INTRODUCTION

Personalized System of Instruction (PSI) was developed prior to the widespread availability of computers, and numerous assistants were required to help administer the system. Traditionally, most of these individuals were students in more advanced courses who received some credit in their course for helping with the PSI procedure in a less advanced course. The amount of administrative work required is likely the main reason for the decline of interest in PSI. Another obstacle may have been the difficulty of obtaining a more advanced course from which to draw students. In addition, students in another course are unpaid labor, which violates contracts that some universities and colleges have with campus unions. The enormous processing capacity of modern computers provides technological solutions to these problems, as well as a means to enhance PSI. We believe that all effective courses include methods for: (1) presenting specific study objectives, (2) providing contingencies that result in verbal activity relating to the objectives, and (3) monitoring this activity and providing the students with feedback. PSI is a highly systematic method for doing each of these things and thus is well suited to computer-aided delivery. This chapter describes Computer-Aided Personalized System of Instruction (CAPSI), an approach to higher education that grew out of this line of reasoning.
Many resources have been invested in generating computer programs that develop a variety of the skills needed to interact with complex phenomena, but higher education consists of more than just learning such skills; it is typically seen primarily as developing the complex verbal abilities that we call knowledge, comprehension, critical thinking, and higher-order thinking. This presumably is the purpose of activities such as listening to lectures, participating in discussions, reading texts, and writing papers, tests, and examinations. In the large majority of classes, computer technology has not had a major impact on these activities. The most dramatic change computer technology has brought to college and university teaching has been in the use of e-mail. On the whole, computers have served in this regard simply as a messaging system—an extremely convenient one both for instructors and students—but not in itself a primary teaching tool. Computer technology in higher education also provides access to online catalogues of library material, databases of literature citations, and articles and other information posted on websites. Direct applications of computer technology in higher education have largely been devoted to an attempt to import standard classroom activities into a web environment. These applications, however, do little to address problems that exist in many standard classroom activities.

If recreating standard course procedures on the Internet is not an efficient use of computer technology in higher education, what is required? We believe that contingencies resulting in course-related verbal activity must be combined with methods for presenting specific study objectives, monitoring the activities of students, and providing them with feedback. PSI is a highly systematic method for doing this and thus should be well suited to computer-aided delivery. PSI and computers appear to be a natural match, and it seems to follow that computerizing PSI is an efficient way to use computers in higher education.

A BRIEF HISTORY OF CAPSI

The first version of CAPSI was implemented in the 1983/84 academic year as a computerization of a PSI method that the senior author used to teach his undergraduate psychology courses at the University of Manitoba (Kinsner & Pear, 1988; Pear & Kinsner, 1988). Rather than using student assistants from another course, this method used students from the same class who had passed a given unit as peer reviewers. To pass a given unit, a student was required to demonstrate mastery on each question assigned on a test for that unit. If a student's test did not demonstrate mastery, a "restudy" result was assigned—meaning that after a period of time allotted for restudying the unit, the student could take another test on that unit. The instructor monitored the
peer reviewers' work and they received a small amount of course credit for each unit test they reviewed.

All questions were of the short-answer or short-essay type. When the computer was introduced, its main function was to randomly select question numbers indicating which study questions a student was to answer and to select the instructor or teaching assistant to evaluate the test or two peers to review it. Both peer reviewers had to independently agree that the student had demonstrated mastery in order for the computer to credit the student with a pass. Students worked at a teletypewriter terminal called a DecWriter; this was connected to the university's mainframe computer by telephone. There were often long lines at the terminal, which were alleviated to some extent by increasing the time that the classroom was available. A second terminal was introduced in the 1984/85 academic year. When e-mail came into effect in the 1986/87 academic year, students who had access to terminals and modems were allowed to write tests outside of the classroom and e-mail their answers to the instructor for marking. In the 1987/88 academic year, the program was integrated with e-mail so that the computer automatically mailed tests to students requesting them and mailed the completed tests to two peer reviewers, a teaching assistant, or the instructor.

In the 1994/95 academic year, a DOS version was implemented on a LAN server in the psychology department of the University of Manitoba. With a few minor upgrades, this is the version currently in use (Pear & Crone-Todd, 1999). Students access the program either from campus labs or off campus via a telnet connection. In a course taught using CAPSI, students complete study questions selected by the program from a bank of essay-type questions on the material they have just learned. Students proceed at their own pace through the study units. Students who demonstrate mastery of a unit (defined as correctly answering all questions on the unit assignment or test) may serve as peer reviewers for that unit (Figs. 1 and 2). The program is applicable to any course topic and any set of questions or problems. It is currently being used successfully in a number of undergraduate psychology courses taught at the University of Manitoba, and it has been used to teach distance education courses as well as on-campus courses. In addition, a collaboration with another university merged two courses on the same topic at the two institutions.

**HOW CAPSI UTILIZES COMPUTER CAPABILITIES**

A course taught by CAPSI starts out in design like a traditional PSI course. A text covering the course material is chosen, the course material is broken down into study units, and study questions are written that define what the student is expected to have learned as a result of reading the text. After this, however, the method takes full advantage of the computer's information-processing, data-storage, and communications capabilities. We consider each of these in turn.
(Top) Computer screen as viewed by a student receiving a unit test consisting of three short-essay questions. (Bottom) Upon opening one of the questions, the student sees the entire question field and a field for typing the answer. The student can scroll in the windows if necessary.

**Information-Processing Capabilities**

Despite the immense information-processing capabilities of computers, their linguistic abilities at present fall far short of those of humans. This might seem to limit the use of computers in PSI to restricted-answer questions (e.g., true/false, multiple-choice, fill-in-the-blank) or, at best, questions requiring just a
few words. There are computerized PSI programs that take this approach. We believe, however, that unless restricted-answer questions are simply being used for screening purposes, this is a step in the wrong direction. The power of PSI, in our view, lies in the large amount of verbal behavior and large number of interactions between learners that it has the potential to generate. It is true.
that restricted-answer questions are sometimes used in traditional PSI courses; however, in these cases, the student assistants ideally provide oral feedback so that there is opportunity for discussion of the material. The generation of expository verbal behavior, not the mere choosing of a provided option, is essential to effective learning. Such behavior is not likely to occur when the answer is constrained by a few choices that are generally provided on the test; or, if it does occur, will be limited to the private level. Moreover, we cannot assume that a correct response on a restricted-answer test necessarily implies that the appropriate private behavior (i.e., thinking) has occurred.

When a student produces expository verbal behavior, however, his or her thinking is in plain view to be reinforced with positive comments if it meets the criterion for reinforcement or changed with corrective feedback if it does not. Despite the computational capabilities of modern computers, they are not yet at the level at which they can, in general, respond to expository verbal behavior with sufficient understanding to reinforce it appropriately. With CAPSI, however, the system as a whole—which is the collection of individuals who are interacting through the mediation of the computer—does contain sufficient understanding. In fact, the system generates understanding of the material at an exponential rate. Although the computer itself as one component of the system may not be able to provide adequate feedback, it can locate the requisite knowledge and understanding within the system as a whole. Less abstractly, the computer knows each student's level of mastery in the course at any given instant. Thus, in addition to being able to assign the instructor or teaching assistant to provide feedback on a given answer, the computer can assign this task to a student in the course—a peer reviewer—who has mastered the unit of the test. All that is needed is an algorithm for assigning tests to qualified peer reviewers. In passing, it might be noted that the system as a whole, consisting of complexly interacting parts resulting in states that are not totally predictable from earlier states, is an example of what is termed a dynamical system (Kinsner & Pear, 1990).

Personalized System of Instruction reorganizes the transfer of expertise that occurs in more traditional classrooms. Instead of a single teacher dispensing knowledge to a large number of students, students in PSI respond to the subject matter throughout the course and receive frequent guidance. The computer's selection of peer reviewers from within a CAPSI-taught course further transforms these relationships, in that students learn directly from each other as well as from the instructor and teaching assistants. The system is also highly scalable within limits—as class size increases, the pool of available peer reviewers keeps pace. In large classes, students continue to receive rapid and frequent feedback in quantities well beyond what a single instructor could provide. There is, of course, an upper limit to the number of students that can be handled by a single instructor who must oversee the whole course; however, class size could potentially be quite large if there is a commensurate increase in number of paid teaching assistants.
Data Storage Capabilities

The computer's vast capacity for storing and accessing data means that every interaction in the course is available for scrutiny by the instructor or by a researcher. Throughout a course, students request tests and answer questions, indicate their availability to review the tests of other students, write comments as feedback to answers on other students' tests, make judgments about the overall mastery shown in those tests, and (sometimes) appeal the results of their tests. The program stores the date and time of all these events, enabling the instructor to evaluate student progress in the course at any time; however, CAPSI provides not merely a transaction record, but a true relational database of course activity. That is, an instructor or researcher can make queries that efficiently reorganize the data into nearly any useful form. It is as easy to generate a longitudinal record of test writing by a single student as to examine all the answers given by a class to a particular study unit or even a single study question. This latter function permits rapid identification of course topics and objectives to which students are not responding effectively. In such a case, the objectives can be restructured or supplemented, and the results of these changes can be assessed with equal rapidity. The activity of peer reviewers can also be related to study units and objectives, which can help to identify objectives that are especially difficult to assess. This information can be used to locate and reduce vagueness in the objectives or to identify skill deficiencies in the peer reviewers. Queries can also be made to monitor the work done by individual peer reviewers, for the important purpose of providing feedback to improve their performance. Ultimately, the data storage capabilities of a computer-aided system should help not only to optimize individual courses but also to discover how course variables affect the learning process.

Communications Capabilities

One of the features of PSI is that the method does not require all students to attend class at the same time. There are, for example, PSI courses in which assistants are available at extended periods throughout the day and students can choose to take tests from a wide range of times. The communication capabilities of the computer extend this convenience without limit. Computers allow access 24 hours a day, 7 days a week, permitting asynchronous testing and feedback. Moreover, students do not have to go to a central location to be tested. This means that CAPSI is employable in a virtual environment, both temporally and spatially. We have used it for both on-campus and distance courses. We have also combined on-campus and distance courses. Students of widely varied backgrounds can interact, even without necessarily realizing that they are interacting with students from other locales. Knowing the background, current location, or situation of the person they are interacting with is not relevant to learning the course material. The heterogeneity thereby introduced
may have beneficial learning effects as well as result in more efficient course delivery.

The flexible design of PSI makes it especially suited to accommodating students with disabilities (Brothen, Wambach, Hansen, 2002). These advantages are accentuated by use of the computer. Students with chronic health difficulties do not need to worry about missing test dates in CAPSI-taught courses and do not even need to be physically able to travel to the classroom. Moreover, the instructor can adjust time requirements on an individual basis so that students with learning disabilities can take as much time as needed in reading the text-based materials and composing answers, without feeling pressure to keep up with a lecture or class discussion. Indeed, the computer-mediated nature of CAPSI permits persons with virtually any disability to be readily accommodated with specialized computer programs or equipment.

REFINEMENTS OF CAPSI

The bare outlines of CAPSI have been sketched above; however, further details and additional features that have been or can be incorporated require clarification.

Higher-Order Thinking

Some computer-mediated instruction can be fairly criticized for emphasizing learning by rote or for developing knowledge without providing the kinds of experiences that develop higher-order or critical thinking. PSI, too, has been criticized on these grounds, although the charge has been answered (Reboy & Semb, 1991; Ross & Semb, 1981). CAPSI avoids this pitfall with at least three major features that explicitly require higher-order thinking by students. The first feature is the systematic incorporation of study objectives that require more than simple knowledge. To teach higher-order thinking, it is necessary to define it precisely enough that it can be recognized when displayed by students. In CAPSI, our approach to operationally defining higher-order thinking is based on Bloom’s taxonomy (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956); however, Bloom’s definitions were not precise enough to allow us to obtain good reliability using them, so the definitions were modified (Crone-Todd, Pear, & Read, 2000; Pear Crone-Todd, Wirth, & Simister, 2001). Table 1 summarizes the modified definitions, and Figs. 3 and 4 show flow charts for defining question and answer levels, respectively. The reliability of these definitions usually ranges from moderate to good, although more work is needed to improve their reliability.

The peer review system makes the use of the higher-level objectives feasible. It may be possible to compose restricted-answer questions that test the higher levels of thinking, but it is certainly more difficult than it is to compose essay-type questions that do so. This is particularly true at the synthesis and
TABLE 1
Modified Definitions of the Categories in Bloom's Taxonomy

<table>
<thead>
<tr>
<th>Categories I and II</th>
</tr>
</thead>
<tbody>
<tr>
<td>The answers to these types of questions are in the assigned material (e.g., textbook or lecture) and require no extrapolation.</td>
</tr>
<tr>
<td>I. Knowledge</td>
</tr>
<tr>
<td>II. Comprehension</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Categories III, IV, V, and VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>These questions go beyond the textual material in that they must be inferred or extrapolated from the information in the assigned material.</td>
</tr>
<tr>
<td>III. Application</td>
</tr>
<tr>
<td>IV. Analysis</td>
</tr>
<tr>
<td>V. Synthesis</td>
</tr>
<tr>
<td>VI. Evaluation</td>
</tr>
</tbody>
</table>


evaluation levels (Bloom’s levels V and VI, respectively). Consider, for example, how one might devise a multiple-choice question that would test a student’s ability to design a self-modification program for washing the dishes after each meal or a multiple-choice question that would test a student’s ability to state and defend a position on the justifiability of a particular war. Creativity, which is required at Bloom’s levels V and VI, is particularly difficult (if at all possible) to test with restricted-answer questions, because by definition creative behavior cannot be specified in advance. That is, one can often recognize a creative answer when one sees it but not predict in advance all possible creative solutions to a particular problem; however, individuals who are well
FIGURE 3
Does the answer use terminology appropriate to the course?

Yes

No

Is it factually and completely correct?

Yes

No

Was a position defended?

Yes

No

Were concepts combined in a new way?

Yes

No

Was a principle or definition created?

Yes

No

Was a comparison or contrast made?

Yes

No

Was an example explained?

Yes

No

Was an original example identified or provided?

Yes

No

Knowledge

Comprehension

Evaluation

Synthesis

Analysis

Application

FIGURE 4
versed in a particular subject can make judgments regarding the creativity of particular answers to questions in that subject. Automation of these judgments is not technologically possible at the present time, and it may be quite some time (if ever) before it is.

The second feature conducive to higher-order thinking and creativity is CAPSI’s peer-review process. Students are instructed to explain the result they assign to each unit test they review. These reviews, therefore, represent activity at Bloom’s level VI.

Finally, CAPSI incorporates an appeal process as a check against assignment of “restudy” results when “pass” is warranted. Students are instructed to appeal a restudy result to the instructor if they disagree with it. Making an appeal involves addressing the reviewer’s comments in a clear and convincing manner. If the instructor accepts the student’s argument, the result of the test is changed. In most cases, arguments provided are above the level of that of the question itself; typically, in fact, arguments of a structured level VI responding are required in order for an appeal to be successful. Feedback on appeals of tests is also provided to the peer reviewers of the appealed tests in the expectation that this will have a positive effect on their thinking levels as well.

**Incentives for Student Behavior**

It is essential to provide incentives (or reinforcers) for students to engage in the behaviors required by the system. Specifically, students must prepare answers to study questions, write tests, and review other students’ answers. A simple point system, translatable into grades at the end of the term, is effective at maintaining these behaviors for most students. Curving of marks is possible but not necessary or recommended; however, it is not enough that students perform certain behaviors—a minimum level of quality is also expected. It is important, for example, that peer reviewers provide feedback promptly. Delays in feedback are alleviated by having students indicate their availability for peer reviewing to the computer, which will then only select students who can respond within a specified time period (e.g., 24 hours). A mild point penalty is levied on a student assigned to review a test if he or she has not done so within a certain specified time after the student has submitted it. The test would also be reassigned to another reviewer or to the instructor or teaching assistant.

Feedback given by peer reviewers should also be accurate. Requiring mastery on a given unit before a student may review tests of that unit is part of the solution to this problem; however, there can be slippage in the system because of its hierarchical nature. Students who are last in finishing a unit may have their tests reviewed by students who were next to last, and so on. Careful spot-checking of peer reviews and providing appropriate feedback can alleviate this problem. The data collection capabilities of computers facilitate this monitoring. Research (Martin et al., 2002a,b) has shown that it is possible to reliably measure feedback accuracy and that a quality control measure of having two
Future versions of the system could allow students to rate the peer reviewers who review their tests, in terms of both quality and promptness of feedback. These interlocking social contingencies would mirror those responsible for much of our behavior in everyday life, while maintaining a formal structure amenable to control and record keeping by the computer.

**Plagiarism**

Administering tests asynchronously at distant locations means that tests are not supervised; hence, answers may be copied from the text or other sources. There are various solutions to this problem: (1) programs can scan for copied or closely paraphrased material that does not credit its source; (2) a time limit, managed by the computer, may be imposed on each test measured from the time it is received to the time it is submitted; (3) heavily weighted supervised exams, separate from the PSI portion of the course, may be given at a centralized location; and (4) test questions should be designed at the higher thinking levels so that answers are not readily available to be copied. While the possibility of plagiarism and cheating is always present, the issue is no more serious for CAPSI than it is in any course that assigns work outside of class. Indeed, the problem may be less for a CAPSI-taught course because the answers to study questions (especially the higher-level questions) may be less readily available than papers on specific topics.

**Preventing Mastery-Criterion Circumvention**

Computer-Aided Personalized System of Instruction employs a mastery criterion; however, it is not feasible to ask every question in a unit on each test of that unit. Questions must be randomly sampled. We have found a random sample of three questions to be convenient; however, because CAPSI provides a student the option of canceling a test after receiving it, a student might in theory continue requesting tests until one occurs that contains only questions he or she can answer. Informal observations, as well as sampling theory, show that most students quickly discover that a strategy of learning the answers to only a few questions in each unit is ineffective. First, there is a minimum time requirement between successive attempts on a unit; trying to obtain favorable questions on every unit greatly slows a student's progress through the units. Second, the most important concepts in a course typically recur throughout the units; thus, it would be very difficult for students to progress through all the units without learning those concepts.

A procedure that is not present in the current version of CAPSI, but which might well be incorporated into future versions, is stratified sampling of questions rather than strictly random sampling. That is, questions would be
randomly sampled from set categories (e.g., the categories in Bloom's hierarchy). Midterm and final examinations integrated with the CAPSI component of a course constitute another check that is used to help ensure that a high grade in the course signifies a high level of mastery of the course content. While we feel that these steps go a long way toward preventing circumvention of the mastery criterion, we acknowledge a need for studies to determine how best to ensure mastery in a CAPSI-taught course.

**Training**

A variety of skills are required to take a CAPSI-taught course. One needs to know how to study for tests, how to write acceptable answers, and how to evaluate one's own work and review that of others. A variety of skills are also required to be an instructor of a CAPSI taught course. The question, therefore, arises as to how individuals are trained to function in CAPSI-taught courses.

**Students**

We find that the system works well with no explicit training outside of the system. This is because CAPSI contains built-in training features. First, the program can be used to teach itself. In our CAPSI-taught courses, the first study unit is about the system itself. In that unit, students learn from a manual containing study questions about how to use the technology, how to write unit tests, and how to review their own work and that of others. Second, students receive feedback through the course on how they are performing as test writers and as peer reviewers. Third, the program is highly menu based, and students have a great deal of experience with menu-based programs, as well as experience answering questions in courses. There is, of course, always room to do more with regard to training of peer reviewers. For example, skill deficiencies could be assessed and special tutorials devised. Students could be given practice taking tests and peer-reviewing them prior to beginning to write tests on actual units.

**Instructor**

Most instructors who use CAPSI have previously been students in CAPSI-taught courses; therefore, it is difficult to say how much training is necessary for instructors who have no familiarity with CAPSI. However, our experience in teaching a new CAPSI instructor suggests that an instructor can be shown how to independently create and administer courses in an hour or two. As with the student portion of the program, the instructor portion is heavily menu based, which is extremely helpful in learning and remembering the functions of the program. Also, like most individuals in our culture, instructors are very familiar with the process of asking and answering questions. Moreover, we hope that
our work on thinking levels of questions and answers will be helpful to instructors in their design of questions that help to establish behavior at the desired levels.

Programming CAPSI

A final problem to be mentioned is that there currently is no commercially available CAPSI computer program, and the CAPSI functions (in particular, the peer-review procedure) cannot be incorporated readily into any contemporary commercial online or web-based instructional program. The only existing CAPSI program is the current DOS program, which continues to function well after 9 years of continuous use at the University of Manitoba but which is not readily transferable to other institutions. However, a web-based CAPSI is currently being developed at the University of Manitoba.

THE PEER REVIEW SYSTEM AT WORK

Figure 5 shows workload dynamics averaged from 3 successive years of a CAPSI-taught class in behavior modification. It can be seen from the figure that the instructor and teaching assistant marked the majority of unit tests.
during the first few weeks of a course. As the course progressed, the instructor and teaching assistant continued to mark tests for the first few students completing each unit, but by the end of the course, 90% on average of all unit tests were being peer reviewed.

Table 2 shows data from a single class on the amount of substantive or rich feedback (defined as a comment that is clearly and explicitly related to the question or the student’s answer) each student in the class received and provided. Note that students are listed in descending order according to the

<table>
<thead>
<tr>
<th>Student</th>
<th>Total</th>
<th>From instructor/teaching assistant</th>
<th>From other students</th>
<th>To other students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>36</td>
<td>29</td>
<td>7</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>38</td>
<td>27</td>
<td>11</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>20</td>
<td>25</td>
<td>48</td>
</tr>
<tr>
<td>4</td>
<td>46</td>
<td>18</td>
<td>28</td>
<td>68</td>
</tr>
<tr>
<td>5</td>
<td>23</td>
<td>15</td>
<td>8</td>
<td>75</td>
</tr>
<tr>
<td>6</td>
<td>69</td>
<td>12</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>9</td>
<td>31</td>
<td>40</td>
</tr>
<tr>
<td>8</td>
<td>29</td>
<td>7</td>
<td>22</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>42</td>
<td>7</td>
<td>35</td>
<td>42</td>
</tr>
<tr>
<td>10</td>
<td>28</td>
<td>4</td>
<td>24</td>
<td>28</td>
</tr>
<tr>
<td>11</td>
<td>33</td>
<td>3</td>
<td>30</td>
<td>47</td>
</tr>
<tr>
<td>12</td>
<td>57</td>
<td>3</td>
<td>54</td>
<td>25</td>
</tr>
<tr>
<td>13</td>
<td>56</td>
<td>2</td>
<td>54</td>
<td>23</td>
</tr>
<tr>
<td>14</td>
<td>66</td>
<td>2</td>
<td>64</td>
<td>20</td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>1</td>
<td>99</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>19</td>
<td>0</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>32</td>
</tr>
<tr>
<td>18</td>
<td>28</td>
<td>0</td>
<td>28</td>
<td>77</td>
</tr>
<tr>
<td>19</td>
<td>30</td>
<td>0</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td>20</td>
<td>34</td>
<td>0</td>
<td>34</td>
<td>73</td>
</tr>
<tr>
<td>21</td>
<td>35</td>
<td>0</td>
<td>35</td>
<td>45</td>
</tr>
<tr>
<td>22</td>
<td>35</td>
<td>0</td>
<td>35</td>
<td>4</td>
</tr>
<tr>
<td>23</td>
<td>39</td>
<td>0</td>
<td>39</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>953</td>
<td>159</td>
<td>794</td>
<td>794</td>
</tr>
</tbody>
</table>

amount of feedback received from the instructor or teaching assistant. It is apparent that all students in the course received the majority of their substantive feedback from peers. Figure 6, which is based on Table 1, contrasts the amount of substantive feedback received from the instructor or teaching assistant with the amount received from peer reviewers. It is clear that the students provided far more substantive feedback than the instructor. We believe that this amount of feedback cannot be equaled in a traditionally taught course.

Figure 7 shows data from the same 23 students from the class used in Table 2. Each student is plotted according to both the amount of substantive feedback received on unit tests and the amount of substantive feedback given while serving as a peer reviewer. Although all students received considerable feedback, a few students gave very little. Nevertheless, a majority of students were near or above a ratio of 1:1 (shown by the diagonal line) for these measures. In other words, most students received as much or more substantive feedback than they gave. Table 3 shows typical examples of the substantive feedback that the students gave. It seems clear that the feedback students gave is of high quality.

Because students provide so much of the feedback in CAPSI-taught courses, it is important that their comments are accurate. A study of feedback accuracy by Martin, Pear, & Martin, (2002b) provided data on two types of errors in feedback given by peer reviewers. Some changes suggested by reviewers were unnecessary because the student's answer in fact contained the required information. Other suggested changes were unnecessary because,
You are on the right track; however, it is not quite clear if you fully understand the concepts. The examples are not complete; check out page 5. The first example does not show circular reasoning and in the second example, pursue what? Using examples is a great way to convey ideas. Just make sure they are complete and get your message across.

Correct. It may be helpful to follow up your answer with examples.

Check out page 44 in the text again. Hint: Reread item 2. I think you misunderstood the identifying clue. Also, explain your examples in more detail; they do not completely illustrate the concepts.

The explanation for FI is incorrect. Refer to pages 76 and 77 in the text book. Reading the example might help clear up things. Also, the question asked for spaced-responding DRL, not DRI.

Some key info in your explanation of what is specifically wrong is missing. Check out pages 109 and 110. Hint: Distinguish behavioral and cognitive psychology. Michael makes some interesting points. Be very careful on your spelling; “cure” should be “cue”—two very different meanings; also, “regence” should be “reference.”

Good. Not that it is asked for, but you may want to explain what an “unprogrammed reinforcer” is as it is a new term. Just a thought, but no biggie. Your answer is a perfect answer on its own anyway. Excellent!

Good answer, and good suggestion on a treatment.

Good job. Great example. Could have mentioned a few other situations. See pages 51 under #3 setting.

although the student’s answer did not contain the information specified in the feedback, that information was not required for mastery. Yet, in the course assessed, 87% of all instances of substantive feedback were free from either kind of error. These results indicate that in this teaching method students receive and give large quantities of high-quality substantive feedback.

Data also suggest that this feedback leads to student learning. Martin et al. (2002b) examined occasions on which students answered a particular study question, received substantive feedback on the answer, and subsequently had reason to answer that question again (either on a repeated unit test or on an exam). The subsequent answers improved in ways consistent with the specific content of the feedback in 61% of these instances.

COMPARISON WITH TRADITIONAL COURSES

Thus far, we have not compared CAPSI empirically with traditional courses. Such a comparison is important. Because it is based on PSI to a very large degree, we expect that CAPSI will prove superior to the traditional lecture method (see previous chapter). In addition, there is considerable evidence demonstrating the effectiveness of both computer-based instruction (Kulik & Kulik, 1991) and mastery learning programs (Kulik, Kulik, & Bangert-Drowns, 1990). This should not be surprising if we think of learning an academic subject matter as the learning of a set of largely verbal skills. As with any set of skills, practice of those skills is important in the development of mastery. We would not expect someone to become a competent driver, for example, merely by listening to lectures on driving or watching demonstrations of it. Thus, systems that require active learning and mastery typically produce better learning of academic subjects just as they do of other skills. As indicated previously, students in a CAPSI-taught course write much more and receive far more feedback than would typically be possible in a traditional course.

COMPARISON WITH OTHER COMPUTER-MEDIATED COURSES

Most computer-mediated courses attempt to adapt the computer to traditional methods of teaching. Courses taught in this fashion may be highly unsystematic with little assurance—or the means to ensure—that students are learning much of anything relevant to the subject matter. The problem lies with trying to adapt traditional educational methods—which already lack effectiveness—to an environment that is unsuited to them. Because of its highly systematic procedure (almost by definition; note that the middle initial of PSI stands for "system"), PSI adapts naturally to the computer environment.
Many computer-mediated teaching methods fail to take advantage of the
tremendous capacity of the computer to provide structure to student–student
and student–instructor interactions. The result frequently is a lack of focus on
the course topic in discussions, lack of involvement by many students in the
process, and a large quantity of verbal material to be evaluated that represents
little actual work by many students. Many courses involving computer-
mediated communications use the computer primarily for extended comment-
tary by the instructor and students on given course topics (see Pear & Novak,
1996). CAPSI, in contrast, takes full advantage of the ability of the computer to
provide structure to the critical interactions that occur in an educational
environment (see Kinsner & Pear, 1990).

EXPANDING ON TECHNOLOGY

We have not addressed the more impressive multimedia capabilities of com-
puters, such as presenting realistic images and simulations. There is no reason
why these features should not be exploited to the fullest, once the basic
aspects of CAPSI are in place. For example, the computer could provide online
teaching sequences prior to a given unit and could test students to determine
whether they are ready to be tested by the human components of the system.
There is no escaping the fact that, for the foreseeable future, judgments
regarding the higher thinking levels will have to be done by humans. Indeed,
even if computers were able to make these types of judgments, CAPSI would
probably still use peer review if for no other reason than students learn to
evaluate their own work by reviewing other’s work. Nevertheless, CAPSI can
only stand to benefit from any advances that may occur in the field of artificial
intelligence. These advances could lead to improved methods for developing
CAPSI-taught courses, evaluating and providing feedback to learners in these
courses, providing tutorials for them, and analyzing the data resulting from the
courses (see Kinsner & Pear, 1988, 1990).

RESEARCH STUDIES ON CAPSI

The system we have described provides many areas for scientific study, includ-
ing the development of higher-order thinking, the optimal operational levels of
the parameters of the peer review system, and the provision of effective
feedback to students. Although CAPSI differs significantly from traditional
courses, the variables in CAPSI are traditional educational variables. Traditional
courses contain textual material on which students are tested, and through
discussions and other activities students, as well as instructors, provide infor-
mation and feedback to other students. CAPSI therefore provides a laboratory
for the systematic manipulation and study of traditional educational variables.
References


The Morningside Model of Generative Instruction: An Integration of Research-Based Practices

KENT JOHNSON  ELIZABETH M. STREET
Morningside Academy  Central Washington University

ABOUT MORNINGSIDE ACADEMY

In 1980, Kent Johnson founded Morningside Academy in Seattle, Washington, to provide behaviorally designed academic and social programs for children and youth and to prepare teachers and other school personnel. Morningside has grown to be a corporation with distinct programs serving each of the original goals. First, Morningside Academy is a school, operating during the school year and in summer. Second, Morningside Teachers’ Academy participates in formal external partnerships with schools and agencies throughout the United States and Canada. It also offers a summer institute for teachers, graduate students, and other professionals. Morningside’s programs are continually evolving to better prepare students for successful schooling and citizenry. This chapter provides an update of a previous article and chapter (Johnson & Layng, 1992, 1994) and describes recent program developments in the Morningside Model of Generative Instruction.
CURRENT WORK

Morningside Academy is a school for elementary and middle school students, most of whom have performed poorly in their previous schools, to catch up and get ahead. Entering students typically score in the first and second quartiles on standardized achievement tests in reading, language, and mathematics. Some have diagnosed learning disabilities (LD); others are labeled as having attention deficit disorder (ADD) or attention deficit/hyperactivity disorder (ADHD). Some lag behind their peer group for no diagnosed reason. Students' IQs range from low average to well above average. A small percentage of students have poor relations with family members and friends, but most do not.

Morningside Academy's elementary school students typically enroll for 1 to 3 years to catch up to their grade level. Many middle-school students enroll for all of middle school. Morningside Academy offers a money-back guarantee for making 2 years of progress in 1 school year in the skill of greatest deficit. In 23 years, Morningside Academy has returned less than 1% of school-year tuition.

The academic program focuses upon the three main foundation skills—reading, writing, and mathematics—including language, facts, skills, concepts, principles, problem solving, and organizational aspects. Literature, social studies, and science are the grist for teaching these foundations. Each student participates in extensive entry assessments of academic, learning, and performance skills. Students with similar needs and goals are grouped together for instruction. Groupings change throughout the day and year, depending on subject matter and student progress.

Morningside Academy's teachers coach students to perform their best using clearly defined rules and expectations for performance and productivity, explicit modeling of high performance, and moment-to-moment monitoring and feedback. Students carry a daily report card throughout the day. Points are earned and recorded for meeting specific academic, learning skills, and citizenship aims that the teacher specifies before each class period. Students share their report cards with their families each day. Many students earn home-based rewards such as extra television, computer access, or telephone time for meeting their aims. In addition, classroom wall charts display the points that each student earns.

In the middle school, in addition to the aforementioned Foundation skills, students learn how to study and perform successfully in content classes in the social and natural sciences and the humanities. Each program explicitly teaches textbook reading, note taking, studying, participating in class discussions, test taking, and essay and report writing.

Morningside Academy offers a 5-week summer school program that focuses on reading, language, writing, and mathematics. The summer school program offers a money-back guarantee for progressing 1 year in the skill of greatest deficit. Morningside Academy has returned less than 2% of summer school tuition.
MORNINGSIDE TEACHERS' ACADEMY

External Partnerships

Morningside Teachers' Academy helps public and private schools and agencies implement its programs through formal external partnerships. Our collaboration with each partner is extensive, usually lasting from 3 to 5 years. Four main goals of external partnerships are (1) to help their students achieve grade level performance; (2) to teach their faculty the assessment, teaching, and learning strategies that will serve them throughout their teaching career; (3) to teach their principals how to become instructional leaders and support their faculty's teaching efforts; and (4) to develop implementation, teacher education, and internal coaching systems to maintain our efforts after we leave. To achieve these goals, we provide 40 to 60 hours of workshops and 30 to 50 days of individualized, in-classroom coaching per school year. Morningside also maintains frequent contact with school district personnel, focusing upon problem solving, program expansion, and self-maintenance, including train-the-trainer models that are suited to a school district's ongoing relationship with its schools. As of this writing, 86 schools and agencies throughout the United States and Canada have contracted for partnerships since 1991.

Summer School Institute (SSI)

Morningside Teachers' Academy offers a 4-week intensive summer institute for individual teachers, graduate students, parents, and other professionals who want to learn our programs. Participants can earn graduate credits. Over 300 professionals have attended SSI since its inception in 1991. Individuals may also contract for school-year internships and sabbaticals and experience a customized institute.

MORNINGSIDE TECHNOLOGY TRANSFER

Each Morningside program is competency and mastery based, with a set of procedures and expected outcomes for both learners and teachers. Morningside takes an applied science approach to teaching the competencies to mastery. In the Morningside system, research-based components of curriculum and instruction are combined into a generic model of teaching and learning. The science of human learning informs the generic model, just as engineering is informed by its parent science, physics. We describe the generic model, the Morningside Model of Generative Instruction, later in the chapter. In this section, we describe the origins of the curriculum and instruction components that we draw upon. In a continuing expansion of best practices, Morningside's leadership (1) seeks out research-based, learner-verified
materials, methods, and tools to use during instruction, practice, assessment, and measurement of performance; (2) selects certain research-based curricula and instructional methods to user-test at the Academy; (3) adapts materials, methods, and tools to Morningside’s behavioral framework for teaching; (4) user-tests the curricula, methods, and tools at the Academy and collects data on student and teacher performance; (5) develops workshops to teach others how to implement programs that offer improvements over current practice; (6) user-tests the workshops and programs with veteran external partner schools; (7) revises the workshops and makes further adaptations to the programs; (8) designates the program a technology, a practice that is replicable and that can be taught to others; and (9) transfers the technologies to others, as part of the Morningside Model of Generative Instruction.

PHILOSOPHICAL AND EMPIRICAL UNDERPINNINGS

In their 1992 article in the American Psychologist, Johnson and Layng described the Morningside Model of Generative Instruction. The Morningside model prescribes a stepwise progression through an instructional sequence from entry to true mastery of an objective and aligns classroom practices with each step in the progression. Instruction has three phases: establishing, practicing, and applying. True mastery is defined as performance that is accurate, speedy, durable, smooth, and useful. Underpinning the model is the selectionist approach to understanding human behavior advocated by psychologist B.F. Skinner and the progressive philosophy of John Dewey.

Skinner first advocated a selectionist approach to understanding human behavior in 1969 in Contingencies of Reinforcement: A Theoretical Analysis. Compared to a structuralist approach, which emphasizes form and process, the selectionist approach emphasizes the function of particular behaviors in meeting environmental contingencies. Skinner draws a parallel between the emergence of complex behavioral repertoires and the emergence of complex forms in evolutionary biology. The environment selects simple forms, and a more complex entity gradually emerges. In the case of human behavior, reinforcement or success selects the element. In evolutionary biology, reproductive advantage is responsible. The Morningside Model represents a selectionist approach to understanding complex human behavior and designing instructional protocols. The program builds complex intellectual skills from combinations of successful elements. In addition, the model is, itself, evolutionary. It responds to emerging evidence, refining some practices and discarding others that are ineffective in meeting important educational contingencies.

Dewey’s American pragmatism and philosophy of education (1900, 1902, 1916, 1938) describes the kind of selectionist process that occurs during the advanced segments of the Morningside continuum of curriculum and instruction. Dewey emphasizes natural influences over learning, taken from the
student's current activity, goals, and values systems, rather than arbitrary parcels of subject matter teaching, teacher-initiated research, and teacher-initiated project assignments. For Dewey, this selection process is automatic and evolving, leading different learners down different functional paths in the real world. Morningside's program begins with basic elements and tools and builds repertoires that make possible the kind of naturally reinforced learning that is characteristic of project-based learning and Dewey's progressive education.

The Morningside model builds on five separate but overlapping streams of research: generativity and contingency adduction; instructional design and implementation; critical thinking, reasoning, and self-regulated decision making; program placement and modification based on continuous measurement; and classroom organization and management. Johnson and Layng (1992, 1994) review the work of numerous behavior analysts and educators whose contributions are reflected in the model, only a few of which will be repeated here.

**GENERATIVITY AND CONTINGENCY ADDUCTION**

The Morningside model is a model of generative instruction. It hinges on evidence that complex behavioral repertoires emerge without explicit instruction when well-selected component repertoires are appropriately sequenced, carefully instructed, and well rehearsed. Ferster (1965) was among the first behavioral psychologists to observe that new learning and novel behavior is a result of a rearrangement of existing repertoires. Epstein, a student of Skinner, used the term *generativity* to describe the unprompted interconnection of existing repertoires to solve a problem (Epstein, 1991; Epstein, Krishnit, Lanza, & Rubin, 1984). Andronis, Layng, & Goldiamond, (1997) applied the term *contingency adduction* to a related phenomenon. In their account, new contingencies or performance requirements may recruit performances learned under other contingencies. The new contingency shares common features with the original contingencies that produced the performance.

Alessi's pivotal paper in 1987 discussed the implications of generativity for the design and power of instruction. He reasoned that most curricular strands have an infinite set of relationships, all of which cannot be taught directly. Instead, children learn to respond to a general case; they learn a pattern of responding that produces effective responding to many untrained relations. The elegance of an instructional program depends on the programmer's ability to detect and teach some minimal response or generative set which can combine and recombine into the universal set of possible relationships.

Although the phenomenon of contingency adduction or generativity occurs naturally, the arrangement of events to maximize its occurrence does not. Morningside programs are generative because they focus on specific sequences of skills, teach minimum response repertoires, and establish the general case. Instructional programs are built on a logical and empirical
analysis of the knowledge, skills, and relationships in a field of study. Skills are introduced in an order that makes it most likely that previously learned skills will be recruited to meet new performance requirements.

A SYSTEM OF INSTRUCTION

Morningside adopts a scientific approach that builds on a system of instruction first outlined by Markle and Tiemann (1967) at the University of Illinois at Chicago. Instructional protocols are developed according to a set of principles, and programs are tested on naive learners to ensure they produce intended results. Designers begin with clearly stated goals and objectives, conduct content and task analyses to identify curricular tasks on which mastery of terminal objectives hinges, construct criterion tests that fairly represent the stated outcomes, specify the entry repertoire the learner must demonstrate in order to be successful in the program, design an instructional sequence using the minimum set of instructional tasks required to achieve the outcome, and adjust the program based on performance data from naive learners.

Central to the system are (1) a thorough analysis of the content area that is the subject for instruction; (2) instructional protocols that match what Robert Gagne (1965) and others have called types of learning, learning typologies, or learning outcomes; (3) ordering of elements in the curriculum to ensure that learners progress through it seamlessly; (4) elegant instructional protocols that achieve outcomes with the minimum amount of instructional intervention; and (5) field-testing and fine-tuning curricular assumptions and instructional programs at every turn. Morningside likes this system because Morningside, too, is an evolutionary system. Programs are evaluated on the basis of their ability to produce happy, competent learners, naturally reinforced by progress. When programs don’t, they are changed.

ESTABLISHING OBJECTIVES AND ANALYZING CONTENT

When Morningside designers develop an interest in a content area, they read research in the field, examine scope and sequence charts from a variety of curricular materials, and review existing instructional protocols. The work of establishing instructional objectives and analyzing content is interdependent rather than linear. Typically, a general goal is established, and the instructional design team begins a dance between understanding the topic of study and setting explicit objectives that derive from it. Content analysis uncovers critical foundational skills that learners must master to meet terminal goals. Content analysis at Morningside is of two primary types: content-dependent and content-independent.
Content-Dependent Analysis

As Morningside’s content analysts become familiar with the overall contour of the content area, they identify major skill sets within it. Major skill sets may be thought of as the socially validated goals of instruction within a content area. Once major skill sets are identified, analysts attempt to discover the broad range of sub-skills and knowledge that make up these socially validated outcomes. Fine-grained task analyses reveal tool skills and component skills that constitute the authentic outcomes. Underlying knowledge and skills are organized to discover common foundational skills (tool skills) and common second-level skills (component skills) that constitute more complex composite or compound skills. In essence, the analyst builds, from the various major skill sets, a series of overlapping pyramids that have common foundational building blocks and component skills. This work serves three purposes. It reveals a hierarchy of foundational skills which, when mastered, will aid in acquiring higher-level skills. It also reveals where particular skills should be inserted into the scope and sequence. Last, it clarifies when order of presentation of a skill in the instructional sequence is critical and when it is not.

Content-Independent Analysis

In addition to content-dependent analyses, Morningside designers analyze content according to content-independent typologies. Two primary typologies form the basis of this work: learning channel analysis and learning outcomes analysis.

Learning Channels

Haughton (1980) applied the term learning channel to a method of describing objectives on the basis of their stimulus and response characteristics. Stimulus characteristics, in this model, are defined in terms of the sense through which the stimulus is experienced. Using everyday language, Haughton referred to them as see, hear, touch, smell, and taste. Later, he added “think” (which now is referred to as “free”) to describe stimuli that are not present in the external environment but rather reside in the history of the learner. Response characteristics are described on the basis of common movements (for example, say, write, point, and do). A stimulus–response pair (for example, see–write, hear–point, free–say) is called a learning channel. Designation of learning channels allows the instructional programmer to ensure that all combinations required to operate effectively in authentic tasks are included in the instructional regimen.

Learning Outcomes

Morningside analysts rely heavily on Tiemann and Markle’s analysis of learning outcomes which first appeared in their text, Analyzing Instructional Content (1983,
Tiemann and Markle's typology is reminiscent of Gagne's types of learning first described in his seminal work, *The Conditions of Learning* (1965). Tiemann and Markle's account extends Gagne's work, provides a different classification scheme that is somewhat more teacher friendly, and, most important, includes extensive training in how to develop efficient instructional protocols. Nine learning outcomes are proposed: three psychomotor skills, three simple cognitive skills, and three complex cognitive skills. Morningside's programs focus primarily on simple and complex cognitive skills. An instructional protocol is matched to each type and channel of learning and becomes a kind of authoring system into which new content can be entered, producing a more or less explicit script for teachers to follow in facilitating the learner's progress.

**Instructional Program Development**

We either develop our own instructional materials or select, enhance, and incorporate promising materials into the existing content analysis. When otherwise promising materials fail to use empirically supported protocols for skill acquisition or provide insufficient opportunities to practice, Morningside designers overlay instructional protocols or design additional practice to ensure that the material meets Morningside's standards.

The Morningside design team develops or selects protocols and programs for both well-defined and hard-to-define objectives. Following Markle's (1990) and Engelmann and Carnine's (1982) recommendations, instruction for well-defined objectives minimizes teacher talk in favor of active responding by students. Designers strive for faultless communications in which only one conclusion is possible. Morningside designers also favor lean programs that move students out of instructional routines into practice routines as quickly as possible. Current instructional design efforts attempt to teach the more difficult-to-define strategic objectives as principle applications, a complex cognitive skill in the Tiemann and Markle model, and then support students to add the novel strategies necessary to solve a series of problems without assistance.

Instructional programs are designed to fit the learner. Desired performances are broken down until they link with an individual learner's entering repertoire. There are no assumptions based on age or grade of the learner. Instead, learners are placed into instructional programs that coincide with their individual entering repertoires. Often corrective or remedial learners can skip segments of instruction because their foundation repertoires are spotty. Performance data collected during instruction ensures that students skip segments at appropriate times. Sometimes performance data reveal that additional lessons or program pieces are needed to form a bridge between the learners' current repertoires and the terminal goal.
The instructional block at Morningside is made up of three primary activities: instruction, practice, and application. Each instructional objective or set of objectives is taught in these three phases.

**Instruction**

Instruction refers to establishing a new repertoire; learners acquire a skill that they could not perform previously. Instructional protocols teach basic academic skills, including associations, sequences, concepts, and principle applications. The specific format of the lesson is a function of the learning channel and learning outcome of the task it is designed to teach.

During instruction to establish acquisition, students engage with a teacher in highly interactive lessons that introduce one skill at a time and combine skills as accuracy emerges. Gilbert’s mathetics model (1962a,b) is the primary method for establishing new well-defined skills. Lessons are fairly tightly structured, and their design is intended to produce the speediest possible acquisition of new skills, judged by the degree to which they match an expert’s performance—in this case, the teacher’s performance or the “answer key.” Each instructional lesson is an interchange between the teacher and a single learner, a small group of learners, or an entire class.

When component skills require brief, uniform responses, we employ Engelmann’s Direct Instruction (DI) variant of mathetics (Englemann & Carnine, 1982). In DI, teachers present scripted lessons to learners, who answer teacher-initiated questions in unison. Teachers and students move through a series of questions in a predictable way, providing attending cues to ensure that all students are focused on the question and signals to ensure that students respond together. Teacher and students volley many times a minute with their questions and answers. Teachers praise and correct student responses until all students are accurate. Logically and empirically validated sequences lead students through nearly errorless learning in which stimulus discrimination is explicitly taught and response forms are shaped. The explicitness and careful progression of DI lessons ensures that students develop flawless skills very quickly.

The challenge for teachers is to faithfully present the lesson as designed, achieve choral responding among learners, listen carefully to the quality of their responses, provide encouraging feedback following correct responses, and apply error correction procedures that effectively reduce errors. Even though teachers may consult the script during the lesson, they must be conversant enough with its pattern to maintain eye and ear contact with learners. The best results occur when teachers are able to implement programs with a great deal of procedural reliability, although the method is robust enough to accommodate small lapses in procedure.

Throughout instruction, teachers address errors as they arise. Errors provide an opportunity to determine imperfections in the instructional lesson or provide for immediate and error-specific error correction. If it becomes
apparent that several elements necessary for correct responding are missing for some or all students, the teacher may conduct a brief review to firm up a tool or component skill or revert to a point in the curriculum where performance stabilizes.

Although they are best known for teaching tool and component skills, Direct Instruction lessons have been designed to teach many complex skills in word problem solving, pre-algebra and algebra, chemistry, earth science, economics, history, logic and argumentation rules, and study skills, to name a few. Once basic foundations skills, including learning skills, are established, teachers can shift instruction from DI to a looser mathetics approach. Thus, we have a continuum of Direct Instruction that starts with Engelmann and Carnine’s very formal DI and graduates learners to the more general mathetics case.

Establishing and acquiring are the gateway to mastery and fluency, but they are not synonymous with it. We have found that students who move from skill to skill with halting, if accurate, performance encounter greater and greater difficulty with each subsequent skill. Thus, when prescribed levels of accuracy emerge, students enter the second stage: the practice stage.

Practice

Following successfully completed instructional lessons, students practice their freshly learned skills until they achieve levels of performance characterized by durability and applicability. Practice is timed, highly structured, goal-oriented, and continuously monitored.

Practice sessions apply Precision Teaching (PT) technology, an approach that defines mastery in terms of rate of response. Morningside adopted PT early on and has been a partner in its continued development. The approach was conceived by Lindsley (1972, 1990) at the University of Kansas in his quest for a mechanism that brought continuous measurement and rate data into educational practice. A student and colleague, Haughton (1972) developed the now-standard 1-minute timing to track performance frequencies during practice. Lindsley and Haughton promoted practice regimens that quickly produced high-frequency accuracy rates and low-frequency error rates on well-calibrated curriculum slices.

Morningside recommends optimal performance frequencies that may be characterized as fluent. As a metaphor, fluent performance is flowing, flexible, effortless, errorless, automatic, confident, second nature, and masterful. Fluent performance is fun, energetic, naturally reinforced behavior. Binder (1993, 1996) coined the term fluency building to refer to practice activities that are designed to achieve these goals. Currently at Morningside, five characteristics of performance determine fluent performance frequencies: retention, endurance, stability, application, and adduction (RESAA). Until a frequency aim has been empirically demonstrated to produce these five characteristics of fluent performance, we use the term frequency building instead of fluency building.
Morningside students spend as much as 40% of their school day practicing skills, much of it in the context of our peer coaching technology (Johnson & Layng, 1992, 1994). During peer coaching, pairs of students work together to build skills to achieve frequency aims, although sometimes they practice alone or in threes. During practice, students time themselves on specially designed frequency-building materials until they can perform a certain amount—accurately, smoothly, and without hesitation—in a certain amount of time. Students record their timed performance on specially designed charts, called Standard Celeration Charts (Lindsley, 1972, 1990; Graf & Lindsley, 2002; Pennypacker, Gutierrez, Lindsley, 2003).

Teachers set both performance and celeration aims. Performance aims tell the student how many of a skill they should be able to do in the timing period. Celeration is a measure of the change in rate over time and indicates whether the student is reaching a performance aim in a timely manner. A celeration aim is a line of progress drawn on a chart at a certain angle from the learner's first performance frequency to the frequency aim. The celeration line tells how many days it should take for the student to reach the performance aim, thus providing an empirical definition of progress. As students practice, they plot their own frequencies and compare their progress to the celeration aim lines. Their comparisons tell them whether they are making sufficient progress or whether they need to ask the teacher or another student for help.

Timings typically are of 1-minute duration, although students will do several timings in a single practice session. For selected tasks, students who meet their performance aim during a 1-minute timing move into an endurance phase. Endurance training ensures that students maintain speed and accuracy for appropriately long periods of time.

**Application**

Application is the third activity that occurs in a daily class session at Morningside. After instruction and practice, students apply the skills they have learned in the context of compound-composite tasks such as games, simulations, and real-world applications. Effective application activities adduce a combination of key component elements already in the learner's repertoire to achieve a certain outcome or effect. Application activities may involve explicit, direct instruction in the necessary recombination of elements or they may function as probes to assess automatic combinations of elements.

Most classroom schedules today are driven by activities, not instruction. The activities are made up of challenging compounds to stimulate creative principle application and problem solving. Project-based learning is currently in vogue from late elementary school through college. Most project-based learning arrangements are an "upside-down" approach to curriculum planning: They assume that students can perform all the component elements and that the composite, compound tasks will produce the appropriate contingency
adduction. Proponents contend that projects are inherently interesting and stimulating and believe these anticipated motivational features outweigh component skill weaknesses. The assumption is that, if the task is sufficiently interesting, learners will employ a battery of skills to figure it out. In the end, some learners do and some learners don't. While we agree that meaningful projects are important educational endeavors, we design Dewey’s progressive, real-world applications by introducing compounds later in a “right-side-up” sequence of instruction that teaches from elements to compounds.

We design at least two kinds of application activities. The first kind requires the student to engage in a previously learned performance in a new context. For example, after reading essays in their controlled reading program and engaging in teacher-directed discussions, students may read a newspaper and discuss the articles with their peers or write a letter to the editor of the newspaper about a particular article after learning and practicing the basic rubrics of writing a persuasive essay.

An important reading application activity in our curriculum involves strategically applying comprehension skills during reading. In this application, students, working in groups, take turns reading a selection aloud. At certain points, a teacher stops the reading and engages in “think aloud” monologues that model applications of comprehension skills that students have previously acquired. The teacher may pause the group reading at various points to make a prediction about what will happen next or what a character will do or may make a connection between the plot or a character and the teacher’s own life experience. After two or three think-alouds, the teacher uses a delayed prompting method to assess and prompt students to apply previously acquired skills. First, the teacher calls on a student at certain points during the group reading to make a prediction or connection that will help to make sense of the reading or help the student relate more closely to it. If the student does not respond competently, the teacher provides very general prompts to adduce the application. If the student’s application still does not meet criterion, prompts become increasingly specific and intensive, eventually moving to a full model of the application as needed for the student to achieve a successful response. Thus, the student stays engaged with the teacher until becoming successful, no matter how many volleys are required between them. The teacher provides increasing support until the student is successful. The relevant data to collect are the number and kind of teacher prompts that are provided, not the accuracy of the student’s response, as all students stay engaged with the teacher until they are successful. This delayed prompting method is a reverse of mathetics, in which the teacher demonstrates, then prompts, then tests. Here the teacher tests, then prompts, then finally models until student performance meets criterion. Delayed prompting could be used to promote any application.

The second kind of application activity we design requires new combinations of previously learned elements. More advanced operations in
arithmetic, such as long multiplication or division of numbers, are combinations of previously taught addition, subtraction, and multiplication elements. More advanced forms of sentences and compositions are combinations of elements learned separately during previous writing instruction. More advanced field sports are combinations of many previously learned motor activities and chains. The compound called debating combines elements such as argumentation rules, oratory style, and quick refutation. The elements in all of these activities can be separately taught; the compound can be taught as an application activity that can recruit the necessary elements.

Both kinds of application activities promote generativity and contingency adduction, helping to evolve creative thinkers and problem solvers over time. Applying generic reasoning and self-regulation skills also greatly improves application performance. Students can learn to monitor their own performance and apply reasoning skills to recruit appropriate component elements they have already mastered.

**Instructional Blocks**

Morningside devotes 90 minutes daily to the primary academic content areas of reading, mathematics, and writing. Time is allocated to ensure adequate time for instruction, practice, and application of skills. Teachers divide time among building tool skills, building component skills, and building compound skills.

**Critical Thinking and Self-Regulation**

Morningside directly instructs and monitors improvement in strategic thinking, reasoning, and self-monitoring skills. Strategic thinking is the glue that allows students to employ component skills and strategies in productive problem solving. Typically, problems provide opportunities for learners to combine known associations, algorithms, and concepts in ways that may not be dictated by an existing formula or that may yield more than one answer. Morningside’s instructional and practice strategies build tool and component skills that are readily accessible for problem solving, but they do not instill the reflective contingencies that are required to adduce relevant knowledge and skills to solve a particular problem.

In opposition to some contemporary educational practice, we directly teach a productive type of thinking called problem solving through reasoning. It involves a dance between a speaker and a listener residing in the same body. Self, as speaker, constantly provides practical, supplementary verbal stimulation to self, as listener. The speaker constructs responses during a series of private volleys, prompts, and probes of the learner’s own behavior. At Morningside, we view the failure to self-monitor and reason during problem solving as a failure of instruction rather than as a failure of the learner. This perspective has
provided a challenge to develop instructional strategies that turn learners into productive thinkers and problem solvers.

In response, Morningside turned to an approach developed by Whimbey and Lockhead in the 1970s (Whimbey & Lockhead, 1999). Procedures for thinking-aloud problem solving (TAPS) were developed originally to improve analytical reasoning skills of college students. Whimbey and colleagues, in *Blueprint for Educational Change* (1993), cite five categories of common errors in problem solving: inaccurate reading, inaccurate thinking, weak problem analysis, lack of perseverance, and failure to think aloud. Robbins (Robbins, 1996; Robbins, 1995) adapted Whimbey and Lockhead's methodology for younger children and has proceduralized the instructional regimen. She has designed scripted instruction to ensure that elementary school students can engage in the reasoning process. During TAPS instruction, teachers model both speaker/problem-solver and active listener roles and verify that students can identify the behaviors associated with the characteristics of each role. Teachers model process errors, and learners catch them. Learners then take turns playing each role, making occasional errors that other learners must detect. The TAPS instructional program begins as a training exercise using complex problem-solving and logic exercises, but as learners’ skills with both repertoires become firm, teachers integrate the procedure into all academic areas. At Morningside, we employ the Robbins TAPS design in three stages. Students first develop each role—problem solver and active listener—in pairs and then learn to serve both roles simultaneously, moving back and forth between offering solutions and then critiquing and encouraging performance, first publicly, then privately.

Morningside teaches TAPS as an underlying repertoire that structures discovery learning. TAPS routines facilitate the combination of elements that have been learned as well-defined tasks. TAPS also helps to reform the isolated skills and clarify how they interlock or chain together. A delayed prompting method can be used to teach such specific reasoning routines, and students can engage in these strategies during TAPS to combine elements and solve problems successfully. Technically, current problems may occasion certain reasoning routines embedded in a generic TAPS repertoire, which in turn may produce the contingencies that result in the appropriate recombination of already-mastered component elements.

**Self-Direction and Independence**

When visitors drop by Morningside Academy, they might very well get the best lesson in the Morningside model from one of the students. This is because Morningside teachers constantly and intentionally strive to make the instructional protocols and decision-making process transparent to parents and students. They encourage students to take charge of their own learning. Students chart their own timings and begin to set their own celeration
aims—sometimes much more ambitious ones than the teacher would have set for them. Students manage their own practice and recommend program modifications. They know when programs are working and when they are not. Students learn important goal-setting, self-monitoring, self-management, organizational, and cooperative learning skills. They also learn self-management and self-determination through the freedom to take their own performance breaks and still meet their expected goals, skip lessons when they can demonstrate mastery, move through the curriculum at their own pace, select their own arrangement of tasks to accomplish in a class period, choose their own free-time activities, and give themselves report card points, among other opportunities.

**PROGRAM PLACEMENT AND MONITORING BASED ON CONTINUOUS MEASUREMENT OF PERFORMANCE**

A hallmark of Morningside procedures is the continuous interplay between instruction and assessment. The primary purpose of assessment at Morningside is to ensure that students are correctly placed in an instructional program and that they are making gains that promise long-term academic benefit. Assessment at Morningside consists of five primary elements:

- Standardized norm referenced assessment, typically administered twice yearly, gauges a learner's performance level in relation to peers and gauges growth using socially validated measures of performance.
- Diagnostic/prescriptive testing determines initial placement in an instructional program.
- Weekly standardized performance probes track progress through a curriculum using the curriculum-based measurement (CBM; Knutson & Shinn, 1991) technology developed at the University of Oregon.
- Daily measurement of accuracy and rate of responding on highly calibrated tool and component skills using PT technology dictate progress through a program.
- Portfolios authentically illustrate and provide social validation of students' progress.

The speed with which students can complete a course of study is dependent, at least in part, on the accuracy of their program placement. Morningside uses a variety of diagnostic and placement instruments to determine a student's placement in a program. Some are more formal than others, and it is common for Morningside teachers and administrators to use more than one measure to ensure correct placement. Morningside also tracks progress on important learning outcomes throughout the year, using CBM protocols. The most common CBM measures are in reading, math, spelling, and writing.
Assessments are timed, and performance is scored according to explicit CBM criteria. Because CBM measures are somewhat more molar than the molecular measures charted on a daily basis in classrooms and because they are more explicitly representative of state curriculum standards and standardized achievement test objectives, they provide external validation of the student’s growth.

By far the most important tool for the teacher in monitoring the effectiveness of instructional programs is Precision Teaching timing and Standard Celeration Charting. The chart provides a picture of the learner’s performance rate, error rate, and growth rate in a standardized format that allows teachers to make decisions at a glance about the learner’s progress (Johnson & Layng, 1992, 1994). Teachers and students use these learning pictures to make daily decisions about what would be best for the learner to do next, how to allocate classroom time, when the curriculum step or slice is too large or too small, and when the timing protocol requires modification.

CLASSROOM MANAGEMENT

Morningside classrooms are busy, high-energy classrooms, and teachers manage many instructional activities simultaneously. Morningside’s commitment to learning and, in many cases, to making up for lost time means that students need to be academically engaged throughout a class period. There is no time to waste, no time for disruptive behavior. All of this means that teachers need to establish workable classroom routines, and students need to develop a battery of skills that support their learning.

The Daily Report Card has evolved since Morningside’s inception (Johnson, 2004) to become the first line of prevention in Morningside teachers’ classroom management toolbox. At the beginning of each day, students receive a daily report card on which is listed their classes and expectations for the day. Students earn points for demonstrating four categories of skills: academic, learning, organization, and citizenship. Students receive academic skill points for their ability to meet or exceed performance aims and to demonstrate adduction or generativity. Learning skills points reward “looking like a student,” answering on signal, following along with the lesson, and following directions. Organization skill points acknowledge students being on time, having their materials, keeping their work area neat, and charting in pencil. Citizenship points are awarded for using appropriate language, respecting other people’s space, using one’s voice appropriately, using one’s body appropriately, and respecting physical property.

Behavioral expectations are clearly stated and posted on the wall. At the beginning of the school year and when new students enter the program, teachers model and prompt the behaviors that are expected. Teachers distribute points in each category throughout each class period. At the end of the
day, students meet individually with their teachers to discuss their academic progress and comportment. The teacher writes summary comments, and students take the report cards to their parents, who in some instances sign and return them. As students transition to project-based learning, daily reports become less structured.

**EMPIRICAL DATA SUPPORTING TECHNOLOGY TRANSFER OF THE MORNINGSIDE MODEL OF GENERATIVE INSTRUCTION**

Four features of the Morningside Model constitute its core: (1) learner-verified instructional methods, tools, and programs are incorporated for basic academic and learning skill development; (2) a significant amount of school time is allocated to practice, using fluency and celeration; (3) children learn reasoning, thinking, problem-solving, research, and cooperative learning skills; and (4) children are transitioned into more independent learning environments in which they apply their basic academic, reasoning, research, and cooperative skills to learning social studies, science, and literature, according to their interests. Morningside Academy arranges such a learning environment for its children and youth, and they make enormous progress in school.

The remarkable results of Morningside Academy's initial 11-year study of its children's mean standardized test gains in reading, language arts, and mathematics have been reported elsewhere (Johnson & Layng, 1992, 1994). Reading averaged 2.5 years growth per school year. By the end of the study, language arts gains approached an average of 4 grade levels and mathematics gains rose to more than 3 grade levels of improvement per school year. Morningside completed its formal lab school evaluation process in the spring of 1992. Currently, it assesses its students in September and June on a variety of in-house, state, and national measures. Children's median achievement test performance remains above 2 grade levels per year in reading, language arts, and math.

Since 1991, Morningside Teachers' Academy (MTA) has successfully implemented programs with over 17,000 students in Illinois, Washington, Georgia, Pennsylvania, British Columbia, South Dakota, and Oklahoma. Students in the Chicago Public Schools, the Nechako School District in British Columbia, the Seattle Public Schools, DeKalb County (Georgia) Public Schools, and elsewhere have profited from our services. MTA has also contracted with several First Nation and American Indian schools in British Columbia, Washington, South Dakota, and Oklahoma, helping them to develop programs in their schools and adult literacy centers. Adult learners in the City Colleges of Chicago and at Motorola Corporation in Phoenix have also made enormous strides in their reading, writing, reasoning, and math skills. Student achievement test results will be reported in a forthcoming book (Johnson and Street, 2004, in press).
CONCLUSION

Morningside has more than 20 years of commitment to the ideals of educational accountability and empirically verified approaches to instruction, one that places the responsibility for student success squarely on the shoulders of educators. Schools are not to act as mere sorting machines for determining degrees of natural, genetic, or pre-existing talent. Instead they must act so that all children are literate and can function as effective citizens in a democracy. We have inherited children who have been left behind, and we have helped them catch up and move ahead. We have believed that if the child was not learning, we were not teaching. We have included learners in the teaching process, as coaches with each other. We have reformed our practices until the evidence revealed that our practices work for kids. And, we have shared the effective practices with others whose children were left behind. We have stood behind our practices, offering parents money-back learning guarantees. In the end, we believe we have helped define what it really means to leave no child behind.

Acknowledgments

This chapter is a short version of a book (Johnson & Street, 2004) that is being published concurrently. Dr. Kent Johnson is the executive director of Morningside Academy and Morningside Teachers' Academy. Dr. Elizabeth Street is a professor of psychology and executive assistant to the president at Central Washington University. She also has been a consultant to Morningside Academy throughout its existence. We are indebted to Drs. Joanne Robbins, T. V. Joe Layng, and Warren Street for their thoughtful comments and editorial assistance during the development of this book chapter. We also are indebted to the many teachers at Morningside and at partner schools and agencies whose feedback and insights have strengthened the Morningside Model, and to the students whose learning has been the most important feedback of all.

References


According to the latest government data from the National Center for Educational Statistics, only 36.5% of students with the goal of earning a Bachelor's degree who enrolled in a 4-year institution in 1996 completed that degree within 4 years (Berkner, He, Cataldi, & Knepper, 2002). This figure is troubling for several reasons. Students who do not complete their degrees within 4 years have wasted a large amount of money and time, both theirs and society's. Colleges and universities are heavily subsidized by taxpayers, and the costs of education are justified only if our institutions are producing graduates prepared for employment in their field of study.

What factors predict timely completion of a bachelor's degree? "The characteristics of students who were mostly likely to graduate within four years with a bachelor's degree fit a commonly held perception of what a college student looks like—he or she receives a good academic preparation in high school, enters college immediately after high school, enrolls in college full time, and is continuously enrolled" (Berkner et al., 2002, p. x). Other factors that predict success include type of institution attended, gender, race, and dependent family income.

As a college professor, there is nothing that I can do to change these predictors, except perhaps for one. "Students who entered college with good academic preparation—those who received mostly As in high school took two
or more Advanced Placement tests, or had high SAT scores—also had higher completion rates than others. About 80 percent competed a bachelor’s degree within 6 years, and more than one-half (55 and 61 percent) graduated within 4 years” (Berkner et al., 2002, p. vii). Although it is clear that good academic preparation increases a student’s chances of success, what is unclear is how that works. Do students with good academic backgrounds have better study and study management skills? If so, perhaps it is those skills that help them persist. Traditional college courses rely much more heavily on student skills than the typical high school course. Most professors still use the “lecture, assign, and test” method. It is up to the students to design their own study materials and use them effectively to prepare for tests.

When I asked the students in my CPSY 101 course (a freshman Community Psychology course) if they had ever been taught how to study, only one or two raised their hands. I then tested their skill at writing study questions by asking them to prepare a test over an assigned reading. Most wrote very simple, definition-type questions, unlike the type that professors typically ask. Few had ever used flash cards, and none reported keeping track of their progress as they studied for tests. Studying is hard work. If students see little or no progress when they study, they may find it difficult to persist.

So, I decided to teach them two skill sets: (1) how to study for college tests, and (2) how to track their own progress as they studied. I taught them how to write study questions like the kind that a professor might ask and how to use those materials to prepare for college tests. I also taught them how to measure, chart, and evaluate their own progress as they studied for tests. I called this progress measure their **learning efficiency**.

**WHAT IS LEARNING EFFICIENCY?**

Learning efficiency is a measure of improvement in performance accuracy and speed per amount of learning time. To calculate learning efficiency, we must measure two things: (1) performance improvement, and (2) learning time:

\[
\text{Learning Efficiency} = \frac{\text{Performance Improvement}}{\text{Learning Time}}
\]

Performance improvement is the improvement in a learner's accuracy and speed as a result of that learner's interaction with a learning program. Suppose two students take their final exams on the first day and the last day of class. The improvement in their test performance is a measure of how much they learned during the course. One student correctly answered 1 question per minute on the pre-test and 10 questions per minute on the post-test, achieving an improvement of 9 correct per minute. The other correctly answered 9 questions per minute on the pre-test and 10 per minute on the post-test, achieving an improvement of 1 correct answer per minute. Which one learned more? The
student that made the most improvement learned the most. Performance improvement allows us to assess the effectiveness of a course but not its efficiency. To measure learning efficiency, we need to measure learning time.

Learning time is the amount of time that learners spend interacting with the learning program. Consider two additional students in our course. Both made the same amount of improvement, from correctly answering 5 questions per minute on the pre-test to 10 questions per minute on the post-test, but was the course equally efficient for both? Suppose that we asked both students to record the number of minutes each spent studying for the course. The data revealed something interesting: One student spent twice as much time as the other to get the same improvement. That second student displayed greater learning efficiency by achieving the same performance improvement in one-half the learning time.

Learning efficiency is a measure of academic progress that takes into account both performance improvement and the student's learning time. The higher the learning efficiency, the less time it takes for learners to achieve competence. If students were to collect data on their own learning efficiency, such data would provide them with more feedback on their progress as they prepare for tests and could also be used to improve the efficiency of their studying.

THREE POSSIBLE CAUSES OF POOR LEARNING EFFICIENCIES AND THEIR SOLUTIONS

We can group the causes of poor learning efficiency into three categories: (1) students who cannot learn efficiently because professors do not provide the information and tools they need; (2) students who do not want to learn efficiently because the consequences produced by attending class and studying are weaker than the consequences for alternative activities; and (3) students who do not know how to learn efficiently because they have never been taught efficient study and study management techniques. To solve the "can-do" problem, professors can provide students with measurable learning objectives, well-designed instructions, practice, and feedback, peer tutoring, and multiple testing opportunities. When professors adopt this approach, sometimes referred to as the Personalized System of Instruction (PSI) (see Chapter 12), more students achieve mastery; however, the cost can be prohibitive. Professors and student helpers devote many more hours designing instructional and testing materials, tutoring, and grading than is required by the traditional college course. Perhaps for this reason, few professors currently run PSI courses.

To solve the "want-to" problem, professors can design their courses to increase the amount of time students spend attending class and studying. They can do this by giving weekly tests and tying the course grade very tightly to
studying for the tests. Some versions of PSI set weekly deadlines for passing tests. Although this approach does get students to spend more time studying, it is also very labor intensive. More frequent testing requires more frequent grading. Professors who use PSI either spend many additional hours grading papers or restrict their tests to machine-graded, multiple-choice formats.

To solve the "know-how" problem, professors can teach their students how to study and how to manage their own studying. One might assume that most students who enter college have already learned these skills, as they have already completed 12 years of schooling; however, as mentioned previously, only one or two students in my CPSY 101 course report being taught how to study. Teaching students how to study and manage their own study is the least expensive solution to the problem of poor learning efficiencies than the other two solutions. If students can be taught how to design and use their own learning materials, professors can continue to teach the traditional way: lecture, assign readings and papers, and occasionally test. For this reason, the author has begun to teach his students how to study for tests in his CPSY 101 course.

**CASE STUDY: EVALUATING LEARNING EFFICIENCIES IN A CPSY 101 COURSE**

At the beginning of the semester, I announced that the course objectives included learning how to study for college tests and how to evaluate the student's own learning efficiencies while preparing for those tests. Students were given examples of different types of test questions that any professor might ask. These included questions that asked students to compare or contrast problem-solving theories, questions that asked students to distinguish between examples of different types of problems, and questions that asked students to design solutions to those problems. Their weekly assignment was to write questions like those that the professor might ask using the information in the assigned reading and the kind of questions provided by the professor in lecture and handouts. I then asked the students to read their study questions to the class and commented on their quality. Students turned in typed copies of their study questions to earn points for both quantity and quality. See Fig. 1 for examples of study questions from three different students in the CPSY 101 course.

Students were also taught how to use flash cards to practice with their study questions. The professor demonstrated how students could use a variation of delayed prompting by looking at a question on the front of a flashcard and then, if unsure of the correct answer, turning the card over and reading it. Delayed prompting is a procedure that has been demonstrated to increase the learning efficiencies of students with learning difficulties (Godby, Gast, & Wolery, 1987).

Students were required to record their number of correct answers to study questions, number of "not yet" (errors), number of corrects per minute,
Sample of RM’s Study Questions for the Final Exam

Card Front: Q: What is pinpointing?
Card Back: Ans: Define behavior and results specifically enough to measure them

Sample of Learner JN’s Study Questions for the Final Exam

Front of Card: Pinpoint
Back of Card:
1. Define results objective-nail length of 8-11 mm
2. How will I measure it?-millimeter side of a ruler
3. Compare it to my current results-Nail length of 6 mm
4. Define performance objective-Crochet Afghan 6 rows/week
5. How will I measure it? Count # of rows per week
6. Compare it to my current performance? # of rows per week is 1.5

Sample of Learner RB’s Study Questions for the Final Exam

<table>
<thead>
<tr>
<th>Front of Card</th>
<th>Back of Card</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pinpoint-Define results</td>
<td>I want to lose 4 pounds by the end of the semester</td>
</tr>
<tr>
<td>Pinpoint-Measure results</td>
<td>Current Weight: 132 pounds</td>
</tr>
<tr>
<td>Pinpoint-Evaluate results</td>
<td>I want to weigh: 128 pounds</td>
</tr>
<tr>
<td>Pinpoint-Define performance objective</td>
<td>Gap: 4 pounds, 100% gap</td>
</tr>
</tbody>
</table>
| Pinpoint-Measure performance objective | A-running  
C-2.5 miles  
E-on cardio machines at Eastman  
S-3 times a week |
| Pinpoint-Measure performance objective | I am going to Eastman 1-2 times per week for 20 minutes  
Gap of 1-2 time per week of 10 min each |
| How/What did I pinpoint? (6 steps) | Define results  
Measure results  
Evaluate results  
Define performance objective  
Measure performance objective  
Compare current results to performance |

FIGURE 1

Samples of study questions for CPSY 101 Final Exam from three different students in the course, RM, JN, and RB.

number of “not yets” per minute, and the total amount of time they spent studying on a special learning efficiency recording sheet. See Fig. 2 for an example of the learning efficiency recording form. The students were then taught how to plot these data on a standard learning efficiency chart, so that they could get feedback on their learning efficiencies for each test.

I have provided examples of learning efficiency charts from three students who prepared study questions and practiced with them for the course final
exam. See Fig. 3 to compare the learning efficiencies of RM, JN, and RB. Each chart shows the improvement in the number of correct responses per minute and the number of “not yets” per minute (errors) per minute of student learning time as a slope that runs from the left to the right side of the chart. A comparison of the slopes of the correct-per-minute lines for each learner reveals that learner RM had the highest learning efficiency (the steepest slope) and learner RB had the lowest learning efficiency (the shallowest slope). The charts also show that RB had much higher corrects and “not yets” per minute at the beginning of practice (on the left side of her chart) than the other two students. However, these practice data do not tell the whole story. Figure 4 shows learning efficiencies plotted from the pre-test score, post-test score, and learning time for each student. In this chart, it is clear that student JN, the student with the second steepest learning efficiency, scored highest on the post-test, because she had a greater improvement from pre- to post-test than RM.

An examination of JN and RM’s flash cards in Fig. 1 can account for the discrepancies between their practice learning efficiencies and their scores on the post-test. Because the exam asked the students to describe how they would solve a performance problem, using an example, JN’s study materials allowed for more appropriate practice than RM’s. JN’s answers much more closely approximated the professor’s.

### WHAT THE DATA CAN TELL US ABOUT HOW TO IMPROVE LEARNING EFFICIENCIES

Collecting data on student learning efficiencies is only the first step in an effort to discover ways to improve learning efficiency. Now that we have a way to measure and evaluate learning efficiency, we can begin our search for those factors that speed or slow learning. What advice could we give to RB, the student with the lowest learning efficiency (the flattest slope in both Figs. 3
FIGURE 3
Practice learning efficiencies for students RM, JN, and RB.
and 4) and the highest "not yets" per minute at the beginning of practice? Here is what we know based on the handful of studies that have been done so far. Several studies have shown that delayed prompting is more efficient than other types of prompting techniques. Not only is delayed prompting more efficient, but it also tends to produce fewer "not yets" (errors) than other practice methods (Godby et al., 1987). Is RB looking at the question on the front of the study card, then, if she is unable to answer the question correctly after a short delay, is she turning the card over to read the correct answer on the back of the card when she practices? Although the professor demonstrated this technique, there was no guarantee that all students were using it correctly. One way to find out is to ask RB to demonstrate her flashcard technique. At least one study suggests that using an error correction technique (additional practice of missed questions) also improves learning efficiency (Johnson, Schuster, Bell, 1996). Is RB spending additional time practicing with her "not yet" cards? We can check whether RB is using the error correction procedure by asking her to demonstrate it to us.

LEARNING EFFICIENCY GOES TO COLLEGE

So, as a reader what should you conclude from this chapter on learning efficiency? (1) College students can be taught how to study efficiently and to evaluate their own learning efficiency. (2) Many reported that the flashcard method helped them prepare for the tests; some even reported that they had begun to use the same method to study for their other classes. (3) This method
of collecting data on learning efficiency is fairly easy to implement, as the students are collecting their own data, but, of course, we cannot be sure that the data are accurate without reliability checks by independent observers. (4) At this point, there is insufficient research on variables that affect learning efficiency to guide our efforts to improve it. I hope I have made a case for the need to conduct such research.

References
INTRODUCTION

Language and cognition are important domains in the discipline of psychology and are often the primary focus in the study of psychological development and in the design of programs of remedial education (Lovaas, 1981; Piaget, 1967). The Applied Behavior Analysis (ABA) approach to autism, for example, considers the establishment of language skills a primary treatment goal because these abilities are prerequisites for most other types of learning and are frequently deficient in autistic populations (Sundberg & Michael, 2001).

Most ABA approaches to language training are based in large part on Skinner’s (1957) definition of verbal operants, including mands, echoics, textuals, transcription and dictation-taking, intraverbals, tacts, extended tacts, and autoclitics. According to Skinner, verbal operants differ from other operant responses because reinforcement in the former is provided indirectly through a social mediator, rather than directly through environmental contingencies. This approach to verbal behavior is functional in that it does not conceptualize
language as a translation of "meaning," and this distinction has important implications for how verbal behavior is established. In contrast, an approach to training language that emphasizes "meaning" (rather than function) rests on the assumption that spontaneous transfer from one verbal operant to another will occur once a child is taught the meaning of a word (Sundberg & Michael, 2001). For example, a semantic theory of language would predict that learning a verbal operant such as a tact (e.g., uttering "juice" in the presence of actual juice) could provide the child with words as tools that can then be applied with other verbal operants such as a mand (e.g., asking for juice). Based on this assumption, training the verbal skill of tacting will likely produce the collateral verbal skill of manding or vice versa; however, children with severe language difficulties do not typically utilize words in the manner in which the words-as-tools analogy suggests (although spontaneous transfer from one verbal operant to another may be facilitated by learning in more sophisticated speakers; Sundberg & Michael, 2001). Indeed, Skinner (1957) pointed to the general imbalances found between listening and speaking skills and between reading and writing skills as evidence for distinct verbal functions that do not automatically transfer across behavioral domains. Training regimes based on the establishment of Skinner’s verbal operants, therefore, characteristically attempt to establish verbal operants such as mands and tacts as distinct verbal repertoires (Greer, 2002).

Although Skinner’s approach to verbal behavior was widely accepted as the first comprehensive account of language based on operant principles and was incorporated into the design of many (if not most) behavioral interventions, it has met with criticism within and beyond the behavioral tradition. Chomsky (1959), for example, argued that Skinner’s approach to verbal behavior could not adequately account for the generativity of language found even in young children. More recently, a number of behavioral researchers have suggested that Skinner’s account ultimately failed as a basis for an experimental analysis of verbal behavior (Hayes, Barnes-Holmes, Roche, 2001a).

The core argument offered by Hayes et al., (2001a) is that Skinner’s definition of verbal behavior was too broad. Specifically, Skinner’s definition of verbal operants rests on the mediation of reinforcement of verbal responses by a listener or listeners who have been conditioned to mediate verbal responses in a particular way. This definition turns out to be extremely broad, however. And, indeed, Skinner appears to have been aware of this fact when he argued that the interaction between a human experimenter and a non-human subject constitutes a small verbal community in which the former supplies reinforcement in much the same way as a listener. Although appealing in its simplicity, particularly for a basic science of behavioral psychology that was primarily focused on the behavior of nonhumans, a definition of verbal behavior that effectively includes most human–nonhuman interactions renders superfluous the conceptual class verbal.
One solution to this problem is to redefine verbal events in terms of what are called derived arbitrary stimulus relations (Barnes-Holmes, Barnes-Holmes, Cullinan, 2000; Chase & Danforth, 1991; Hayes et al., 2001). According to this view, the derivation of multiple stimulus relations (explained subsequently) is believed to be the core process involved in human language and cognitive abilities. As well as offering a new functional definition and experimental analysis of verbal behavior, some researchers have suggested that it may be possible to harness these newly defined verbal processes in the development of programs for remediating deficits in language and cognition (Barnes-Holmes, Barnes-Holmes, Cullinan, 2001a). In the first part of this chapter, we outline the basic features of this approach, known as Relational Frame Theory (RFT). We then turn our attention to RFT-driven research that has direct implications for the design of empirically supported educational methods.

RELATIONAL FRAME THEORY

Defining Derived Relational Responding

At its simplest, relating may be defined as responding to one event in terms of another and is demonstrated readily by nonhuman and human organisms alike (Reese, 1968). Most living organisms, with a history of appropriate training, are able to respond to the nonarbitrary relations among the physical properties of stimuli or events. For example, adult rhesus monkeys can be trained to select the taller of two stimuli (Harmon, Strong, Pasnak, 1982), and this type of behavior may be described as relational responding because it relies upon relative comparisons among the stimuli involving discriminations based on their formal properties (i.e., one stimulus is actually physically taller than the other).

Relational responding, however, becomes more complex when it is brought under the control of contextual features beyond the formal (nonarbitrary) properties of the stimuli. Arbitrarily applicable relational responding is the term used to describe relational responses that can be brought to bear on any stimuli presented in an appropriate context. For example, if you are told that "X is taller than Y," you can derive that Y is shorter than X without actually seeing what X and Y refer to. In this way, arbitrarily applicable relational responding may be controlled by contextual cues that are modified entirely by social whim. For instance, in a game, children may be instructed that "Tall means short and short means tall." In this case, the relational functions of tall and short are applied in a purely arbitrary fashion and are not governed by the actual formal dimensions of the stimuli.

In order for relational responding to come under appropriate forms of contextual control, as would be required for competent performances in the
example of the children's game above, children must learn to discriminate between the relevant features of the task (i.e., responding relationally to events in the presence of appropriate contextual cues) and the irrelevant features of the task (e.g., responding to the physical properties of the stimuli). This training history is clearly illustrated in the establishment of the bidirectional stimulus relations that emerge between words and their referents that form a large part of early naturalistic language interactions.

Early experiences of learning to name objects comprise a wealth of name-object and object-name relations across an extensive range of objects and names. That is, young children are trained as follows: Given the name of an object → select the object (the name–object relations are explicitly trained), and given the object → select the name of the object (in this case, the object–name relations are explicitly trained). In essence, reinforcement is being provided for responding in accordance with the bidirectional relations between object names and actual objects and vice versa. Reinforcement for such bidirectional responding is rich in early natural language interactions but occurs only in certain contexts, such as in the presence of phrases such as “What's that?” and the juxtaposition of objects and words. According to RFT, this training in bidirectional relations ensures that in certain contexts name–object relations reliably predict object–name relations and vice versa, and generalized bidirectional responding emerges. For example, explicit training in a new name–object relation (given the name "teddy" → select the teddy) may result in the derived or untrained object–name relation (given the actual teddy and asked "What’s this?" → say the name "teddy"). This training history is brought to bear on novel stimuli by the presence of specific contextual cues (e.g., “What’s this?”) that control responding in accordance with the bidirectional stimulus relations. Note that in the example of naming, the stimulus relations are entirely arbitrary because, in practically all cases, the words do not bear any formal resemblance to the actual objects to which they refer (i.e., the word "teddy" looks nothing like an actual teddy).

From the perspective of RFT, arbitrarily applicable relational responding has three defining properties: mutual entailment, combinatorial entailment, and the transformation of stimulus functions. The term **mutual entailment** is used to describe the basic bidirectionality of relational responding outlined previously. Arbitrary stimulus relations are always mutual—if A is related to B, then B is always related to A. That is, if the first relation is specified, the second relation is entailed (hence the term **mutual entailment**). Technically defined, mutual entailment applies when, in a given context, A, for example, is related in a characteristic way to B and, as a result, in that context B is related in another characteristic way to A.

The term **combinatorial entailment** is used to describe a derived stimulus relation in which two or more relations mutually combine. For example, if you are only told that A is more than B and B is more than C, then you can derive that A is more than C and C is less than A. From a developmental or educational
perspective, it is likely that combinatorially entailed relations emerge later than mutually entailed relations.

The third defining feature of arbitrarily applicable relational responding is the *transformation of stimulus functions*. This term is employed when the functions of a given stimulus are modified or changed as a result of derived relations with other stimuli. If, for example, an individual is told that B is the opposite of A, and a conditioned *punishing* function is attached to A, the functions of B may be transformed such that it becomes a conditioned *reinforcer* because of its participation in a relation of opposition with A (Roche & Barnes, 1997; Roche, Barnes-Holmes, Smeets, Barnes-Holmes, & McGready, 2000).

Just as the relational response is controlled by context, the transformation of stimulus functions must also come under contextual control. Consider some of the perceptual functions of a lemon, including its bitter taste, its rough texture, and the fact that it is bright yellow in color. When an individual is asked to imagine a lemon, many of these perceptual features become psychologically present. In the technical language of RFT, this psychological event is described as follows. The word “lemon” and actual lemons participate in what is called a relational frame of coordination. In addition, the words “imagine a” function as a context in which some of the perceptual functions (especially visual functions) are elicited based on the relational frame. In another context (e.g., “imagine tasting a . . .”), other functions would be elicited. Contextual cues, therefore, control not only the type of relational frame involved but also the transformation of functions that are enabled by the frame in question.

**Relational Framing**

Relational Frame Theory employs the generic term *relational frame* to describe particular patterns of arbitrarily applicable relational responding (Hayes & Hayes, 1989). A number of generic relational frames have been discussed in the literature (although others may yet be identified). These include frames of coordination, opposition, distinction, comparison, hierarchy, and deictic frames of perspective-taking (Hayes, Barnes-Holmes & Roche, 2001b). Perhaps the most commonly known pattern of relational responding involves the frame of *coordination*, in which the relations are ones of identity or similarity. The example of naming described previously is an often-cited instance of the frame of coordination, and this frame is probably one of the first to be established.

The relational frame of *opposition* requires the abstraction of a dimension along which stimuli can be ordered and distinguished in equal ways from a reference point. In natural language, *opposite* typically implies the relevant dimension. For example, saying that “fast is the opposite of slow” implies that speed is the dimension along which the related events are to be ordered. Furthermore, RFT suggests that the frame of opposition will emerge later than coordination because the combinatorially entailed relations within frames of opposition are frames of coordination. For example, if fast is the opposite of
slow, and slow is the opposite of quick, then fast and quick are the same (i.e., coordinated), not opposite.

Relational frames of **distinction** involve responding to the differences among stimuli or events, typically also along a particular dimension; however, in frames of distinction, the relevant dimension is rarely implied. Consider, for instance, the statement: "This is not a person of average intelligence." Based on only this information, it is not possible to determine whether this person is of extremely high or extremely low intelligence. Furthermore, a combinatorially entailed difference relation is unspecified. For example, if you are told that A is different than B, and B is different than C, then A and C may be the same or different. This type of unspecified relation is a defining property of the frame of distinction.

There are many specific types of frames of **comparison**, including bigger and smaller and faster and slower. These frames involve responding to events in terms of a quantitative or qualitative relation, again along a specified dimension (e.g., speed). If I say that "a lion is faster than a dog and a dog is faster than a mouse," the events can be compared along the dimension of speed, and you can derive that "the lion is faster than the mouse and the mouse is slower than the lion." Furthermore, comparative relations may be made more specific by quantification of the dimension. For example, if you are told that "a lion is twice as fast as a dog and a dog is twice as fast as a mouse," you can now derive that the lion is exactly four times faster than the mouse and the mouse is four times slower than the lion.

The final family of relational frames that we will describe here involves **deictic relations**, which appear to be involved in perspective-taking (Barnes & Roche, 1997; Hayes, 1984). The three key deictic frames seem to be I and YOU, HERE and THERE, and NOW and THEN. Relational Frame Theory argues that taking a perspective involves responding in accordance with these relational frames. In other words, taking a perspective involves responding from I, HERE, and NOW with respect to events located THERE and THEN. For example, saying "I cannot see what you see" requires that the speaker distinguish between I and YOU (i.e., that we do not always see the same things), and that what I am seeing HERE and NOW is not what YOU are seeing THERE and THEN. We will return to this type of relational framing toward the end of the chapter.

There is much more to RFT than has been presented thus far. For example, RFT also describes the relating of relations and the relating of relational networks with other relational networks. These processes are believed to account for competence in verbal skills such as analogy, metaphor, and story telling, but it is beyond the scope of the current chapter for these issues to be described fully (but see Hayes et al., 2001a, for a book-length account). The important point to be gleaned here, however, is that from the RFT perspective deriving relations underpins developmental or educational achievement, and a small number of psychological processes are sufficient to yield the full gamut of cognitive skills.
RESEARCH IN RELATIONAL FRAME THEORY AND ITS IMPLICATIONS FOR EDUCATION

From the perspective of RFT, over-arching relational skills can be taught, and subsequent improvement in relational responding should lead to improved abilities in areas of cognition and language, as well as in intelligence in general (Hayes, 1994). The RFT approach to education has two core assumptions. First, skills in relational responding provide the basis for a wide range of cognitive abilities that correlate with educational achievement. Second, multiple-exemplar training provides an important method for harnessing these cognitive skills and building up flexibility in relational repertoires. This section of the chapter briefly describes a number of RFT studies in which multiple-exemplar training was successfully employed to establish novel relational repertoires. In each study, a relational or cognitive deficit is identified in the behavior of the experimental participants, and then the study seeks to remediate this deficit by reinforcing one or more exemplars of the relevant relational repertoire. These studies thus provide examples of how the RFT approach can inform educational practice. We should add that multiple-exemplar training is not exclusive to RFT and indeed is an inherent feature of traditional educational practice (Englemann & Carnine, 1982). For example, children are often presented with tasks that are grouped by content that establish flexibility in over-arching relational skills (e.g., adding numbers together, filling in missing numbers in a sequence, or identifying the nouns in sentences). However, an approach to education based on RFT seeks to identify tasks that can be grouped according to the relational skills involved rather than according to traditional content areas.

Derived Transformations of Function in Accordance with Symmetry

One of the first RFT-based studies that attempted to analyze the development of relational framing in young children involved a systematic analysis of the role of multiple-exemplar training in establishing simple derived relations (Barnes-Holmes et al., 2001c). In this study, young, normally developing children between the ages of 4 and 5 years old were selected and a task was designed to investigate the transformation of function in accordance with symmetry. The children were first trained in an action–object conditional discrimination task. During this training, when the experimenter waved, choosing a toy car was reinforced with feedback (i.e., the trained relation was wave–car), and, when the experimenter clapped, choosing a doll was reinforced (i.e., the trained relation was clap–doll). Following this training, the children were tested without feedback for the derived object–action symmetry relations. That is, when the experimenter now presented a toy car, the child was required to wave (i.e., the tested relation was car–wave), and, when the experimenter presented a doll, the child
was required to clap (i.e., the tested relation was doll–clap). Of the 16 children, 11 failed to show the target derived symmetry performances on the first test. At this point in the study, a multiple-baseline design was used to phase in the introduction of explicit symmetry or object–action training for those children who failed the symmetry test. In other words, after failing to demonstrate symmetry, some children were then re-exposed to the same trials as in the symmetry test, but corrective feedback was provided after each trial. In order to test the effect of this training, the children were thereafter exposed to an entirely new set of actions and objects in the same training and testing format. In simple terms, the children were trained on one exemplar and then tested on another. With the multiple baseline design, some children were exposed to several sessions of training and testing with the novel sets of stimuli prior to receiving the explicit symmetry training in order to determine whether these children would improve in the absence of explicit object–action training across exemplars.

The results of the studies overall showed that for all 11 children who failed the first symmetry test, explicit symmetry training was effective in establishing the derived transformations of function in accordance with symmetry, and that for the majority of children only one exemplar of training was necessary for the derived performance to occur on a novel set of stimuli. As an aside, a number of similar experiments employed an alternative naming intervention commonly used in education and found this to be much less effective than the multiple-exemplar training in establishing the derived test performances (Barnes-Holmes, Barnes Holmes, Roche, Smeets, 2001b).

One issue that arose from this series of studies was the very limited number of exemplars required for the children to demonstrate the target derived performances. This suggested that the exemplar training simply activated a previously established relational repertoire of symmetrical responding, and the age and verbal sophistication of the children supported this conclusion. The obvious limitation of this work, therefore, is that it did not demonstrate the establishment of previously absent repertoires of relational framing, which is often what is required in educational programs. The three studies outlined subsequently address this concern.

**Teaching Derived Manding**

Establishing a manding repertoire is very important for children with language deficits, because it provides immediate control of the social and non-social environment and facilitates the development of speaker and listener repertoires (Sundberg & Michael, 2001). It is not surprising, therefore, that mands are typically the first verbal operants humans acquire naturally or are trained to acquire in educational programs (Bijou & Baer, 1965; Skinner, 1957).

One of the most common difficulties in educational programs that attempt to build manding repertoires in individuals for whom they are found to be
absent is the identification or establishment of a variety of deprivation states. Although many kinds of deprivation may already be present in a child with autism, for example, these may remain unknown to the instructor who is presented with generic responses such as crying and pulling, but not with specific indicators of deprivation (Sundberg & Michael, 2001). Given that children with autism have successfully demonstrated derived relational responding (Eikeseth & Smith, 1990), a recent study attempted to establish derived manding via relational frames in young normally developing and autistic children (Murphy, Barnes-Holmes, & Barnes-Holmes, 2003).

The experimental sequence employed in this study consisted of three phases: mand training, conditional discrimination training, and testing for a derived transfer of mand functions. During Phase 1, participants were trained to use two stimulus cards, each with an abstract symbol on it, to mand for a pink token and a yellow token, respectively (these two cards will be referred to as A1 and A2). A “state of deprivation” was created by presenting participants with a task that required them to mand for the appropriate number of either pink or yellow tokens. That is, the participant was presented with a token mat that contained a number of pink and/or yellow tokens and, to complete a mand training trial successfully, the participant had to mand for only those pink or yellow tokens that were needed to complete the missing set (i.e., if a participant manded for a token that he or she did not need, the trial was recorded as incorrect).

Participants who successfully completed mand training were trained in two conditional discriminations using a matching-to-sample procedure. During this training, the children were taught to relate the symbol on the A1 stimulus card to a second symbol (B1) and to relate this second symbol to a third symbol (C1). The training also involved teaching the children to relate the A2 symbol to the B2 stimulus and the B2 stimulus to a C2 stimulus. In this way, two relational frames of coordination were established (A1–B1–C1 and A2–B2–C2). The critical test from an RFT perspective involved determining if the children would spontaneously use the two C stimuli to mand for the appropriate colored tokens. In other words, would the yellow manding function of A1 transfer via coordination to C1, and similarly the pink manding function transfer from A2 to C2?

The three normally developing children and two of the autistic children readily demonstrated the predicted derived transfer of mand functions on their first exposures to the test; however, one of the autistic children completely failed the derived transfer test (the data for this child are presented in Fig. 1). Consequently, this child was exposed to exemplar training in the derived transfer of mand functions. That is, when the child failed to show the derived transfer, he was immediately exposed to the derived transfer tasks again, but this time corrective feedback was provided after each response (labeled as “Transfer Training” in Fig. 1). Subsequently, the child was re-exposed to the mand training, conditional discrimination training, and derived
**FIGURE 1**

**Mands:** Percent correct across blocks of 6 trials. A1-B1, A2-B2; B1-C1, B2-C2; D1-E1, D2-E2; E1-F1, E2-F2; J1-K1, J2-K2; K1-L1, K2-L2; M1-N1, M2-N2; and N1-O1, N2-O2: Percent correct across blocks of 10 matching-to-sample training trials in each set. **Derived mands:** Test for derived transfer of mand functions A1-C1, A2-C2; D1-F1, D2-F2; J1-L1, J2-L2; and M1-O1, M2-O2. **Transfer training:** Percent correct across blocks of 6 training trials involving direct training of transfer of mand functions. Asterisks indicate break of at least one day between sessions.
mand testing but with a completely novel set of abstract symbols for the A, B, and C stimuli (e.g., Set D, E, and F). In total, the child required five exemplars of explicit derived mand training before successfully demonstrating a derived transfer of mand functions on a novel set of stimuli (Set M, N, and O) in the absence of corrective feedback. As can be seen from Fig. 1, the improvement in derived manding was gradual across exemplars, and this suggests that a genuinely novel relational repertoire was established in the behavior of this child.

This recent research provides a good example of how the relational concepts of RFT, and its emphasis on exemplar training, can be brought to bear on more traditional behavioral approaches to the teaching of verbal behavior. The one child who repeatedly failed the derived transfer of mand functions test clearly demonstrated that directly trained and derived manding may exist as functionally distinct verbal skills and that the latter may require extensive remediation in an educational context to become firmly established in a child’s behavioral repertoire. Clearly, much more work remains to be done, but this recent study indicates that the application of RFT methodologies and strategies to educational research and practice may be of considerable value.

Establishing the Relational Frames of More-Than, Less-Than, and Opposite

An important cognitive skill that children are required to master involves relational reasoning, particularly when that reasoning gives rise to conclusions that were not explicitly taught or instructed. The ability to take generic relational skills and apply them to new content across a range of contexts constitutes an important educational goal. A recent RFT study systematically examined how generative relational reasoning might be established when it is found to be absent in the behavior of young children (Barnes-Holmes, 2001). This study used a basic problem-solving task to test and train derived relations in accordance with the relational frames of more-than, less-than, and opposite. The basic task employed for establishing both frames involved presenting a child with a number of identically sized paper circles (these were referred to as "coins" because the task involved choosing one or more of the circles on the basis of their stated value). On each trial, the experimenter described specific more-than, less-than, or opposite relations among the coins in terms of value (because the coins were actually the same physical size, the comparative values were entirely arbitrary, as is the case with real money). Based on this comparison, the child was then asked to pick the coin that would buy as many sweets as possible. For example, during a more-than trial, the child might be instructed as follows: "If this coin [experimenter points to the first coin designated as coin A] buys more sweets than this coin [experimenter points to coin B], and this coin [experimenter points to coin B again] buys more sweets than this coin [experimenter points to coin C], which would you choose to buy as
many sweets as possible?" In this case, a correct response consisted of the child selecting the first coin (A). This was the format employed for all trials, and each training trial provided corrective feedback. Numerous sets of coins were employed to create multiple exemplars for training the more-than, less-than, and opposite relations and testing the appropriate derived relations.

Three normally developing children each required 30 to 40 experimental sessions before demonstrating responding in accordance with the target arbitrary relations of more-than and less-than. After this extensive training and testing, the children also demonstrated flexible relational repertoires in that they could respond appropriately: (1) when the experimenter pointed to the coins in any direction from left to right or vice versa and from top to bottom and vice versa; (2) when presented with a novel set of three random objects instead of coins (i.e., a different set of objects such as three cups instead of the three coins that had been presented previously; and (3) when asked which coin(s) they would not choose in order to buy as many sweets as possible. The issue of establishing flexibility in relational responding is important from the perspective of RFT because of its emphasis on the relational skill per se rather than the content involved in the response. For example, if a child has learned to correctly derive opposite relations among three stimuli and the training responsible for establishing this skill was conducted with three coins, the child who has learned the relational skill in question should in principle now be able to demonstrate that type of responding with any three objects (such as three cups) without requiring explicit training with those objects. The presentation of tests involving novel stimuli not presented during explicit training, therefore, is an important feature for RFT in testing the emergence of new skills because this type of performance provides evidence of the child's ability to derive the relations in question.

Although the generative and flexible nature of the performances that were established via exemplar training were impressive, given the complete absence of more-than and less-than relational reasoning skills in the children at the outset of the study, the analysis of opposite relations produced an even greater level of relational complexity. In order to study opposite relations, the task was modified. During the first test for opposite responding, a child may have been presented with four coins and asked: "If this coin [coin A] buys many sweets and is opposite to this coin [coin B], and if this coin [coin B] is opposite to this coin [coin C], and if this coin [coin C] is opposite to this coin [coin D], which would you take to buy as many sweets as possible?" The correct answer on this trial involved selecting coins A and C, because A buys many and is opposite to B (so B buys few), B is the opposite to C (thus C is the same as A and buys many), and C is the opposite to D (thus, if C buys many, D buys few).

Three normally developing children each required extensive exemplar training before demonstrating a complex and flexible repertoire of responding in accordance with the target arbitrary relations of opposite. In the final test phases, the children demonstrated appropriate responding: (1) in the presence
of a novel experimenter; (2) when the experimenter pointed to the coins in any
direction from left to right or vice versa, from top to bottom or vice versa, or in a
completely random sequence; (3) when presented with a set of novel objects
instead of coins; (4) when asked which coin(s) they would not choose in order to
buy as many sweets as possible; and (5) when presented with various numbers
of coins or other items up to and including ten. For example, the children were
trained explicitly with three coins, and they were repeatedly tested to deter-
mine that they could derive the relevant opposite relations among these items.
The children were then presented with four coins, five coins, six coins, and so
on up to ten coins. On each number of coins, they were tested first and, if the
performances were weak, they were then trained explicitly on the opposite
relations among that number of coins (e.g., with four coins). After testing on
four coins, for example, they were then presented with five coins and asked to
derive the opposite relations among these. This type of testing and training
continued until the children had been trained and tested in a series of six coins.
After this number, none of the children needed further explicit training. In other
words, when they were subsequently presented with a series of seven coins,
eight coins, nine coins, and even ten coins, they passed the tests immediately
by deriving all of the target opposite relations. After these tests, they were then
presented with a sequence of ten alternative objects, such as ten pens, and
they were immediately able to derive the correct opposite relations among
these items without training. This type of performance clearly indicated the
presence of a complex and flexible repertoire of derived relational responding
in accordance with the frame of opposite.

The foregoing studies provided evidence that highly complex and flexible
repertoires of relational responding in accordance with the relational frames of
more-than, less-than, and opposite may be effectively established with very
young children by a history of multiple-exemplar training. Although the same
basic methodology of training across multiple exemplars was employed to
establish these relational repertoires, a number of features specific to the
various frames were observed. First, responding arbitrarily in accordance with
more-than and less-than appeared easier than responding in accordance with
opposite. Second, responding in accordance with nonarbitrary more-than and
less-than relations was useful in establishing the more complex arbitrary rela-
tions (e.g., different numbers of sweets were placed on top of the coins to
create actual comparisons of more-than and less-than in order to facilitate the
transition between nonarbitrary and arbitrary responding). Third, many exem-
plars of training were needed to establish even mutually entailed opposite
relations. Fourth, training combinatorially entailed opposite relations was even
more difficult than mutually entailed relations. Fifth, explicit instructions with
regard to the relation of "same" helped to facilitate combinatorially entailed
opposite relations (i.e., if A is opposite to B, and B is opposite to C, then A and
C are the same). Sixth, participants required many exemplars of training to
derive the opposite relations between two, three, four, and five coins but
needed few or no exemplars of training when working with six, seven, eight, nine, or ten coins (i.e., this would suggest that increasing the number of trained coins helped establish opposite responding as a generalized cognitive skill that could be applied arbitrarily, in principle, to any number of stimuli).

The types of studies outlined so far address what appear to be clear examples of generative or novel verbal behavior and relational reasoning. Relational Frame Theory, as an account of human verbal behavior, is directly relevant to these domains; however, RFT also approaches cognitive skills that do not immediately appear to be primarily relational in largely relational terms. Although such an approach may seem counterintuitive, preliminary empirical work in the domain of perspective taking, or what cognitive psychologists call the Theory of Mind (ToM), suggests that there may be some value in adopting a relational interpretation of this phenomenon.

**Relational Responding and Perspective Taking**

According to RFT, cognitive perspective taking (Howlin et al., 1999) involves increasingly complex forms of contextual control of the perspective-taking relational frames of I–YOU, HERE–THERE, and NOW–THEN. As was the case with the establishment of the frames described previously, RFT would predict that the most effective means of establishing these repertoires would be to target the relational frames or generic relational repertoires directly.

As part of a complex research program on perspective taking in children, Barnes-Holmes (2001) attempted to establish the relational skills that appear to underlie perspective taking in young children. In this study, responding in accordance with the relational frame of I–YOU was the first perspective-taking frame to be targeted directly. For illustrative purposes, consider the following simple scenario. The participant was presented with two colored blocks and asked: "If I [experimenter] have a green block, and YOU have a red block, which block do I have? Which block do you have?" If the child responded incorrectly to either question, corrective feedback was provided until correct responding was established. Once these simple I–YOU relations were established, the I and YOU relations were reversed in order to facilitate flexibility in this type of relational responding. The participant, for example, was asked: "If I was YOU, and YOU were ME, which block would YOU have? Which block would I have?"

Once simple and reversed I–YOU relations were established, responding in accordance with HERE–THERE and its combinations with I–YOU was targeted directly. Consider the following scenario: "If I am sitting HERE on the black chair, and you are sitting THERE on the blue chair, where are YOU sitting? Where am I sitting?" With simple HERE–THERE relations in place, reversed HERE–THERE relations were then targeted. During these trials, for example, the participant may have been asked: "I am sitting HERE on the black chair, and you are sitting THERE on the blue chair. If HERE was THERE, and THERE was HERE, where would YOU be sitting? Where would I be sitting?" In this
particular trial type, it is apparent that the HERE–THERE relation is reversed, but the I–YOU relation remains simple. When responding to this type of complex HERE–THERE reversal was established, the task was made even more complex by reversing both I–YOU and HERE–THERE statements simultaneously. Consider the following example of what was called a double-reversed I–YOU/HERE–THERE relation: "I am sitting HERE on the blue chair and YOU are sitting THERE on the black chair. If I was YOU and YOU were ME, and if HERE was THERE and THERE was HERE, where would YOU be sitting? Where would I be sitting?"

Once the perspective-taking frames of I–YOU and HERE–THERE were established, the relational frame of NOW–THEN was targeted. One feature of NOW–THEN responding that differed from HERE–THERE responding is that I and YOU could not be presented together in each trial because responding to I–YOU and NOW–THEN simultaneously renders some of the relations unspecified. In order to establish simple patterns of responding in accordance with NOW–THEN, the participant, for example, was presented with the following scenario: "Yesterday I was watching TV, today I am reading. What am I doing NOW? What was I doing THEN?" Once this pattern of simple NOW–THEN responding was established, the relation was reversed as follows: "Yesterday I was watching TV, today I am reading. If NOW was THEN, and THEN was NOW, what would I be doing NOW? What would I be doing THEN?"

With flexible patterns of NOW–THEN responding established, NOW–THEN and HERE–THERE were mixed to produce new types of double-reversed relations. Consider the following example: "Yesterday, I was sitting THERE on the red chair; today I am sitting HERE on the green chair. If HERE was THERE, and THERE was HERE, and if NOW was THEN and THEN was NOW, where would I be sitting NOW? Where would I be sitting THEN?"

In the Barnes-Holmes study, two normally developing children were exposed to these relational perspective-taking procedures. One 7-year-old female mastered the entire training protocol but required training on the reversed and double-reversed relations. A 3.5-year-old boy was also exposed to I–YOU and HERE–THERE trial types and required extensive training across exemplars, particularly on the reversed and double-reversed relations. In a more recent replication of this work, McHugh, Barnes-Holmes, & Barnes-Holmes, (2003) demonstrated that extensive and systematic exemplar training was necessary to establish even simple NOW–THEN relations in a 4-year-old child.

Although these RFT data on the teaching of perspective-taking are preliminary, the protocols that have been developed in this research have been subjected to systematic empirical analysis using cross-sectional developmental methodologies (McHugh, Barnes-Holmes, & Barnes-Holmes, in press). The findings from this research suggest that the relational skills that are involved in the perspective-taking frames are required in order to successfully complete ToM tasks that have typically been used to study and teach perspective taking.
in educational contexts. Furthermore, additional RFT protocols are currently being developed to study more advanced forms of perspective taking, including false belief and deception (Barnes-Holmes et al., in press). Treating perspective taking as an inherently relational activity, therefore, appears to promise new insights and methodologies for studying and teaching this poorly understood and complex human skill.

**SUMMARY AND CONCLUSIONS**

Based on a relatively small array of psychological and behavioral processes, Relational Frame Theory allows even complex verbal events, such as cognitive perspective taking, to be approached behaviorally and established systematically. In this chapter, we have outlined preliminary findings from a research agenda in the experimental analysis of human behavior that has clear and widespread implications for empirically validated educational practices. This exciting research initiative consists of studies in which both simple and complex forms of derived relational responding were targeted for assessment and remediation using interventions indicated by RFT. A key theme running throughout the diverse content areas covered in this chapter is the role of a basic understanding of relational responses in teaching important cognitive skills in both children and adults. It is our belief that identifying the core relational units involved in these cognitive skills and targeting their fluid and flexible development with appropriate training will lead to significant improvements in the methods used in many educational settings.

**References**


