep it heating up and it works fine -- OK,

S: Hmm
S:  Um.

T:  so we do have a temperature factor going on

S:  You saved me 10 minutes.

These tutorial illustrations give some flavor to the interactions that can take place among teacher and student. So far, we have tried to characterize the teaching/learning process as one of communicating the knowledge structures of the teacher to the student. These tutorial examples illustrate how important (and how difficult) it is for the teacher to know what knowledge the student has about the subject matter.

In the next (and final) example, we examine the learning of a different topic matter: an elementary programming language. This next topic introduces as much more directly to the role that the schemata acquired by the student play in the learning process.
Learning a Programming Language

Computer programming represents a different type of subject matter than does cooking, for it emphasizes the problem solving skills of the student and has less reliance upon a set of fixed principles and recipes. We have examined the learning of one rather unique language that has a number of distinct advantages as a tool for the investigation of learning principles. The language is called FLOW. It was developed at the University of California, San Diego by Professor Jeffry Raskin to teach programming to students in the visual arts. These students had little knowledge of science and mathematics, and moreover, they did not particularly care to get any.

We have found that the study of how a student learns FLOW has important advantages as an experimental vehicle:

* The task is well defined.
* It is reasonable in size and complexity: simple enough to be analyzed, complex enough for our studies. A student with no previous experience can learn a significant amount in 30 minutes, but it usually takes about 20 hours to master the language.
* The task seems to be interesting. Students with wide ranges of backgrounds work enthusiastically at it.
* The task gets at both conceptual and procedural knowledge.
* The fact that the task is computer-based aids the experimental facilities, allowing automated collection and recording of protocols.
FLOW

FLOW has been designed to simplify the process of entering information into the computer. At any point in the program, only the typewriter keys which lead to legal commands are operative. When a key which would lead to an illegal character is depressed, it has no effect. In fact, it is impossible to make an error in syntax. In addition the entire command appears on the screen as soon as the student has typed its first letter. Thus, by these two features, the most common problems for the beginner are eliminated: typing errors and difficulty with the keyboard. In addition, we have modified the system to add several other useful features for our studies.

The command set of FLOW is illustrated in Table 1. In this table, the part of the command that the student must type is underlined. The language is essentially self-explanatory, except perhaps for the commands that refer to "it." "It" is the name of a pointer that refers to a single letter in a string of text (the text is always the string of letters within the "Text is....." statement that was encountered most recently in the stream of processing). When the command "Get it" is first invoked, the "it" is used, the pointer moves one letter to the right along the text string. (A text string is assumed to contain an indefinite number of blanks at its right, so that when repeated use of the "get it" command runs out of letters on the text string, it will then continually point at a blank.)
Table 1
The FLOW Language

The student only has to type the underlined letters. (Some commands fall into more than one category, and so they are repeated.)

CONTROL STATEMENTS

If it is 'E' jump to 235
If counter is 42 jump to 240
Jump to 10
Stop

COUNTER CONTROL

Make counter zero
Add one to counter
Decrease counter by one
If counter is 7 jump to 290
Print counter

TEXT MANIPULATION

Text is 'THE HOUSE IS RED.'
Get it
If it is 'E' jump to 325
Print it

PRINT STATEMENTS

Print 'THAT IS CORRECT:'
Print return
Print counter
Print it

SYSTEM COMMANDS

Run
Walk
List
New
Help
(Escape)
Backspace -- Line numbers
For example, a program to count the number of times the letter 'E' occurs in a sentence looks like this:

10 Text is 'THIS IS A SAMPLE SENTENCE.'
20 Make counter zero
30 Get it
40 If it is '.' jump to 200
50 If it is 'e' jump to 100
60 Jump to 30
100 Add one to counter
110 Jump to 30
200 Print 'The number of E's is '
210 Print counter
220 Stop

This language allows many fundamental properties of programming to be taught while maintaining a simple structure.* The concepts of conditionals can be taught, as well as simple text manipulation. A pointer is present.

* To the reader who believes this to be too simple a language to take seriously, we urge him to attempt these two problems.

a. Print 'yes' if the last two letters of the text are 'ie' or 'ei'; print 'no' otherwise.
The program should print 'yes' for dei, die, diie and diei, and 'no' for died, dice and deli.

b. Print the first word that contains an 'e'. If the text is 'This is a sample sentence', the program should print the word 'sample'.
Algorithms and iterations can be taught. Some things cannot be done, such as letting one program be called by others (subroutines, co-routines, recursion). Any program that requires more than one pointer or more than one counter at a time cannot be performed. But despite these limitations, FLOW teaches many of the basic programming concepts that are used in the more advanced algebraic languages such as ALGOL, BASIC or FORTRAN. Classroom experience supports this statement.**

**Learning FLOW**

When a student who has no previous experience with computers learns a computer language such as FLOW, he is required to master a number of major new concepts and procedures. It is a novel concept for most people that they may give a machine a set of instructions to be performed at a later time. Program loops lack any close parallel with most people's experience. New types of procedures must be learned, such as the use of a computer terminal and the running of a stored program.

Although a person may have never programmed a computer before, he still has a large body of previous experience which he brings to the task. Almost everyone has some concept of a computer and the things that it can do. This

**A number of people have criticized any choice of FLOW, arguing that the pedestrian structure of this class of languages does not allow the student to develop the "powerful ideas" that are so important in general cognitive skills. In this sense, we are told, FLOW has no future. We should be using LOGO, or SMALLTALK, or even IPL. We agree with these critics. But our purpose here is to see how a student acquires a new skill, not to change (or even enrichen) the thought style of the student. For our purposes, any task that is reasonably complex but yet self-contained will serve. FLOW has taught us about learning: that it is sufficient justification for its use.**
is sometimes a problem. In some of our studies, after the student had learned to print a word, we asked him to print the same word a large number of times. The straightforward way to solve this problem would be to write a program that had many print commands; one for each word that was desired. (The student did not yet know about loops.) Indeed some students do this. But a number of students don't bother with this method; they are sure that a computer must have a better way, and typically search (in vain) over the different keys on the typewriter keyboard for one that will do the job. (Unfortunately, one key is indeed labelled "repeat", but it has a different function, and in any event, it is disabled during our experiments.)

The commands in FLOW bear a modest resemblance to English, and therefore language has a strong effect on the early learning of FLOW. For instance, suppose a student's only exposure to the print statement is the single example:

```
010 Print "Hello"
```

He runs this example program and sees that the computer displays

```
Hello
```

What does a student conclude, tentatively, for this example? Would the same thing happen if we tried the example again? What if we put a different word between the quotes? What about two words? Spaces? What is the relevance of the number 010? Students, in fact, generalize very readily in these circumstances. To a great extent the generalizations are guided by the word "Print". A person will have a certain class of things which are thought of as being printable, and this class will influence the generalization he makes from an example of a print statement. Thus we see that the word "Print" used in a computer statement serves not only as a mnemonic, reminding the student to use this statement when he wants to print or display something, but it also
interacts with the student's previous knowledge of the word to influence the learning of the computer language.

Some examples

Here are some examples of the problems encountered by students who are attempting to learn FLOW. We concentrate on the learning of the very first few concepts, (such as the way in which simple iterations can be used with the print statement). We are primarily concerned with examining the interaction between the students' understanding of the concepts through the schemata that they have developed and their attempts to generalize these schemata to new experiences.

In the first example, we follow a student who has learned to use the print command with several different programs, but where each program was no longer than two lines long, and where many had only a single print statement. At this point, she is instructed to type the following program onto the display terminal,

Program 4

010 Print "Rochelle"
020 Jump to 010

Experimenter:* This program will make the computer repeat the printing of the word "Rochelle". What do you think the output will look like?

Student: The computer will print the word "Rochelle" twice.

* Footnote: This one protocol, unlike all the others in this chapter, is a paraphrase of the original situation. For technical reasons, a verbatim transcription is not available here.
The answer is consistent with the ordinary sense of the word "repeat". It is also consistent with the student's prior experience, for in previous programs where there were no Jump statements and where at most there were two print statements, any program that repeated the same printout printed the same word twice. If we could characterize this student's schema for the purpose of the "Jump to" instruction, it probably would look something like this:

Schema 1: If the instruction is "Jump to n", then the computer does instruction number n.

Now the student was instructed to run program 1. When she did so, the output that appeared looked like this:

RochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelle

Student: I guess it keeps repeating until someone tells it to stop.

By her comment, the student has clearly learned something more about the "Jump" statement. To test what she had learned, we asked her to enter a new program into the computer and to predict its outcome.

Program 2:

010 Print "Hi"
020 Print "Rochelle"
030 Jump to 010

Experimenter: What do you think this program will do?

Student: Its first instruction is to print "Hi" so it will do "Hi", then it will (pause) there's no space, so it will just go "HiRochelle" for the second instruction. And then it will go back to the first instruction which was Print "Hi", so it will just write "Hi" until we tell it to stop.
We see from this example that the student has modified schema 1 to something like this:

Schema 2: Do each instruction in order unless the instruction is a JUMP-TO. If the instruction is JUMP-TO n, then continue doing instruction n until told to stop.

Note that this schema, even though incorrect, is perfectly consistent with everything this student has seen up to this point. She has derived her notion of sequential order of execution from earlier programs and has used it here to predict the first two elements of the output. From Program 1 she has seen that the JUMP-TO in that program caused the corresponding instruction to be repeated. Combining these results in Schema 2.

When the student was then asked to run Program 2, this was the result:

```
HiRochelle HiRochelle HiRochelle HiRochelle HiRochelle HiRochelle HiRochelle HiRochelle HiRochelle HiRochelle
```

Once again the result was not what was expected. Once again the schema for "jump" had to be modified.

Student: When you say jump to the first instruction, it will go to that and then I guess it goes to the second one and if there isn't a second one it will just keep repeating the first one. Otherwise it will repeat both.

This is a rather complicated and highly conditionalized notion, but it is perfectly consistent however with all examples she has seen. When she was asked to describe how the computer actually performed Program 2, she provided a correct line by line description. Her schema now might be characterized like this:
Schema 3: Do each instruction in order unless the instruction is "Jump to".
   If the instruction is "Jump to n", then begin doing instructions at number n.
   If there are no more instructions, stop.

Again, we tested her knowledge by asking her to type a specified program and to predict the result:

Program 3: 010 Jump to 030
            020 Print "Hi"
            030 Print "Rochelle"

Student: The computer will go to the third instruction and print "Rochelle" then to the second and print "Hi" and then to the third again and print "Rochelle".

The correct result is this:

Rochelle

Only the one word is printed, and then the program halts. Why did the student make the prediction that she did, when according to schema 3, she should make the proper one? Evidently she has other schemata about the operation of the computer. Many students seem to believe that every statement must be executed at least once, and this schema would explain the result here. If so, this causes a conflict with schema 3, which might possibly be resolved by a reversion to one of the earlier schemata for "jump". Whatever the reason, it was a simple matter for the student to modify her schema for "jump". When she saw that the output was the single word "Rochelle," she was readily able to determine why:
Student: The first instruction tells it to go on to the third and then there is no instruction to tell it what to do so it stops.

Now, finally, she seems to have a complete and correct schema for the "jump" instruction. When given two more tests, she predicted the results correctly.

Program 4:
010 Print "Hi"
020 Print "Rochelle"
030 Jump to 020

The predicted (and correct) result is:
HiRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelleRochelle

This shows that the student doesn't believe that each repetition needs to be the same.

Program 5:
010 Print "Hi"
020 Jump to 010
030 Print "Rochelle"

The predicted (and correct) result is:
Despite the fact that we have only illustrated the travails of one student working with one simple set of problems which only involved two instructions, the major points that concern us have been illustrated.
Towards a Theory of Learning and Teaching

The preceding sections have provided suggestions about the role of the tutor in the teaching process and, more importantly, about the way that the previous knowledge of the student shapes the process of acquiring new knowledge. First, we saw that a person can use old information to generate a plausible structure for a new situation, even when he has no actual knowledge of that new situation. Thus, in the example taken from the attempt to imagine how mayonnaise was created, our two subjects created mythical, incorrect, but logically plausible stories. Next, we saw that this same process of applying old schemata in new situations can cause unexpected difficulties when both tutor and student believe the student understands the lesson that is under study. Thus, in the protocol of the tutoring of a beginning cook by an advanced one, we saw that it took about 45 minutes to clear up the confusions that resulted when the student had mis-applied a former schema in a manner that seemed appropriate, but in actuality was not.

A good interaction between teacher and student requires that there be good understanding of the knowledge base of the student. Our examination of the interactions between an expert tutor and an advanced student showed just how difficult it can be to determine this knowledge. In that protocol, most of the time was spent while the tutor continually revised his estimation of the capability of the student. Even at the end of the entire session, it was not completely clear that there was complete understanding, although
sufficient knowledge had been acquired that the session turned out to be a valuable one for both participants.*

The analyses of the learning of FLOW programming once again made explicit the need to understand the development and use of the knowledge schemata. We saw how a student constructed rules for the operation of the concepts under study that were consistent with past experience, but that were not necessarily correct. Just as when our beginning student of the French White Sauces applied an inappropriate schema, at first it appeared to work satisfactorily. Only later, when either a novel situation or a specific question about the understanding arose did the programming student and the cooking student discover the error and then attempt to modify their schemata to more appropriate ones.

---

*Footnote: A major problem that we have not explored concerns the actual use of newly learned information. When a person acquires a new understanding of a subject that he had previously known in a different manner, what must happen before the new schemata replace the old? In the concluding section of the protocol between the expert and advanced cooks, the student had learned that a cold liquid need not be heated before it was added to a hot roux if it were added very slowly, heating all the while. The student commented, "You saved me 10 minutes." About a year later, when that "student" was preparing a Sauce Parisienne for a group that included one of us, he carefully heated the cold liquid to a boil before adding it to the roux. When he was reminded of his protocol and his comment about "saving ten minutes," he was able to recall the incident. Although the information was clearly part of his retained knowledge, it had made no impact upon his actual performance of the relevant tasks.
We begin to see what one must do to construct a model of the learning and teaching process. The examples of protocols, the consideration of schemata, and the structural network representation of a person's knowledge hint at possible ways to synthesize our knowledge together into a formal model that captures the process of learning. Unfortunately, the hint is not sufficient. Although we have the feeling that we are close to the development of a formal system, there still remain sufficient gaps in our knowledge that we cannot yet complete the job.

A Structural Network Analysis

A start towards a formalization of the process of learning comes from a consideration of structural networks. We can distinguish the structures that a student starts with from the structures that the teacher wishes him to have. Now consider the difference between the two: the difference structure. This is what is to be taught. One approach to the specification of learning and teaching is to systematize the descriptions of these structures and then formalize the path of teaching.

To illustrate the points, we will take a simple network structure and show how it may be analyzed. Figure 2 shows the prerequisite structure for the general topics covered under the title of French white sauces. The formal contents of each node in Figure 2 are not important for this analysis: all that matters at this point is the interrelationship among the nodes of the diagram.

Insert Figure 2 about here
Figure 2

- White stock or hot milk (S2)
- White roux (S1)
- Basic white sauces (BWS)
- Purpose and uses (S3)
- Egg yolk techniques (S4)
- Butter-cream-egg yolk enrichment (S5)
- Enriched white sauces (EWS)
- Error correction (S8)
- Flavorings (S7)
- Cheese gratinee (S6)
A: THE CONTENT NETWORK

B: THE STUDENT'S KNOWLEDGE

C: THE DIFFERENCE NETWORK

Figure 3
In this figure, the arrows that interconnect the nodes represent ordering of concepts. That is, if some node A has an arrow leading from it to some other node B, then the conceptual knowledge structures represented by node A are prerequisite to an understanding of the conceptual structures represented by node B.

The Difference Network

It is important to distinguish between the networks that represent the subject content that is the topic of the learning and the amount that a particular student knows. Clearly, the concepts of teachability and input value depend critically upon the amount known by the student. What is needed is a map of both: the knowledge of the student (the student network) and the content network. What is to be taught is the difference network: that part of the concept network that is not a part of the student's knowledge. This is illustrated in Figure 3.

We define a teachable node as one that has no prerequisites (after the work of Sayeki and Ura 1966). Figure 3C shows a situation where there are immediately 3 teachable nodes: nodes S4, S7, and S8. As the teachable nodes are taught, they are added to the student network and, therefore, deleted from the difference network.

Linear and web teaching. The order in which new material is presented is clearly of some importance. A simple structural analysis of the difference network indicates that, at the extremes, there are two widely differing styles of presentation. One is the linear, orderly presentation of the developing structure. Each new piece of information is added in
sequence. This might be called **linear teaching**. It is the style that characterizes most lectures or textbooks: each item follows from the first. This is illustrated in Figure 4. The other procedure would be to try to give a general overview of all of the components of the topics that are to be discussed, then review them all in more depth, and then repeat again, with each pass through the material deepening the depth of discussion and the amount of detail. Because this procedure acts as if it is applying a coarse web of structure, and then repeatedly going over the web again and again to fill it in with more finely meshed structure, Norman (1973) called this **web teaching**. It is illustrated in Figure 5.

Web and linear teaching would appear to be at the extremes, so that almost all real examples are not pure instances of either. But the different characterizations appear to be useful. Web teaching seems to be a promising approach, for it is clear that in linear teaching the soundness of the structure depends upon the soundness of each node in the sequence. In web procedures, there is sufficient redundancy, that ill-learned concepts will not seriously affect the whole. Web teaching, however, turns out to be less natural and more difficult to do than is true of linear teaching.
Figure 4
Then

Later

Web Learning

Figure 5
The Learning Process

When a student learns a new complex topic, information that takes weeks, months, or even years to acquire, one of the critical aspects of the learning situation has to do with the creation and use of appropriate schemata: the set of rules and structures for dealing with complex situations. When a particular problem is presented, the student appears to begin by relating the one schema that he believes either will work directly or is close to the desired result. Then he proceeds to modify the schema by inserting the particular aspects of the present situation into its framework, as well as modifying the framework as is necessary. This modification process often results in new schemata that can then be added into the set of acquired knowledge. Thus, each new problem solved by a student potentially can add to his general knowledge.

The selection of appropriate schemata for a problem is only one aspect of learning and performing. A second aspect is that of comparison of the actual result of applying a schema with the expected or desired result and then, when there is a discrepancy, modifying the underlying conceptual frameworks. The discrepancies between the actual outcomes and the desired outcome may in fact constitute the most important information that a student receives, for not only do these discrepancies point out new issues about the topic that is being learned, but they also lay bare the structure of the student's own conceptual framework, allowing him thereby to modify them. Thus ability to modify one's own knowledge structure, to take existing schemata and transform them to new structures, is clearly of critical importance.

The learning process, then, must consist of several different cognitive skills. There must be a process of interpreting new information in terms of known concepts, constructing structural representations that can be used
at later times. A good part of learning may consist of the modifications and additions to previously acquired conceptual structures to allow them to form new interpretations of information. Much of what we learn consists of acquiring useful or efficient ways to interpret the input. Learning to solve a new type of problem, for example, often does not involve mastering a new problem solving technique, but rather involves learning to conceptualize the new problems in ways that permit its solution by previously mastered techniques.

In our view, the memory system is organized around conceptual frameworks—schemata—all incorporated within an active structural network of interconnected information units. When a problem is posed to a person, or for that matter, whenever some new event is experienced, it must be possible to make contact with appropriate memory structures, to select and modify the appropriate ones, and then to apply them to the situation. The retrieval process must operate quickly, but with flexibility, for most experiences will be similar to, yet somewhat different from previous ones. The selection and modification processes are clearly important components of the learning process.

Finally, it must be possible to predict the results of operations without actually performing them. Thus, a learner is also a simulator, able to perform mental simulations of the solutions to problems before him, and thereby able to modify the selected schemata on the basis of both simulated and actual results. Of course, the simulated performance will often differ from actuality, for not only may the conceptual structure of the student be deficient, but the actual performance of the mental simulation may exceed a person's cognitive capabilities (the most common failure is usually short-term memory). Nonetheless, the ability to predict the result of a postulated course of action is an essential part of the cognitive repertoire.
Overview

In this chapter we have examined the process of learning large complex subject matters, asking about the ways by which teacher and student communicate the necessary knowledge structures to each other, asking how a student comes to select an appropriate paradigm for solving a problem and to revise inappropriate ones. We have examined numerous protocols from several different learning situations, and we have made the first few stumbling attempts to put together the pieces into a cognitive theory of learning and teaching.

We do not yet know how to put together the missing pieces. We still lack a method for combining together and working with the components of learning. The model of the student is incomplete. The component parts for learning a programming language hang in the air, disconnected and unstructured parts of what should be a consistent, cohesive conceptual structure. We need to formalize our statements, and to devise theories that guide our understanding, that yield observable predictions, and that help advance our further studies of the learning and teaching process.

It is obvious that we are still far from the goal.

Surprisingly, though, we remain hopeful. We have found that the view of the learning process as one of communication, with all participants attempting to form new mental structural representations that will account for the information which they experience, offers insightful probes into the nature of both tracing and learning. That our approaches are far from being complete can be seen as a challenge for the future.
Bibliography


