DOCUMENT RESUME

ED 278 358 IR 012 458

AUTHOR	Lesgold, Alan M.
TITLE	Human Skill in a Computerized Society: Complex Skills and Their Acquisition.
INSTITUTION	Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY	Department of Education, Washington, DC.; Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE	86
NOTE	13p.
PUB TYPE	Journal Articles (080) -- Viewpoints (120) -- Reports - Research/Technical (143)
JOURNAL CIT	Behavior Research Methods, Instruments, & Computers; v16 n2 p79-87 1984
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Apprenticeships; Artificial Intelligence; *Automation; *Cognitive Psychology; *Computers; Correlation; *Epistemology; Interpersonal Competence; Learning Theories; Longitudinal Studies; Path Analysis; *Professional Training; *Radiology; Research Needs; Research Tools

ABSTRACT

This paper discusses some of the issues raised for cognitive psychologists by the computer revolution together with the role that psychologists with computer training ought to play, especially in the study of how people acquire complex skills. The issues addressed include: (1) the competition between humans and intelligent machines; (2) the destruction of apprenticeship opportunities by automation; (3) the flexibility of human intelligence and transfer of skills to different situations; and (4) the use of the computer as a research tool in the study of learning and thinking. A research project on the acquisition of radiological diagnostic skills is then described as an example of research on the learning processes involved in acquiring expertise in complex domains. This description includes the types of skills involved in diagnostic radiology, the methodologies used, and the findings of several related studies. A discussion of the need for the large scale longitudinal studies now possible with the use of computers to verify theories about complex skill learning concludes the paper. Several tables present data from the project and five references are listed. (DJR)

LEARNING RESEARCH AND DEVELOPMENT CENTER

1986/5

HUMAN SKILL IN A COMPUTERIZED SOCIETY: COMPLEX SKILLS AND THEIR ACQUISITION

ALAN M. LESGOLD

2

3

HUMAN SKILL IN A COMPUTERIZED SOCIETY: COMPLEX SKILLS AND THEIR ACQUISITION

Alan M. Lesgold

Learning Research and Development Center
University of Pittsburgh

1986

Reprinted with permission from Behavior Research Methods, Instruments, & Computers, 16(2), 79-87.

The research reported herein was conducted at the Learning Research and Development Center, and was supported by the Office of Naval Research Personnel and Training Group, The National Institute of Education (NIE), U. S. Department of Education and the Air Force Human Resources Laboratories. The opinions expressed do not necessarily reflect the position or policy of the sponsors and no official endorsement should be inferred.

Behavior Research Methods, Instruments, & Computers
1984, 16 (2), 79-87

SESSION I
PRESIDENTIAL ADDRESS

Russell M. Church, *Presider*

Human skill in a computerized society: Complex skills and their acquisition

ALAN M. LESGOLD
University of Pittsburgh, Pittsburgh, Pennsylvania

Some of the issues posed to cognitive psychologists by the computer revolution are explored. These include competition between humans and machines, destruction of apprenticeship opportunities by automation, the issue of flexibility of intelligence, and related issues. This is followed by a discussion of a program of work on the acquisition of a complex skill, radiological diagnosis. Differences and similarities of that domain to other recently studied domains are discussed. The concluding section discusses the need for longitudinal approaches to verifying theory about complex skill learning.

In preparing for this talk, I studied the discourse structure for presidential addresses to see what one is supposed to do. The following rule emerged as the first rule for an address generator:

presAddr → visionaryStatement + talkAboutYourResearch

So, I first have some thoughts to offer about the computer revolution and the role we psychologists with computer training ought to play. After that, I'll turn my attention to some work I have been doing on the acquisition of complex skills, an endeavor in which people and machines now compete.

IMPLICATIONS OF THE COMPUTER REVOLUTION FOR PSYCHOLOGY

It is almost a cliché by now to proclaim that we are about to undergo a socioeconomic revolution in which information will be a major item of trade and will be packaged in small containers, not human heads. Indeed, for the members of this Society, who have relied on computer software for many aspects of their work, the revolution is almost at hand. At least a few of us have been in positions in which what we needed most was a particular software system, not a human assistant or a piece of hardware. The computer has changed our lives in a manner reminiscent of the effects of the printing press or the steam engine. As students of human behavior and cognition, surely we need to attend to the effects of the computer on our fellow human beings, too. Indeed, the problems we pick to study will be driven by the needs of society as always, and society now needs to understand and prepare for the computer revolution.

Some effects of widespread computer usage are now painfully apparent. Many factory workers are being replaced by robots, and others are finding their skills devalued as they compete with robots. I am concerned in this presentation, though, with a second displacement that is also beginning, in the white-collar work force. Many routine jobs were necessary to the running of large-scale companies but were also a cause of overly complex management structures. Competitive companies will gain greatly by trimming management positions, especially those that contribute information processing skill only at the expense of introducing further layers of authority into the management hierarchy. People who thought they were set for life will be out of work and in need of retraining.

Obviously, there will be Luddite reactions to these events, and thus the effects of the computer revolution

This address reports work supported by the Office of Naval Research Personnel and Training Group, the National Institute of Education, and the Air Force Human Resources Laboratory. The research results all come from collaborations with Bob Glaser, Dale Klopfer, Harriet Rubinson, Paul Feltovich, Carol White, Lloyd Bond, Lauren Resnick, Kathleen Hammond, Christopher Roth, and other LRDC colleagues. I take complete responsibility for the assertions made in this paper, with which neither sponsors nor collaborators necessarily agree, but the work would have been impossible without them. My mailing address is: Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA 15260.

will be slowed slightly. However, the changes I just mentioned are already starting to happen, and we can make an important contribution by studying issues related to training of valued workers for the computer age. This concern has motivated much of the research my colleagues and I have been doing. We have been trying to better understand the nature of complex cognitive skills and how they are acquired. As artificial intelligence techniques improve, there will be both an enlarging of the total range of intelligent tasks people and their machines can perform and some natural selection of human and machine intelligence for different productive roles. This will result in a number of situations that are problems for society, as well as in opportunities for psychologists to do useful research. Here are a few examples of what I mean.

Humans and intelligent machines will compete. People select into different occupations in part on the basis of their ability or inability to learn from inadequate instruction. Indeed, some have defined *intelligence* as that ability. Instruction in many schooling and training environments is inadequate for at least some students. As a result, some people, lacking appropriate learning skills, are likely to end up in jobs that require minimal training. These are exactly the jobs most likely to be automated away. The dilemma I see is that those least able to learn from inadequate instruction will be most likely to have to undergo such instruction frequently, as they move from one job to another under chase from automation. This suggests that efforts to improve instruction and training in order to assist poor learners will be worthwhile.[1] It also suggests that we might devote resources to research on the skills of learning and on efforts to teach those skills. A panel that Fred Reif and I chaired has recently published recommendations for federally funded research on possible roles for intelligent computer systems in solving these problems (Lesgold & Reif, 1983).

Automation destroys apprenticeship opportunities. Most knowledge is acquired after a combination of learning by being told and learning through practice. One solution for those who have trouble learning from lectures is the apprenticeship, in which learning is gradual and combined with practice. The apprentice starts with the simplest chores and gradually is allowed to assume more and more of the range of performances that characterize expert or master status. Unfortunately, progress in building intelligent computer systems also proceeds from the simple to the complex. In many skill areas, some of the apprenticeship opportunities are being destroyed by advancing automation.

One example involves the training of electronics technicians. Once, technicians did serious problem solving from the first day they were on the job. They had to understand how the devices they were servicing worked and how to diagnose them, and this understanding was reinforced by regular opportunities to solve problems. Today, new technicians are prevented from having many of the learning experiences that were once possible. They are taught procedures that test for faults in order of probability of occurrence, not in groupings that relate to the structure of the system. They use diagnostic programs that can suggest which circuit board should be replaced. These programs are likely to be accurate most of the time, but their suggestions provide no clues about diagnosis strategy. Often there is not even an explanation of why a board should be swapped. "Swaptronics" is economically effective so long as there remain master diagnosticians to take over from the poorly trained first-line people. I worry about how we will build a next generation of such masters, though, when the on-the-job learning environment in which they are being trained is so impoverished.

Human expertise is flexible. Following the lead of expert systems developers, most psychologists studying cognitition have concentrated on specific bodies of expertise and on how that expertise is acquired. However, if we are concerned with areas in which people are especially important, then we should be taking a broader approach. Expertise captured in a thousand-rule production system is likely to be very specific, very brittle. Human expertise, in contrast, is often more flexible, more capable of transfer to new situations. We need to understand this flexibility and find ways to build it in the course of training people in skills. We need to emphasize transfer in our studies of learning.

Skills may not lead to virtues. Finally, we must recognize that cognitive psychology is generally ignoring an area in which machines are least likely to tread. Machines will acquire intellectual skills; it is less likely that they will acquire certain social skills and virtues. When we glibly say that automation leads to expansion of the service sector of our economy, we are ignoring the extent to which interpersonal skills may be more important to an average person in an automated society than modest intelligence. A look at our governmental institutions and at the problems of our larger businesses quickly tells us that research on leadership skills and on collaboration skills is sorely needed. Perhaps as computer scientists learn to build parallel system architectures in which multiple intelligent entities work together, they can spin off a few metaphors to energize new work on social skill and its acquisition. After all, a few computer metaphors drove cognitive psychology a long way.

Computer tools will allow us to do such work. As psychologists who are a bit ahead of the pack in adapting to the availability of computer assistance, we are the people who should be pioneering work on these problems. We now have computer tools that allow us to study learning and thinking on an appropriate scale. Intelligent machine extensions of our theoretical reasoning and mental modeling capabilities, database systems, and graphics tools, among others, make it possible for us to attempt to do good science that is also relevant to serious social concerns. This Society exists largely to promote the use of such tools, hopefully in research that can eventually help humankind learn to live in automated settings.

ACQUIRING EXPERTISE IN COMPLEX DOMAINS

Continuing to follow the structure I mentioned at the beginning of this talk, it's now time to remove the mitre and put on the lab coat. I want to talk about some research on the acquisition of radiological skill and then briefly about an approach to studying learning that I think is very promising. This work relates to the point I just made about the need to understand expert flexibility and how it is acquired.

Radiological Expertise

In the last few years, several colleagues and I have been trying to understand how the skill of radiological diagnosis is acquired (Lesgold, 1983a; Lesgold, Feltovich, Glaser, & Wang, 1981). Radiology was chosen for two important reasons. First, it requires the ability to access a complex mental representation (of a patient's innards) in response to highly ambiguous visual displays and the ability to tune that representation to specific cases. Second, it involves several domains of knowledge that have very complex interfaces. Visual patterns learned by prototype learning do not map cleanly onto disease variations; often two examples of the same disease look more different than two examples of different diseases. Similarly, the interrelationships among structural anatomy, the physiology of normal function, and the nature of disease are also complex. To the knowledge engineer, it looks like thousands of separate rules are required to capture such a skill. However, having only such a large and cumbersome rule set would leave an expert with very brittle knowledge, knowledge that does not adapt readily to new situations. We were interested in the general flexibility of radiological expertise as well as specific diagnostic knowledge.

I can't provide a complete account of this work in this talk. Rather, I want to present a few highlights, in the form of assertions presented with snippets of the empirical support we have seen for them. The empirical support comes from two sources. First, we have conducted a naturalistic study of expert and resident radiologists, in which we took verbal protocols as they performed a slightly modified version of the diagnosis work that constitutes much of their daily activity (see Lesgold et al., 1981, for details). We later conducted modified study in which we tried to verify what subjects were seeing in our test films by having them draw on the film itself. This way, they could indicate what they were seeing in the film. We also prompted them to draw structures that pilot work had shown were sources of possible skill differences. Here are a few of our conclusions.

The forest-trees problem.

Novices have difficulty constructing representations in which the same stimulus feature is mapped onto two different objects.

The first assertion is a variation on the old problem of not being able to see the forest for the trees. We have strong evidence that radiology residents can only assign a particular region of an x-ray film to one piece of structure. As a result, they often misapprehend the true shape of abnormalities. In some cases, surrounding structures "recruit" parts of the film that also include an abnormality. This makes the abnormality seem smaller than it really is, leading to misdiagnosis. In other cases, abnormal regions grow by recruiting normal regions they abut. We saw both phenomena in the drawings and verbal protocols of our subjects.

A particularly nice example comes from a film showing a collapse of the middle lobe of the right lung. This film showed a large, sail-shaped wedge in the part of the lung that is near the heart (the wedge is produced by the increased density of the lung material after it collapses). Unfortunately, even though this pattern is a relatively common variant of the classic pattern displayed by such a collapse, not every instance of the pattern is caused by collapse. Certain tumors and several other disorders can produce the same appearance. In fact, our residents sometimes interpreted the sail-shaped region in a way that precluded a conclusion of collapsed lung.

Figure 1 shows overlaid tracings of the sail-shaped region drawn by several expert and resident subjects. The experts saw a region bounded by the pulmonary artery, and it was the right size to be considered as a collapsed lung. In fact, they only saw a small collateral of the artery as actually entering the region of high density. The residents tended to follow various bits of noise and to recruit much of the critical region into their representation of the pulmonary artery, leaving

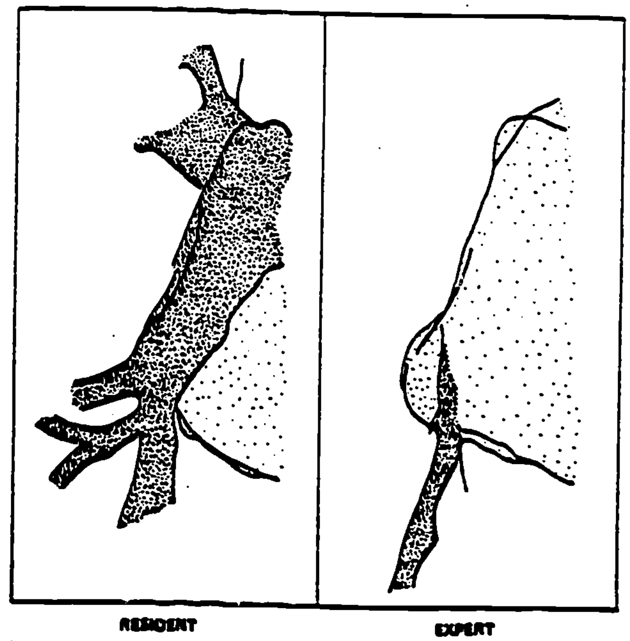

Figure 1. Overlaid tracings of the sail-shaped region drawn by several experts and several residents.

very little to be an unaccounted-for abnormality. They were incapable of seeing both blood vessel and some other density in the same space. Left with a small mass to explain, they tended to call it a tumor. In capsule summary, residents have not yet formed the automatic and complex recognition skills that would keep diagnoses like these from becoming embedded-figures tests in which the Gestalt laws of general stimulus organization play a bigger role than recognition for specific features.

The don't-bother-me-I'm-busy problem.

Experts are opportunistic planners. They have rich, flexible disease schemata. In contrast, novices have classic, less tunable schemata and are more likely to maintain bad film interpretations in the face of discrepant evidence from the patient's clinical history.

This second broad assertion could itself be the topic of an hour or two of discussion. Instead, I will present only some hints of what the data we have been gathering suggest. First, there is the issue of schema flexibility. Earlier, I discussed the issue of flexibility versus brittleness in expertise. Our experts showed a number of examples of flexibility in their diagnoses. They were fast because they were able to quickly recognize general film features that could trigger an appropriate schema to guide further viewing, a matter I'll turn to shortly. On the other hand, when some aspect of what they saw conflicted with the invoked schema, they were able to tune the schema or abandon it, as appropriate.

Table 1 provides some protocol extracts for an expert who showed this capability. The case concerned a patient who was currently healthy but whose right upper lung lobe had been removed a decade earlier. Because organs shifted to fill the space left over from the missing lung lobe, the film had many abnormal features—things in the wrong place, densities that were unusual. The film seems to show a collapsed lung, when actually the lung lobe was surgically removed. Even some experts were fooled and tended to diagnose the patient as having a disease-induced lung collapse. Most residents reached a similar conclusion, although some concluded that the patient had congestive heart failure. However, one expert put on a virtuoso performance.

Almost immediately after his 2-sec first look at the film, the expert drew a conclusion that many subjects failed to reach. He decided that the film showed a chronic problem, one that did not suddenly appear. He then tried to find an appropriate schema into which he could fit what he was seeing. To make schemata fit, he tried applying rotational transformations, considering the possibility that the subject had not been facing directly toward the x-ray plate. However, that didn't help. Finally, he began to let the collapsed-lung schema guide his thinking.

However, he didn't just accept that schema; he kept trying to both test and elaborate it. Soon, he noticed the irregularities of some ribs that had been broken a decade ago by the surgeon. He then very quickly changed schemata and cracked the case. Even then, though, he kept testing his new conclusion and seeking a hybrid of the collapse and lobectomy schemata. This was quite reasonable, since a collapse of a lobe has many of the same long-term effects as removal of a lobe.

Throughout, the expert displayed flexibility, the ability to push, tune, and retreat from a schema that was guiding his thinking. In contrast, some residents held to their pessimistic diagnoses of congestive heart failure even after we told them that the patient was perfectly well, had had a similar-appearing film done a year ago, and had had a lobectomy performed long ago. They still said she had heart failure, even though we had just told them she felt fine. They seemed unable to test the schemata or the specific hypotheses they had invoked or to take account of new data we provided.

Does learning always mean improvement?

Trainee performance in nonclassical cases may not vary monotonically with experience. This is because the tuning of diagnostic schemata both improves their potential accuracy and temporarily increases the conscious processing load they impose.

Table 2 shows another finding, one a bit more surprising for us. Our naturalistic studies involved films that were quite difficult. Indeed, it is close to impossible to make an absolutely definite diagnosis from some of our films without additional data. With several of these

Table 1
Expert Diagnostic Performance

Something is wrong, and it's chronic: "We may be dealing with a chronic process here ....
Trying to get a schema: "I'm trying to work out why the mediastinum and the heart is displaced into the right chest. There is not enough rotation to account for this. I don't see displacement of fissures.
Experiments with collapse schema: "There may be a collapse of the right lower lobe but the diaphragm on the right side is well visualized and that's a feature against it ....
Does some testing; schema doesn't fit without a lot of tuning: "I come back to the right chest. The ribs are crowded together .... The crowding of the rib cage can, on some occasions, be due to previous surgery. In fact, ... the third and fourth ribs are narrow and irregular so he's probably had previous surgery ....
Cracks the case: "He's probably had one of his lobes resected. It wouldn't be the middle lobe. It may be the upper lobe. It may not necessarily be a lobectomy. It could be a small segment of the lung with pleural thickening at the back.
Checks to be sure: "I don't see the right hilum ... [this] may, in fact, be due to the postsurgery state I'm postulating .... Loss of visualization of the right hilum is ... seen with collapse .... ."

Table 2
The Get-Worse Phenomenon (in Percent)

Case	1,2 Years	3,4 Years	Expert
I.3	100	57	100
I.8	27	14	80
II.9	50	0	75

films, we noticed that subjects with 3 or 4 years of training did worse than 1st- and 2nd-year residents; performance was not a monotone function of training. Now, we had only 23 subjects in this study and only 10 films, so this finding should be treated as very tentative. Indeed, our efforts to replicate it in the training research I will discuss in a moment have thus far produced plateaus but not reversals. However, the finding is consistent with a view of skill acquisition that I happen to like, so I will press on.

The interpretation we have been toying with in our research group[2] is that the apparent accuracy reversal is due to a progression from probabilistic to reason-supported decision making. New residents learn classic schemata first. When they examine a film, they try to fit it to a schema they have already learned. Generally, this is equivalent to making the high-probability guess for a set of noticed surface features. As they become more skilled, they are able to do deeper analyses of the feature details that do not quite fit the classic schema—decision making becomes more rational and complete. However, until they become more proficient at this deeper level of analysis, they make errors, perhaps even more than would be made by simply picking the high-probability bet based on a few features. This analysis would lead us to expect a nonmonotone relationship between experience and accuracy for more complex films.

Planning: Knowing where to look and what to see. *The precise, rapid recognition skill that characterizes expertise involves interactions between higher and lower processes. It is not purely top-down or bottom-up.*

Experts recognize constraints on problem solution early but defer decisions until they are necessary.

Again, I've made a very broad claim, and we have only begun to gather an appropriate range of data. Let me present three bits of data that help to support the assertions I've just made. First, I want to show you more evidence that expertise involves more rapid shaping of diagnostic behavior by a preliminary schema. The data I've already presented also showed related effects. Second, I will show you evidence that such schemata are triggered very early, even in the first 2 sec of viewing. Third, I'll show you evidence that experts know where in the film to look. Schemata triggered during the first seconds of scanning a film guide their subsequent search of the film.

The first piece of data is in Table 3. For a film involving chronic lung disease, we tabulated the proportion of subjects in each condition who mentioned findings that were either consistent or inconsistent with a chronic lung disease schema. As you can see from the table, the consistent findings were more probable in experts' protocols, and mention of some findings that are neutral or inconsistent with the chronic lung disease schema was more frequent for residents. Now, you may wonder about an M.D. with an internship who says the hilar area is overly prominent when it is actually less prominent than usual. In fact, it is not that easy to decide what is hilum and what is extra mass, and what is other structure. This decision is what a schema helps drive.

The second piece of relevant data is meant to show that schemata are triggered earlier in experts. We looked at protocols from another chronic lung disease case (not the same one as in the previous figure) and tabulated the frequency with which various schema-consistent features were mentioned after seeing the film for only 2 sec. As you can see from Table 4, experts were more likely to report these features after only 2 sec than were residents; indeed, they seldom failed to report any of the four features listed. Even in a final diagnosis after extensive viewing, the residents are not all that likely to include chronic lung disease in their diagnoses.

The third piece of relevant data is meant to show that schemata help determine which film regions the expert will examine further. Our collapsed-lung film had many general signs of chronic lung disease along with the signs of collapse. One clear indication of schema-driven

Table 3
Early Interpretations Show Schema Invocation (in Percent)

Feature Mentioned	1,2 Years	3,4 Years	Expert
Consistent with Chronic Lung Disease Schema			
Small heart	18	14	40
Heart shifted to left	0	14	40
Right hilum less prominent	0	14	60
Right hilum medially displaced	0	0	40
Neutral or Opposed to Schema			
Normal heart size	73	71	40
Density in heart region	36	43	0
Normal hilar appearance	73	57	0
Increased hilar prominence or hilar density	45	43	0

Table 4
Two-Second Encodings of Film With Tumor and Chronic Lung Disease

Observation	1,2 Years	3,4 Years	Expert
Chronic Lung Disease or Emphysema	18	29	80
Hyperexpanded lung	27	57	100
Low diaphragms	18	43	100
Small heart or narrow sup mediastinum	0	29	80
Final decision includes Chronic Lung Disease	36	57	100

Table 5
Regions Mentioned for Collapsed Lung Film (in Percent)

Region	1,2 Years	3,4 Years	Expert
Lungs and Pleurae	41	41	93
Other Regions	60	59	8
Total	100	100	100

processing would be for looking, and therefore noticing and ascribing of features, to be concentrated in the lungs. Table 5 shows that, for experts, almost all reports of things seen, normal or abnormal, were about the lungs and the pleurae (which are the sacs that surround the lungs). In contrast, residents actually talked more about nonlung observations than about the lungs themselves. Again, we see that expertise involves procedures that are schema driven. Of course, when the schema is inadequate, or overly classic, experts also are able to patch it, to tailor a generic set of procedures to the case at hand.

The Need for Longitudinal Research on Training

I view the radiology work as an accommodation of a perceptually driven domain into a scheme that other researchers had already proposed. We've said a few things that are new about expertise, but we've mainly shown how previous findings by others could be extended to a very different domain. What we're now trying to do is to go beyond studying expertise to studying its acquisition. Thus, we've created an artificial domain in which the films are based upon a computer model and a replicable and controllable set of abnormal perturbations are performed to produce films of "disease."

Pseudo-x-ray training studies. Figure 2 shows an example of a film that we're calling "normal," and Figure 3 shows our version of a collapsed lung. The severity of the "diseases" we employ is varied by using a numerical parameter to determine the extent of a perturbation. For example, a raised diaphragm on the patient's right side (viewer's left side) is one symptom of collapsed right middle lobe. We simply make the extent of perturbation of the diaphragm from its normal location a function of the severity parameter. Basically, the computer model has a "political" map of the chest showing organ boundaries and a "physical" map of the chest showing the gray levels those organs project onto an x-ray film. The levels are approximately authentic, but medical authenticity is irrelevant to our experimental approach. A "disease" is manufactured by specifying a set of features, where each feature is a distortion of either a boundary in the "political" map, a density level in the "physical" map, or both. The exact shape of a perturbation, and the way in which density levels are represented by combinations of black and white dots, is randomly varied over instances of the film types. In addition, we can vary "exposure" by multiplying all the density values by a constant. The programs that develop files of graphics commands for these stimuli are simple and relatively short FORTRAN programs.[3]

We've now given a group of college students extensive training in recognizing four different pseudodiseases in these pseudo-x-ray pictures. They've thus far had about as many hours with us as they would get in a two-credit course—not expert-level training, but much more than the usual learning experiment. Some have been taught mechanistic explanations for the four diseases and the film features that signal them, whereas others have only been taught the physical features to watch for. In the training, we ask them to mention any disease features they see in a film, and after each diagnosis, we ask them to tell us which features were key sources for their decisions.

We are trying to understand the retention and transfer effects of the two forms of training, which tend, in the short term, to produce approximately equal recog-

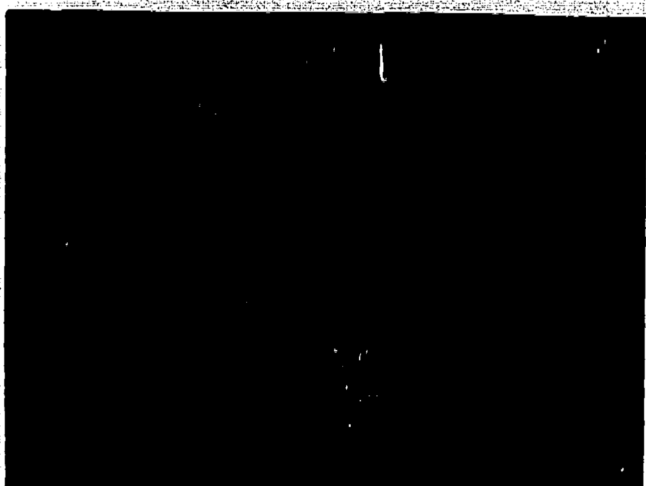

Figure 2. "Normal" pseudo-x-ray stimulus.

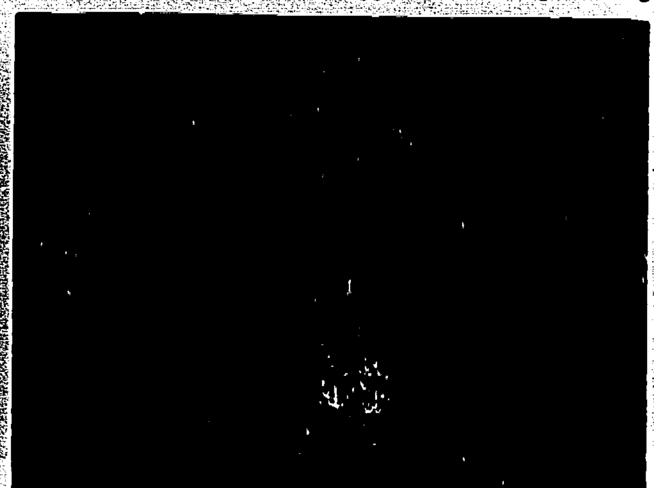

Figure 3. Pseudo-x-ray of "collapsed lung."

nition accuracy. There are some differences that are already apparent. For example, there is an apparent tendency for deep knowledge about our pseudodiseases to result in somewhat more likelihood of noticing features that are central to an explanation of the physical dynamics of a disease and somewhat less likelihood of mentioning features that are, in my opinion, especially physically salient but not reflective of an immediate effect of the basic physical disease mechanism. This is strikingly close to the sort of schema-driven processing effect we got in the studies I just told you about.

The real payoff, I think, is just beginning to appear from this sort of research. We can trace the longitudinal pattern of sensitivity to specific features in different films. For example, Figure 4 shows a matrix of all the features mentioned for a specific type of film over all its occurrences in the course of 8 days of 1-h/day training. The rows represent the instances of the film type being presented, from first to last. The columns represent specific features that were mentioned. Features with the letter M in their code were actual features of this mitral stenosis film, as was A6. You can see that features M1, M4, and A6 were present from the beginning, but that it took some time for the subject to learn to notice P1M5 and P2M2, which were also found in embolism films. Figure 5, in contrast, shows an example of seeing stuff that isn't there, A features in an E film, during the first days of learning. We are just starting to look at this sort of longitudinal data.

Features Reported for Moderate Mitral Stenosis With Normal Exposure by Subject U

FILM	Diagnosis	P1M5	P2M2	M1	M3	M4	A6
140.	Mitr Sten	+				+	+
141.	Atr Sept D						+
142.	Mitr Sten	+				+	+
493.	Mitr Sten	+		+		+	+
518.	Mitr Sten	+		+		+	+
496.	Atr Sept D						+
510.	Mitr Sten	+				+	+
519.	Mitr Sten	+				+	+
498	Mitr Sten	+		+		+	+
489.	Embolism		+		+	−	
511.	Mitr Sten	+			+		+
491.	Mitr Sten	+		+	+	+	+

Figure 4. Longitudinal trace of features mentioned for normal exposure, moderate mitral stenosis over eight training sessions.

Since we can control the relationship between the features we put into our pseudodiseases and what the subject is taught about them, we have the basis for quite a range of longitudinal studies of the effects of theoretical versus more practically specific forms of instruction on recognition and diagnosis. When the possibilities of both near transfer, to milder forms of a disease, and far transfer, to the same disease in different loci (e.g., collapse of different lungs), are considered, it is apparent that this paradigm should be very helpful in studying the effects of different forms of training.

As a methodologically driven organization, I think it is important to be sensitive to the role the computer has played in making such research possible. The mass of data from our longitudinal studies requires a database system. The stimuli are computer generated, and it would be impractical to do them by hand. The modeling work that has led to this whole domain of research rests on a variety of tools developed by artificial intelligence researchers and psychologists who interact with them. It is almost a cliché for a scientist to say that his ability to see a bit further is due to his standing on the shoulders of giants. It seems that we're now able to find both human and machine giants to give us a boost.

Reading skill. With these new tools, I argue that we must start to do studies on the scale that matches the kinds of learning that are important in our society. Just as the physicist sometimes requires a Fermilab or a tokamak to study basic physical forces, so we must be prepared to scale up our experiments to study learning that takes place over hours, days, even months and years. We have the ability to conduct extensive experiments with computer-controlled conditions and com-

Features Reported for Moderate Emphysema With Normal Exposure by Subject K

FILM	Diagnosis	E1	E2	E3	E4	E5	E6	P2M2	P3	A1P5	M1	M3	M4	A2	A3	A5	A6
106.	Mitr Sten	+					+										
105.	Collapse			+							+		+			+	+
107.	Atr Sept D	+						−		−				+	−	−	+
336.	Emphysema	+		+	+	+	+										−
331.	Emphysema	+	+	+		−	+										
349.	Emphysema	+	+	+		−	+										
351.	Emphysema	+	+	+		+	+										
354.	Emphysema	+	+	+		+	+										
352.	Atr Sept D							−	−			−					
339.	Emphysema	+		+		+	+		+								
332.	Emphysema	+	+	+		+	+										
338.	Emphysema	+	+	+		+	+										

Figure 5. Longitudinal trace of features mentioned for normal exposure, moderate emphysema over eight training sessions.

puter-sited databases, and we should use this ability. I want to reinforce this point by presenting very briefly another example of the utility of longitudinal research, this time research on how reading is taught and learned.

For decades, debate has raged over how to teach reading. Designers of basal reading series argue that one should emphasize reading as language, quickly getting children to read whole texts that are meaningful, teaching words as holistic recognition units. Others argue that children should concentrate on learning the basic spelling patterns of which words are composed, so that they can "sound out" a word that they haven't seen in print but may recognize in speech. The issue is even political, with conservatives favoring phonics more often than liberals do. There are a few data, but not many. A couple of decades ago, phonics-rich programs had a slight edge in levels of reading test scores they produced. However, teachers select curricula, so there are all sorts of confoundings of curriculum with teacher qualities, especially in the period when the original research was done, a period during which use of phonics was somewhat restricted to tough, no-nonsense teachers.

Lauren Resnick and I decided a number of years ago that the way to settle some of these issues was to watch the changes in word processing skills and overall reading ability as children advanced through the primary grades. We designed a longitudinal study that followed children through the primary grades in one school that had a phonics program and another that used a standard basal program. We've published some of our findings (Lesgold & Resnick, 1982), and more will be out soon. Today, I can only give you a very brief sense of the use of the longitudinal approach.

We were able to do correlations and various forms of path analysis among the many measures we took, with each measure made at several points in time, on the same subjects. These measures allowed us to conclude, for example, that in both basal and phonics-loaded programs, facility in word recognition is an important precursor of later reading comprehension facility. This supports the idea that general skills of word recognition, as opposed to recognition for lots of separate words only, are important. On the other hand, by second grade, and certainly by third grade, the best predictor of reading success later on is current reading comprehension ability, not word automaticity. There is no evidence that those who are behind in reading should be relegated to phonics drill *instead of* reading sensible text. The general pattern of our data is that word processing facility provides a boost early in the course of training. We suspect that word recognition practice could be useful later if it can be provided without prejudice to time spent on serious discourse processing.[4]

We also were able to trace many of the effects of different forms of instruction. For example, students in the basal classrooms initially were faster at oral reading. We think this is because they didn't have any time-demanding strategies for sounding out words that they didn't know. However, by third grade, there were no striking differences either in reading speed or in performance on comprehension tests. Further, both groups showed the same general pattern of early word recognition facility's predicting later comprehension skill. We continue to examine the data for evidence that specific children require one kind of training or another. This is where differences in curricular success are likely to appear.

What Resnick and I did not do was was to collect sufficiently rich qualitative data on the development of reading skill; at least we have yet to analyze or report any (we do have some preliminary indications that qualitative modeling of the pause structures of early oral reading may be interesting). What I am suggesting today is that we begin to commit ourselves to longer term, richer studies of the course of learning. I think that we will need broader and more refined knowledge of how flexible, extensible skill is acquired if we are to make the same progress in improving human opportunities for intelligent activity as has occurred in research on machine intelligence. Even though such research requires complex computer tools and teams of researchers, we must attempt it.

NEW STYLES OF RESEARCH

This Society has an important role to play in matching the scale of research to the problems for which experimental psychologists can provide solutions. As computer networks, more powerful machines, and artificial intelligence tools appear, we will be capable of conducting studies whose time scale and complexity match that of human learning. By doing so, we can match the new theoretical power of cognitive psychology with the empirical basis that can preserve it as a science. My experience is that this requires more collaborative efforts, and perhaps more interdisciplinary efforts, than were once the case for our field. Nothing I've reported would have been possible had I tried to do it myself, even in the rich computer environment available to me. Fortunately, I've had colleagues skilled both in their disciplines and in the social skills of collaboration. Someday, perhaps we'll understand how to help people acquire both kinds of skill.

REFERENCES

Lesgold, A. M. (1983a). Acquiring expertise. In J. R. Anderson & S. M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower*. San Francisco: Freeman.

Lesgold, A. (1983b). Intelligence: The ability to learn, or more? A review of *Handbook of Human Intelligence* by Robert S. Sternberg, *Contemporary Education Review*, 2, 111-119.

Lesgold, A. M., Feltovich, P. J., Glaser, R., & Wang, Y. (1981, September). *The acquisition of perceptual diagnostic skill in radiology* (Tech. Rep. No. PDS-1). University of Pittsburgh, Learning Research and Development Center.

Lesgold, A., & Reif, F. (1983). *Computers in education: Realizing the potential. Report of a research conference* (U.S. Department of Education Publication Nos. 1983-381-054-134 and 1983-381-054-135). Washington, DC: U.S. Government Printing Office.

Lesgold, A. M., & Resnick, L. B. (1982). How reading disabilities develop: Perspectives from a longitudinal study. In J. P. Das, R. Mulcahy, & A. E. Wall (Eds.), *Theory and research in learning disabilities*. New York: Plenum.

NOTES

1. See (Lesgold, 1983b) for thoughts of mine on some related issues.
2. Known locally as the Department of Cognitive Surgery.
3. Available from the author (at LRDC, University of Pittsburgh, Pittsburgh, PA 15260) on request.
4. My colleagues Isabel Beck and Steven Roth have recently developed game software to provide such practice.