THE QUARTERLY NEWSLETTER OF THE
LABORATORY OF COMPARATIVE HUMAN COGNITION
William S. Hall and Michael Cole, Editors
Sondra Buffett, Managing Editor
Center for Human Information Processing
University of California, San Diego
January 1980, Volume 2, Number 1

Everyday Memory Tasks in Classrooms for TMR Learners*

Harold G. Levine
Graduate School of Education
University of California, Los Angeles

Andrea G. Zetlin
Mental Retardation Research Center
Neuropsychiatric Institute
School of Medicine
University of California, Los Angeles

L. L. Langness
Mental Retardation Research Center
Neuropsychiatric Institute
School of Medicine
University of California, Los Angeles

It is now relatively common to find critiques of the strict, laboratory-based experimental approach to the study of learning, problem-solving, and memory. The two major sources of concern are that the structure of laboratory tasks are not representative of the everyday tasks with which people in real-life settings must contend, and that with few exceptions laboratory studies seldom control for setting or experimenter effects.

One solution to the dilemma posed by basing our laboratory studies of learning on tasks and procedures which are not "ecologically valid" is to undertake fine-grained, environmentally-based observational work (e.g., Brown, 1975; Brown & Campione, 1978; Butterfield, 1978; see also Cole & Scribner, 1975; Cole, Hood, & McDermott, 1978). Unfortunately this is far more easily called for than accomplished. One reason is that we lack an accumulated body of descriptive accounts of such behavior which would make comparison and contrast possible (though the work of Charlesworth, 1978, is promising in this regard), and which would thereby allow us to test our conceptual notions about what distinguishes one kind of task, performance, or behavior from another. As a result we cannot yet answer the question of whether a laboratory task is "representative" of real-life tasks, because (a) we do not have any systematic knowledge of what real-life tasks exist "out there," and (b) we do not know what demands, cognitive or otherwise, such tasks make on individuals and therefore have no conceptual scheme for gauging degree of "representativeness." In addition, while we appreciate that everyday tasks, by their nature, are embedded in an on-going "stream of behavior" we are not yet able to reliably predict in what way(s) the environment actually impinges on the person or on the task itself. Thus while observational researchers typically include setting and participant effects in their studies they have yet, in this area of research at least, to systematically analyze these effects. As in laboratory settings they remain ill-understood confounding variables.

The present paper provides descriptive material of everyday "memory" tasks collected from observations of public school classrooms for trainable mentally retarded (TMR) students[1]. In addition it proposes a multi-dimensional classificational scheme for understanding these tasks which includes aspects of the task and of the task environment as well. We believe that school-related tasks *cannot* be defined or categorized through the exclusive use of unitary "objective" criteria; and that any classificatory scheme ultimately has value only insofar as it mirrors how the culture structures, and the student perceives, such tasks.

*The research of the second author was partially supported by an NICHD Post-doctoral Traineeship. We would like to thank Ronald Gallimore, Kazuo Nihira, and Ray Perez for their helpful comments.

[1]For the purposes of this discussion we consider "everyday" tasks to include both school- (or curriculum-) related tasks and "daily living" tasks. In reality there may be some overlap between laboratory and school-related tasks which is not shared with living tasks in that many laboratory tasks are often derived from or considered parallel to formal schooling experiences (see Cole, Sharp, & Lave, 1976). However, the curriculum in these TMR classrooms tends to emphasize pragmatic, everyday skills, and a large percentage of classroom time is undirected time which seems to allow the students to engage in self-initiated activities. In this sense the classroom and the tasks found there share characteristics with "everyday" settings and tasks.

0160-3361/80/1-1 $1.00 © LCHC

Design and Procedure

Our research was conducted in three age-graded classrooms of a special day school for trainable mentally retarded learners. The students ranged in age from 6 to 14.9 years and had IQ scores spanning 25 to 55. Most of the children and adolescents were from middle-class families and lived at home.

We used naturalistic observation techniques—the observation and recording of verbal and nonverbal behaviors in their natural settings (Bogdan & Taylor, 1975)—carried out over a three-year period. During that time the researchers were able to develop an extensive data base and to become familiar with their subjects so that their presence was no longer obtrusive (Pelto, 1970).

Our initial guiding strategy in these observations was to focus on a wide variety of everyday tasks. We used two criteria for regarding a behavioral episode as a task: (1) the student must recognize a behavioral demand (whether self- or other-imposed) which we could also recognize through the student's subsequent verbalizations and/or other behaviors; and (2) we must be able to define two separately identifiable states which are functionally linked to each other through the student's behavior. The actual functional behavior we define as performance. These definitions are purposely broad so as not to ignore any situation properly considered a task for the child or any behavior regarded in some sense as performative.

Observers spent hundreds of hours in the classrooms observing all major periods of classroom activity such as reading lessons, workshop simulation activities, recess, lunchtime, music periods, etc. They maintained field notes of the behaviors they observed, noting with care any aspect of the situation, the participants, their actions, the task, or any materials in use which could conceivably have been noticed and hence evaluated and/or acted upon by the student learner.

From the total record of everyday tasks which we are now in the process of assembling and ordering, we selected the representative sub-sample of "memory task" items listed in Table 1. We decided a task had a memory component if (1) a teacher (or other adult) keyed a student's task-related performance with some verbal instruction to remember or with some reference to a past skill, project, event, etc. (or the student him/herself did so); and/or (2) we were able to identify the skill or experience undertaken by the student as something already encountered by him/her in the past—that is, we used our long-term involvement with the knowledge of these students and their learning environment to build up a "file" of items of which students should have memory.

TABLE 1
Memory Tasks at a School for TMR Children

1. In a typical "Going-on-a-trip" game teacher asked W. to repeat the five items other children before him have mentioned.
2. C. wrote his name on his completed art project when not specifically told to do so.
3. The teacher asked M. to say his telephone number.
4. In a verbal response game the teacher varied the order and combination of "address category" questions—e.g., "What is your street?" "What is your state?"
5. B. had to repeatedly respond in the refrain section of the song "This Land Is Your Land"...(see Item #6).
6. (From Item #5) with the lyrics: "This land is your land; this land is my land...(etc.)."
7. The speech teacher told M.: "When you see D..." (see Item #8).
8. (From Item #7) "...tell D. to come see me" (continued in Item #9).
9. (Continued from Item #8) When M. saw D... (see Item #10).
10. (From Item #9) M. said: "D., go see Mrs. G. [the speech teacher]."
11. The teacher asked the class if they recalled a previous lesson.
12. M. had to set up the school phonograph so that he and others could dance to music during free time.
13. Class had to recite the Pledge of Allegiance.
14. Teacher asked K. if she remembered the names of her siblings who would be attending "Sibling Day" at the school.
15. Teacher asked J. to remember a list of food items which J. had "written" on a sheet of paper (child can neither read nor write) when teacher was absent.
16. J. passed through cafeteria line getting all items (food, utensils, napkins) he needed for lunch.
17. During daily graduation rehearsal S. had to attend to the musical cue which would indicate it was time for her performance, and...(see Item #18).
18. (From Item #17) S. had to remember the actual dance steps which she was to do during the song "Did You Ever See a Lassie?"
19. F. wished to neatly tear a page from a magazine as he had been shown the day before.
20. F. tried to find recipes in a magazine for a self-generated project.
21. At the end of the school day M. had to remember to do his weekly assigned classroom "chores," which included...(see Item #22).
22. (From Item #21) inverting the chairs and placing them seat-side down on the tables.
23. Teacher told L. to put classroom objects away where they belong.
24. N. decided to put a classroom object which was out of place on the floor onto an empty shelf.
25. In the math bingo game M. was designated the "caller" and had to remember the name of a number and geometric shape.
26. In the math bingo game M. (and the other players) had to remember how to both play and win.
27. In the math bingo game M. had to remember to turn over the numbered cards after announcing them (as the teacher had shown him so that he wouldn't get confused).
28. In doing her math workbook problems L. had to remember the meaning and use of "+" and "−" symbols.
29. In doing her math workbook problems L. had to remember the steps necessary to solve the following math problem:
$$N(\phi \Upsilon \phi) - N(\Omega) = \_\_\_\_\_.$$
30. The teacher told D. to use her "arithmestick" (a plastic, abacus-like counting device) to solve the addition problem.
31. M. began to count on his fingers when he saw the math problem on his worksheet (see Item #32).
32. (From Item #31) in counting on his fingers M. had to remember how to use his fingers to count and manipulate the numbers from the math problem before him.
33. During the morning calendar exercise J. had to identify the numbers which the teacher had removed from the calendar while J. followed instructions and kept his eyes closed.
34. W. had to identify functional words (e.g., "poison," "stop," "danger") from flashcards presented by the teacher.

Insofar as possible, our description of task items in Table 1 reflects how the task was originally presented to the student. Although our methods do not permit us to know whether the student defined the task in the same way we did, we can at least make some judgment as to the task with which the student was initially confronted. Some of these tasks are ones which teachers would define as "memory development" tasks, while others seem to involve memory as only one component of a larger task. Some items are highly routinized and are probably overlearned by the students; others involve novelty in some way. Our operating assumption is that all memory tasks, no matter how different from each other or how intractable to a standard universal definition, bear a "family resemblance" to one another which makes them amenable to conceptual analysis.

Dimensions of Everyday Tasks

We have examined the variety of tasks listed in Table 1 as well as the classroom situations in which these specific tasks were actually embedded. Our intention has been to develop as many dimensions of contrast as possible to describe these tasks. By doing so we hope to discover new ways of thinking about everyday tasks, to isolate features of tasks or task-environments which elicit learning success or failure by students, and to provide direction in a future probe of those dimensions which actually are salient for student learners. To date we have identified five dimensions which describe our data. These are listed in Table 2.

TABLE 2
Dimensions of Everyday Memory Tasks

a. Problem presenting environment:
- a_1—teacher or other familiar adult-structured problematic situation
- a_2—child-structured problematic situation

b. Task demands calling for memory of or awareness of:
- b_1—unconnected words, numbers, symbols
- b_2—semantically connected words (e.g., messages, song lyrics)
- b_3—location
- b_4—an indication that some action, verbal or nonverbal, is required
- b_5—events to be remembered as whole
- b_6—number of steps sequentially arranged

c. Frequency of task occurrence:
- c_1—daily occurrence
- c_2—frequent occurrence but not daily
- c_3—irregular occurrence

d. Predictability of response set:
- d_1—type of response and task content are always the same
- d_2—type of response and task content always vary
- d_3—type of response is the same; task content varies

e. Feedback system:
- e_1— teacher and child come to "agreement" as to what counts as correct or acceptable response
- e_2— child is left on his/her own to decide what "counts" as acceptable response
- e_3— child is "informed" by the situation that his/her response is acceptable or unacceptable

Dimension *a* represents the way in which tasks come into the learner's awareness. Tasks are either structured by the teacher or another familiar adult (a_1) as in having the learner state her address (Item #4 of Table 1) and telephone number (Item #3), or by the individual students (a_2) as in writing his name on a finished art project (Item #2) and in recalling the steps to operate the record player (Item #12). With some of these student-structured tasks, although the teacher may be the one to initiate the action, as when she takes the class to the cafeteria line (Item #16) or assigns them a number of math workbook problems (Items #28 and 29), it is then left to the individual students to proceed with the task demands on their own.

Dimensions *b* through *d* document aspects of the task itself. Category *b* refers to the kinds of memory demands which this school culture places on its students. Unconnected words, numbers, and symbols (b_1) are stressed to be remembered as a child's name (Item #2), a telephone number (Item #3), a list of foods (Item #15), or when confronting the + and − symbols in math workbook problems (Item #28). The students are also called upon to remember semantically connected words (b_2) as in delivering a verbal message (Item #8) or reciting the Pledge of Allegiance (Item #13); places where objects are located (b_3—e.g., Items #20 and 23); an indication that some action, verbal or nonverbal, is required (b_4—e.g., Items #5 and 17); and certain events which are meant to be remembered as a whole (b_5—e.g., Item #11). Finally, the students are called upon to remember how to do something which involves a finite number of steps sequentially arranged (b_6), as the steps required in setting up a phonograph (Item #12).

We have also been able to meaningfully group tasks on the basis of the frequency of their occurrence in the classroom. Some tasks (c_1) occur daily such as the Pledge of Allegiance (Item #13) or on a daily basis for a short period of time while the class is engaged in a particular project such as the graduation rehearsals (Items #17 and 18). Other tasks, such as the "Going-on-a-Trip" game (Item #1), are frequent but not daily occurrences (c_2). Finally, a third group of tasks (c_3) occurs very irregularly, and are therefore not as likely to have a patterned response set available to the student (Items #15 and 20).[2]

Dimension *d* groups tasks on the basis of how predictable the response is to the student. We distinguish tasks in which (1) both the type of response and its actual content are always the same (d_1) such as the Pledge of Allegiance (Item #13); (2) both the type of response and the actual content vary with each new problem-presenting instance (d_2) such as the placement of an object which was found on the floor onto an available empty shelf (Item #24); and (3) the type of response is typical but the actual content varies (d_3) such as a student

[2]Additionally we feel that the difference between a learner's first encounter with a task and subsequent encounters may prove to be important. Our intention is to examine this distinction more systematically when we look at actual performance.

having to do a daily classroom chore while the chore itself changes from week to week (Item #21).

Category *e* reflects the type of feedback (if any) the student receives on his/her task performance. We have found that feedback to the student comes about in one of three ways: (1) the teacher and student come to an agreement as to what counts as an acceptable response (e_1); (2) the student is left to him/herself to judge whether the demands of the memory tasks have been fulfilled (e_2); and finally (3) something intrinsic to the task provides feedback, such as a phonograph which will not work if the steps necessary to start it were not followed (e_3).

RESULTS AND DISCUSSION

Each of the 34 memory items in Table 1 were scored by two independent raters according to the dimensions of contrast listed in Table 2. Overall interrater reliability was .96. Agreement indices for each dimension are as follows: .97 for dimension *a*; .97 for *b*; 1.00 for *c*; .94 for *d*; and 1.00 for *e*. A list of these dimensional sets is given in Table 3.

TABLE 3
Dimensional Sets for 34 Memory Tasks at a School for TMR Learners

1.	$a_1\ b_1\ c_2\ d_3\ e_1$	18.	$a_1\ b_6\ c_1\ d_1\ e_1$
2.	$a_2\ b_1\ c_1\ d_1\ e_2$	19.	$a_2\ b_6\ c_3\ d_1\ e_3$
3.	$a_1\ b_1\ c_2\ d_1\ e_1$	20.	$a_2\ b_3\ c_3\ d_3\ e_3$
4.	$a_1\ b_1\ c_2\ d_3\ e_1$	21.	$a_2\ b_4\ c_1\ d_3\ e_1$
5.	$a_1\ b_4\ c_2\ d_1\ e_2$	22.	$a_2\ b_6\ c_2\ d_1\ e_1$
6.	$a_1\ b_2\ c_2\ d_1\ e_2$	23.	$a_1\ b_3\ c_2\ d_3\ e_2$
7.	$a_1\ b_4\ c_3\ d_2\ e_2$	24.	$a_2\ b_3\ c_2\ d_3\ e_2$
8.	$a_1\ b_2\ c_3\ d_2\ e_2$	25.	$a_1\ b_1\ c_2\ d_3\ e_1$
9.	$a_2\ b_4\ c_3\ d_2\ e_2$	26.	$a_2\ b_6\ c_2\ d_1\ e_1$
10.	$a_2\ b_2\ c_3\ d_2\ e_2$	27.	$a_2\ b_6\ c_2\ d_1\ e_1$
11.	$a_1\ b_5\ c_3\ d_2\ e_2$	28.	$a_2\ b_1\ c_2\ d_1\ e_1$
12.	$a_2\ b_6\ c_2\ d_1\ e_3$	29.	$a_2\ b_6\ c_2\ d_3\ e_1$
13.	$a_1\ b_2\ c_1\ d_1\ e_2$	30.	$a_1\ b_6\ c_2\ d_3\ e_1$
14.	$a_1\ b_1\ c_3\ d_3\ e_2$	31.	$a_2\ b_4\ c_2\ d_1\ e_2$
15.	$a_1\ b_1\ c_3\ d_2\ e_2$	32.	$a_2\ b_6\ c_2\ d_3\ e_1$
16.	$a_2\ b_6\ c_1\ d_3\ e_1$	33.	$a_1\ b_1\ c_1\ d_3\ e_1$
17.	$a_1\ b_4\ c_1\ d_1\ e_1$	34.	$a_1\ b_1\ c_2\ d_3\ e_1$

The 34 items selected for analysis here were intended to be representative of the range of tasks and task environments in these classrooms. Although we probably have not exhausted the variability to be found in tasks and task environments we feel that these do offer a useful guide to understanding everyday tasks. In addition, even though it was not our specific intention to examine the levels of demand placed on students in TMR classrooms we can use these 34 dimensional sets as basic data for understanding how memory demands are actually structured there. In this context, a number of noteworthy findings emerged.

First, the most obvious finding was the high degree of routinization in this setting for teacher and student alike. Of the 34 tasks, 25 (.74) have values of either c_1 or c_2 and d_1 or d_3 indicating that the vast majority of tasks occur daily or "frequently," and require the same type of response (from the teacher's point of view) while the actual content of the response may vary. In contrast only 6 of the 34 items appear to be truly "unfamiliar"—occurring infrequently (c_3) and requiring a novel response (d_2).

The second interesting finding was that three values of dimension *b* account for an extremely large share of the variability, 24 of the 34 items (.71). In other words, the greatest memory demand (in terms of frequency of occurrence) involves the recollection of words, numbers, and symbols remembered individually (b_1) or in some longer, semantically meaningful sequence (b_2), or of procedures for doing things (b_6). Clustering b_1, b_2, and b_6 values with values of other dimensions provides limited but additionally suggestive evidence bearing on the nature of this task environment. Of 14 b_1/b_2 tasks, 11 (.79) are a_1 tasks and 10 (.71) are c_1 or c_2 tasks as well. That is, the majority of verbal memory demands are teacher-initiated and occur on a daily or frequent basis. Much the same pattern emerges with b_6 tasks. While b_6 tasks tend to be student-initiated (8 of 10, or .80) they also tend to be daily or frequent in occurrence (9 of 10, or .90). In addition all b_6 tasks (10 of 10) are either d_1 or d_3 tasks (that is, the type of response is always the same though the actual content of the response sometimes varies). Combining all 24 b_1, b_2, and b_6 tasks we see that 19 of them (.79) have values of either c_1 or c_2 and d_1 or d_3, further highlighting a very "familiar" environment.

Third, an unexpected outcome of our observations of the task environment was an inability to consistently divide task responses into categories of "right" or "wrong." Of greater salience within the classroom setting, was the question of whether a response was acceptable and who decided that it was so. Whether a response was "actually" correct or incorrect (such as the answer to a mathematics problem) seemed to be less important to the teacher than the fact that a response was called for and given.

In making a distinction between correct responses and acceptable ones it is interesting to note that our sample of 34 items contains no example of "acceptability" as solely determined by the teacher. When it is not situation-dependent (dimension e_3) or student-dependent (e_2), acceptability comes about through an interactive process between teacher and student (e_1). We have come to think of this process as a dance-like "pas-de-deux" (see also Cicourel, 1974; Mehan, 1979). In this process the teacher typically makes a demand, the student further adjusts the demand, the teacher responds, and so forth until the two reach an agreement on what is to count as acceptable performance. The final task for the student is therefore two or more steps removed from the original demand.

Though we are still documenting the frequency with which this pas-de-deux occurs, initial indications are that it is relatively high (17 of the 34 tasks in Table 1 have an e_1 value). If this does indeed prove to be the case we feel justified in arguing that the seemingly poor classroom performance (from an "objective" point of view) of TMR learners even on highly routinized tasks is partially a result of the greater value attached to "acceptableness" over "correctness." For example, neither teacher nor student seems to care that the Pledge of Allegiance be learned "by heart." Apparently each is mainly concerned that an acceptable level of sound be produced which, in its modulation and with a few "correct," audibly distinct words here and there, follows the recording of the Pledge which is playing in the background. It is fair to say, we feel, that these TMR students find few classroom situations in which they are expected to remember "correctly." The implications of this observational finding for various kinds of memory testing with their emphases on correct answers should be obvious. It should be equally clear that any cognitive analysis, even if methodologically feasible with everyday tasks, may be directed to the wrong task—one with which neither the student nor the teacher is concerned.

Finally, our micro-level approach to everyday memory demands reveals additional complexity with the cognitive analysis of everyday tasks. We believe that although many everyday demands involve a kind of "set" problem for the student, many others involve attention to shifting values of the dimensions which define the task and its environment and which become part of the cognitive "load" on the student. Items #5-6 of Table 1, for example, define two components of what might otherwise be mistakenly regarded as one single task. This is true for Items #17-18 as well. While values for the *a, c, d,* and *e* dimensions remain constant, the *b* values change from memory for a "cue" (b_4) to (1) memory for semantically connected words (b_2) in Item #6 and (2) memory for a number of sequentially arranged steps (b_6) in Item #18.

A more complex "translation" process may be found in Items #7 through 10. In this case only 3 dimensions (*c, d,* and *e*) remain constant; the actual memory demand alternates between a cue (b_4) and a sequence of words (b_2) and the initiator of the action changes from the teacher (a_1 in Items #7 and 8) to the student (a_2 in Items #9 and 10). We believe that these sometimes subtle shifts in task definition are relatively common in everyday settings and produce meta-level cognitive demands—e.g., the student's "recognition" that s/he is the initiator of an action when originally it was the teacher—which transcend the "simple" demands of the task itself.

CONCLUSION

The goal of our analysis has been to develop a scheme for categorizing everyday memory tasks and discovering any underlying patterns in these tasks. We also feel that our data directly bear on the more general issue of the relationship between "everyday" and "laboratory" tasks.

FIGURE 1
Different Approaches to the Study of Tasks

PLACE WHERE STUDIED	ORIGINAL LOCATION OF TASK: LABORATORY SETTING	ORIGINAL LOCATION OF TASK: NON-LABORATORY SETTING
LABORATORY	A	C
"REAL WORLD"	B	D

To date, we believe researchers have been concerned with four basic types of studies of tasks. These are represented by the 4 cells in the 2 × 2 matrix in Figure 1. We compare the kinds of tasks—i.e., whether they were developed for use in the laboratory or in other settings—with where the tasks are actually studied—i.e., in the laboratory or in non-laboratory settings. In Cell A we place typical memory tasks such as serial recall, problem-solving tasks such as the Tower of Hanoi puzzle, and the like. These evolved in the laboratory and are used to study cognitive processes. In Cell D we place the kind of tasks studied here, and represented by Table 1. These are tasks created by teachers and students for each other and themselves. The tasks in Cell B are adaptations of laboratory-based tasks for use in other settings. This strategy is common in cross-cultural studies, and usually involves manipulating the content and materials of the task so that they are "culturally appropriate" (for one example of such an approach see Cole, Gay, Glick, & Sharp, 1971). Often such studies explicitly examine performance on these tasks over time in order to make comparisons across settings. The strategy for studying Cell C tasks involves assessing how certain tasks are formatted and contextualized in a given "natural" setting and then modeling one's experiments on this pattern (see, e.g., Lave, 1977).

Since our initial interest was the interrelationship of Type A and D tasks we analyzed three typical Type A memory tasks in terms of the conceptual scheme developed here. Interestingly the three tasks—free recall, serial recall, and paired-associate learning—all have similar dimensional sets. Thus we found that all three laboratory tasks emphasize recollection of isolated words (b_1) and initially occur quite irregularly (c_3). Two of the tasks (serial recall and paired-associate learning) have the same response type (i.e., the verbalization of a single word) with variable content (i.e., different words on each of the trials) (d_3). The third task (free recall) requires that the type of response and task content are always the same (d_1). In terms of the memory demand, the novelty of the task, and the type of response called for, these laboratory tasks have counterparts in the classroom environment. Indeed we found one task (Item #4 in Table 3) which has a pattern identical to these laboratory tasks—$b_1c_3d_3$.

Type A and D tasks differ in the *a* and *e* dimensions. All Type A tasks involve an outsider who is responsible for presenting the task, and we created a new value (a_3) to represent this dimension. We do not feel, however, that this difference is a significant one as students in classrooms frequently encounter new persons—substitute teachers or aides, new volunteers, administrators—who assign them tasks. Dimension *e* represents a more important difference. Laboratory tasks require a single person to make a unilateral judgment on the correctness of the response. Not only is the emphasis on correctness, but also there is no feedback to or negotiation with the subject on the satisfactoriness of the latter's response. As we have seen, this is quite unlike the classroom setting. As a result no classroom task we observed has a structure identical with these three laboratory memory tasks, and none of these laboratory tasks taps the wide range of different task types (28 in even our limited sample) actually confronted by student learners.

Although studies of tasks as represented by Cells A and D are not without their problems, they seem to be "pure" forms of a method of analysis and type of task. The approaches represented by Cells B and C on the other hand seem to be "compromises" desirable on methodological grounds—Cell B studies achieve the generalizability that conventional laboratory tasks are said to lack and Cell C studies provide more control over and reliability of task performance in which non-laboratory tasks are said to be wanting. Thus both Type B and C tasks attempt to bridge the "gap" between laboratory and non-laboratory settings. Our limited goal in this paper has been to suggest one classificational framework within which all tasks can be grouped. We feel that further research needs to be done to fully understand "tasks," regardless of the setting in which they were originally developed or actually studied. This is a critical first step in our eventual understanding of the demands we place on youngsters and the nature of their task-related competencies.

REFERENCES

Bogdan, R., & Taylor, S.J. *Introduction to qualitative research methods.* New York: Wiley, 1975.

Brown, A.L. The development of memory: Knowing, knowing about knowing, and knowing how to know. In H.W. Reese (Ed.), *Advances in child development and behavior.* New York: Academic Press, 1975.

Brown, A.L., & Campione, J.C. Permissible inference from the outcome of training studies in cognitive development research. *Quarterly Newsletter for the Institute for Comparative Human Development,* 1978, *2*(3), 46-53.

Butterfield, E.C. On studying cognitive development. In G.P. Sackett (Ed.), *Observing behavior: Theory and applications in mental retardation* (Vol. 1). Baltimore: University Park Press, 1978.

Charlesworth, W.R. Ethology: Its relevance for observational studies of human adaptation. In G.P. Sackett (Ed.), *Observing behavior: Theory and applications in mental retardation* (Vol. 1). Baltimore: University Park Press, 1978.

Cicourel, A.V. Some basic theoretical issues in the assessment of the child's performance in testing and classroom settings. In A.V. Cicourel, K.H. Jennings, S.H.M. Jennings, K.C.W. Leiter, R. MacKay, H. Mehan, & D.R. Roth (Eds.), *Language use and school performance.* New York: Academic Press, 1974.

Cole, M., Gay, J., Glick, J.A., & Sharp, D.W. *The cultural context of learning and thinking.* New York: Basic Books, 1971.

Cole, M., Hood, L., & McDermott, R.P. Concepts of ecological validity: Their differing implications for comparative cognitive research. *The Quarterly Newsletter of the Institute for Comparative Human Development,* 1978, *2*(2), 34–37.

Cole, M., & Scribner, S. Theorizing about socialization in cognition. *Ethos,* 1975, *3*, 249–268.

Cole, M., Sharp, D.W., & Lave, C. The cognitive consequences of education: Some empirical evidence and theoretical misgivings. *The Urban Review,* 1976, *9*(4), 218–233.

Lave, J. Tailor-made experiments and evaluating the intellectual consequences of apprenticeship training. *The Quarterly Newsletter of the Institute for Comparative Human Development,* 1977, *1*(2), 1–3.

Mehan, H. *Learning lessons: Social organization in the classroom.* Cambridge, Mass.: Harvard University Press, 1979.

Pelto, P.J. *Anthropological research: The structure of inquiry.* New York: Harper & Row, 1970.

Toward a Standardized and Objective Methodology for Studying Children's Distributive Justice Reasoning

Robert D. Enright
Department of Educational Psychology
University of Wisconsin

Lesley A. Manheim
Department of Educational Psychology
University of Wisconsin

Christina C. Franklin
Department of Psychology
Tulane University

Those who study cognitive development with a structuralist model have traditionally relied on Piaget's *mêthode clinique* as the primary method for gathering data. In the social cognitive realm, this bias has continued. For instance, Kohlberg (1969) assesses moral reasoning through the presentation of up to nine dilemmas followed by probe questions to assess children's moral structural development. Selman (1976) assesses similarly in interpersonal conceptions or children's understanding of friendship, while Damon (1977) does the same in assessing children's distributive justice development.

All of the above researchers, while carving out different domains of social reasoning, have relied exclusively on dilemmas followed by open-ended questions which become idiosyncratic for each child depending on his or her answers to the original set of open-ended questions. Certainly this approach has advantages in the pilot phases of research. The open-ended format allows the child to "construct" his or her own conceptions of the social world rather than forcing the child to react to experimenter-created statements. This allows the experimenter to obtain a complete picture of a stage sequence which is missing with forced-choice scales. For instance, if the theorist hypothesizes in domain X the stage sequence of A,B,C, and D, he or she may never discover E with a forced-choice scale that includes items for stages A-D only. Or, if the hypothesized sequence is A,B,C,D, there may be an A/B transitional stage which is missed with the forced-choice procedure.

While there are advantages, then, to the traditional assessment approach in cognitive and social cognitive development, there are also some glaring weaknesses if the open-ended approach as the only viable test-construction procedure continues beyond the pilot phase. The weaknesses are as follows: (a) Since the interview is not standardized, each child takes a somewhat different "test," which contributes to measurement error. (b) Replication in a strict sense is impossible since any two studies are likely to have different experimenters with different interview styles. (c) Structural scoring is difficult as Rest (1975) notes since many children's responses are difficult to categorize. (d) Clinical interviews are often lengthy, contributing to fatigue effects especially in young children. (e) Transcriptions of taperecorded interviews, training of interviewers, and training of scorers are both time consuming and expensive. (f) Finally, because the open-ended, clinical method relies on the child's verbal production, the interview may be confounded by verbal ability.

Given the formidable problems to be overcome, is it possible to build an instrument for assessing children's social cognitive development that is more scientifically sound than the clinical method? While Rest (1974) has successfully done so in the moral domain, his was scaled for adolescents and adults. Children create other kinds of assessment problems such as attention span, primacy and/or recency effects, and reading level which excludes the possibility of a paper-and-pencil standardization procedure.

The instrument described below was developed to assess 4- to 10- or 11-year-old children's distributive justice reasoning about the fair allocation of goods in a group or society. Damon (1977) has described via the clinical method the following stage sequence:

- 0-A: The child believes that whoever wants the most money or goods should have it.
- 0-B: The child bases distributive decisions on external characteristics. The oldest one, for example, should get more than the others.
- 1-A: The child believes everyone should receive the same amount regardless of other characteristics.
- 1-B: The child bases distributive decisions on behavioral reciprocity. For example, the child believes that those who work harder or do more than the others should get more.
- 2-A: The child bases distributive decisions on psychological reciprocity. That is, the child believes that those who are most in need should receive more than the others.
- 2-B: Compromise is the key to distributive decisions. Here, for example, the child thinks the neediest and most hard-working children should get the most. It means a little less for either the needy or hard-working children than is the case in a 2-A or 1-B decision, but it shows greater cognitive complexity in decision making.

This sequence was assessed with a standardized, forced-choice procedure described below.

The Distributive Justice Scale

The Distributive Justice Scale (DJS) is based on Damon's (1977) recommendation of a comprehensive assessment strategy in distributive justice. In his interview, Damon suggests that once a child has made a distributive justice decision, the experimenter should pre-

0160-3361/80/1-7 $1.00 © LCHC

sent other alternative decisions to see if the child continues to hold the original belief when faced with new possibilities. Damon, in effect, is presenting a paired-item test, the pair at any time being the child's current distributive justice belief and the interviewer's probe in which a different alternative is suggested. Since each interview is different, however, not all children get the exact same alternatives presented to them, nor in the same order.

For the DJS, pictures like that in Figure 1 were drawn to represent the different stages of distributive justice for a given dilemma. Stage 2-B was excluded from the DJS since it could not be represented pictorially in such a clear-cut way as the other stages which have only one child getting the most or all getting the same. Figure 1 refers to the following dilemma patterned after Damon (1977):

> These boys and girls all go to the same camp. This is Betty, she's the oldest one at the camp; this is Jennifer whose family does not have much money; this is David who made the most paintings; and this is Matthew. One morning they all thought it would be a good idea if they got out their paints and painted pictures of what they saw around the camp. When they were done, Betty made 2 paintings, Jennifer made 2, David made 4 paintings, and Matthew made 2. After they did this, they asked the person who runs the camp if he would like some of the paintings. He bought all the paintings and gave some nickels to the children. The children, then, had to decide how to split up the nickels. What is the best way to split up the nickels?

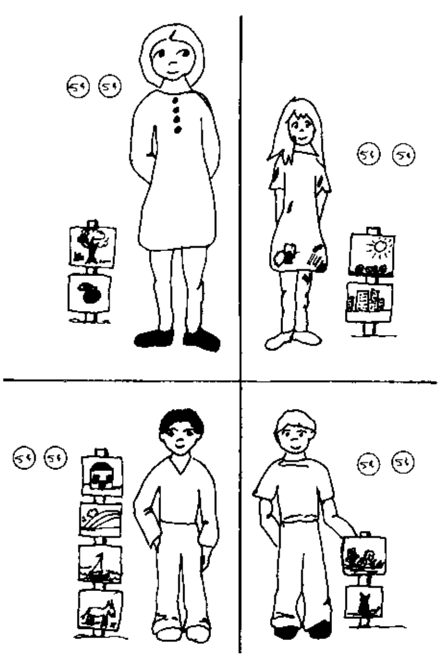

Figure 1

Figure 1 represents stage 1-A since all children get the same amount despite David (lower left) doing more pictures than all the others. Each picture has a standardized statement with a decision and reason for the decision read to the child as the picture is presented. For the stage 1-A example the experimenter says, "In this picture, all the children get the same number of nickels so they won't fight about who gets more." As another example, the stage 0-A picture shows Matthew with five nickels and the other children with one each accompanied by the statement, "In this picture, Matthew gets the most nickels because he wanted those nickels more than anything else in the world."

Another dilemma is also presented to assure generality beyond one dilemma. The other dilemma has the males at the top and a female in the lower left doing the most. This reversal of roles was done to control for subjects' possible sex role biases influencing their distributive justice responses. As can be seen in Figure 1, the drawings were done to be as racially non-specific as possible.

For each dilemma, a paired-comparisons format is used in which each picture and statement is paired with each other picture and statement. For any given pair, the experimenter places the two pictures in front of the child, says the two statements corresponding to the respective pictures, and asks, "Which picture ends the story the best?" For each pair, a random selection was used to decide which picture was presented first. The order of presentation for the 10 pairs was also decided via random selection. These random orderings constitute the final orderings for the test rather than selecting a new ordering for each child.

Once the child is shown the 10 pairings for any given dilemma, he or she is presented with 3 pairs repeated from the original 10 to check for consistency. These repeated pairings are presented in reverse order of the original pairings to control for primacy or recency effects which may confound the consistency check. For instance, if a recency effect were operating with a particular child with a particular pairing and that same pair was again presented, the child would again choose the second picture presented, thus appearing to be consistent. If the child fails 4 of 6 repetitions, he or she is eliminated from the sample. It should be noted that out of 258 children tested in our research programs to date with the DJS, only 6 have had to be removed because of lack of consistency. Total administration time for both dilemmas is approximately 12-15 minutes per child.

The DJS is scored by selecting the child's preferred stage via the picture comparisons for each dilemma. For example, if the child chose 1-B over all other stages, the child would be assigned that stage for the dilemma. A complication arises as in any paired-comparisons test if a triangular relationship exists such as 2A > 1-B, 1B > 1-A, 1A > 2A. In such a case, the lowest stage in the triangle is chosen since the child is, at best, in transition and the only stage on which the child is consolidated is probably the lowest of the three. The final score is obtained by first converting the preferred stage of each dilemma to a numerical value (e.g., 0-A = 0.0, 0-B = 0.5, 1-A = 1.0, and so forth). A mean of the two dilemmas represents the total score. Such a score assumes that developmental levels are continuous rather than discontinuous. For example, a value of 1.75 is interpreted as the child being between 1-B (1.5) and 2-A (2.0), showing evidence for both kinds of reasoning (the reader is referred to Flavell, 1971 for further discussion on the continuous vs. discontinuous controversy).

Two studies which attempted to validate the DJS are described below. A more detailed account of these can be found in Enright, Franklin, and Manheim (1979).

A Cross-Cultural Comparison of Distributive Justice Development

In Study 1, 66 middle-class children from the midwestern United States participated. There were 22 children, half male and half female, from first, third, and fifth grades. All were given the DJS and the Peabody Picture Vocabulary Test to assess the degree to which the DJS overlapped with verbal ability. The reliability of the two DJS dilemmas via the Spearman-Brown formula was adequate, being in the high .60s. There was also a strong linear trend found for the DJS, $F(1, 63) = 24.23$, $p < .001$. No sex effects were apparent. The means were as follows: first grade = .98, third grade = 1.46, and fifth grade = 1.65. The DJS and PPVT correlated .25 and even lower when age was partialled out. This first study showed age trends that match Damon's (1977) norms, and good convergent-discriminant validity suggesting little confound of verbal ability in the DJS.

For Study 2, 88 children from Kinshasa, Zaire, Africa took part. They comprised similar age groups as the American sample and all were tested in school. There were 29 6-year-olds, 29 18-year-olds, and 30 10-year-olds. There are both Belgian and African tribal influences in the sample. The children attend a school owned by Belgians but run by native Zairians. At the time of testing, the political structure was a dictatorship since the Zairian leader had proclaimed himself "President for Life." The economic structure was a blend of socialism (nationally owned businesses) and capitalism (privately owned businesses). There is a social class structure in Kinshasa and so we chose the middle class for comparison purposes with the American sample. The DJS was translated into Lingala, the native language of Kinshasa. The results replicated Study 1. The reliability was in the .70s and there was a strong, upward linear trend, $F(1, 85) = 15.74$, $p < .002$. The means were as follows: 6-year-olds = .93; 8-year-olds = 1.36; 10-year-olds = 1.57. The replication of findings is remarkable when it is realized how different were the political/economic structures in the two cultures. Also, the experimenters in either study were not aware of the results of the other study either before or during data collection, thus ruling out experimenter bias as a competing hypothesis.

RESEARCH IMPLICATIONS

It would have been difficult to do these cross-cultural comparative studies without a standardized measure such as the DJS. The amount of training needed for an open-ended approach would have been too expensive and time consuming. For instance, it would be very difficult for an American, who speaks little Lingala, to explain to a Zairian, who speaks some English, how to clinically probe when each child's interview is somewhat different from all other children's. Structural probing is an art that takes work and feedback from others. Thus, a Western vs. non-Western study that employs the clinical method is at a disadvantage due to the verbal production needed both to explain the test to the experimenter and to give the test to the children. It is interesting to note that in all of the structural, social-cognitive domains such as moral development, role-taking, interpersonal conceptions, and distributive justice, very few studies have appeared in the published literature which were done in a non-Western or non-industrialized country. The structuralist method, therefore, may be a handicap to generalization of results. The current studies were done to show that a structuralist model need not be equated with one and only one methodology.

Besides the methodological implications for the DJS, these studies suggest that distributive justice in these samples is parallel for the middle-class Western and non-Western children. Piagetians are known for their universality assumption and these studies do not contradict that assumption. It may be the case that young children regardless of culture go through the hypothesized distributive justice sequence. Such research into the early developments should be continued in other cultures as well, such as socialist cultures and lower-class cultures in capitalist countries, to see if the sequence continues to hold. The DJS can help answer these questions because of its standardized and easily translatable form. If the sequence does hold across cultures through middle childhood, the next important question to answer is when and how such convergence across cultures begins to diverge. We would not expect a continued stage convergence across cultures through adulthood if we can infer different distributive justice cognitions in adulthood from the radically different political and economic structures created across the world. Again, these questions might only be answered by a standardized, objective methodology.

REFERENCES

Damon, W. *The social world of the child.* San Francisco: Jossey-Bass, 1977.

Enright, R., Franklin, C., & Manheim, L. Children's distributive justice reasoning: A standardized and objective scale. Manuscript submitted for publication, 1979.

Flavell, J. Stage related properties of cognitive development. *Cognitive Psychology,* 1971, *2,* 421–453.
Kohlberg, L. Stage and sequence. In D. Goslin (Ed.), *Handbook of socialization theory and research.* Chicago: Rand McNally, 1969.
Selman, R. A structural approach to the study of developing interpersonal relationship concepts. In A. Pick (Ed.), *Xth annual Minnesota symposia on child psychology.* Minneapolis: University of Minnesota Press, 1976.
Rest, J. Manual for the Defining Issues Test. Unpublished manuscript, University of Minnesota, 1974.
Rest, J. New options in assessing moral judgment and criteria for evaluating validity. Paper presented at the meeting of the Society for Research in Child Development, Denver, April 1975.

Group Process and Learning in an Interacting Group

Noreen M. Webb
Graduate School of Education
University of California, Los Angeles

Research in the classroom has traditionally focused on methods of teaching, without analyzing the nature and impact of the classroom environment on a student's learning. Yet, students do not learn alone, as subjects in the laboratory usually do. Some part of the student's time is spent interacting with other students; what each student learns will be affected by other members of the class. The instructional outcome for the individual student, then, depends upon the characteristics of the class itself and on the student's experience within the class.

Social science researchers have begun to recognize that the context in which students learn will influence their motivation and learning (e.g., Coleman, 1966; Anderson & Walberg, 1974; Bronfenbrenner, 1976). But little research has examined *how* the context influences learning. In particular, little attention has been given to the effects of interaction with other students in a group setting on individual learning. The literature on learning in groups includes comparisons of performance in individual and group laboratory settings (e.g., Amaria, Biran, & Leith, 1969; Klausmeier, Wiersma, & Harris, 1963; Laughlin, McGlynn, Anderson, & Jacobson, 1968; Laughlin and Sweeney, 1977), comparisons of achievement under lecture and discussion methods (e.g., Dubin & Taveggia, 1968), and examination of achievement in tutorial situations (e.g., Devin-Sheehan, Feldman, & Allen, 1976; Gartner, Kohler, & Riessman, 1971). Most of this research is inconsistent. The only consistent finding is that tutors showed improved achievement as a consequence of "teaching" another student or group of students. None of these studies examined group process in any systematic way to explain differences in achievement among learning conditions.

Webb (1977) systematically examined group process to explain the results of a study comparing mathematics learning of eleventh-grade students in small groups of four to learning individually.[1] In the group condition, students were encouraged to help each other learn. Some groups had mixed ability. Each mixed-ability group consisted of one high-ability, one low-ability, and two medium-ability students. Compared to learning in uniform-ability groups, learning in mixed-ability groups was beneficial for high-ability and low-ability students, but typically was detrimental for medium-ability students.

Webb's analysis of group process showed that better performance was associated with active verbal participation in the group. In mixed-ability groups, high-ability and low-ability students interacted with each other; highs helped the lows. Unless medium-ability members aggressively asked for explanations or took part in the explaining, they were ignored.

The finding of the importance of active verbal participation is tantalizing. However, the relation of group process to individual achievement needs to be examined in depth to show how participation was beneficial. This paper analyzes in detail the group process in one typical mixed-ability group. Group process is related to achievement in terms of the specific components of the algorithm used to solve the mathematical problem.

Procedures in the Experiment

In the group analyzed here, one high-ability student, two medium-ability students, and one low-ability student were assembled to learn a mathematical task. All students were female and were not acquainted before the start of the study. The students were members of a pool of 181 students who had taken a battery of aptitude and achievement tests. A factor analysis using all scores in the test battery produced two orthogonal ability factors: scholastic ability and nonverbal spatial-analytic ability. High, medium, and low ability strata were defined using the scores on the two ability factors.

There were three phases of work on the task. During the first phase students worked individually on work booklets which contained instruction on component concepts and skills of the task. The work booklets consisted of several pages of text followed by exercises for the student to do. Students worked alone on the exercises and asked questions of the experimenter whenever they were confused. By the end of the time allotted for this phase, all students had completed several exercises correctly. During this phase they learned the components they would need in order to solve the complex problem in the next phase.

[1]See the Annotated Bibliographies section of this issue for an abstract of this report.

0160-3361/80/1-10 $1.00 © LCHC

During the second phase students worked together to help each other learn how to solve the complex problem. Students were instructed to ask questions of teammates and to explain how to solve the problem to any group member who was confused. Students were told not to divide the work, but to work as a four-person group. The group practiced solving several problems during this phase. For each problem there was a booklet of hints consisting of a step-by-step solution. The group used the booklets to settle arguments about the definitions of terms and to check the accuracy of its solutions to the problems. What group members said during group work was recorded on an audio recorder.

During the third phase students returned to their seats to take a test on the complex problem they had just learned. The test consisted of a complex problem similar to the ones they had solved during the second phase. One week later they solved another test problem similar to the original test problem. On the immediate and delayed tests, partial credit was given for each step of the problem performed correctly.

The Mathematical Task

Students were asked to calculate an algebraic expression for the n^{th} polygonal number. The n^{th} polygonal number is the total number of dots in an array of polygons, in which the outermost polygon has n dots in each side. A sketch of the array for the n^{th} *triangular* number is given in Figure 1. In the array the smallest triangle has two dots on each side. Each successively larger triangle has one more dot on each side. The n^{th} *triangular* number, then, is the total number of dots in an array of triangles in which the outermost triangle has n dots on each side. The number of dots in the array forms an arithmetic series: $1 + 2 + 3 + \ldots$. The n^{th} *triangular* number is the sum of the arithmetic series. Solving the problem involves describing the array of dots, specifying the arithmetic series, determining the n^{th} number in the series, and, with that information, calculating the sum of the series. The algorithm that students were taught to use to find the n^{th} *triangular* number is given in Table 1.

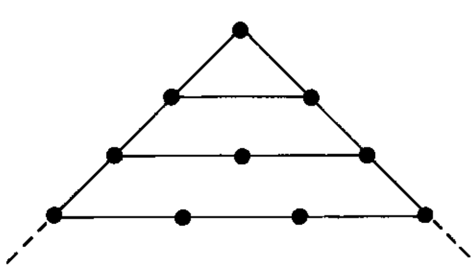

Figure 1. Array for the *nth triangular* number.

TABLE I
Algorithm for Calculating the n^{th} *Triangular* Number

Step I. Draw the array of triangles.
- (1) Draw a triangle.
- (2) Starting at one vertex, draw a line from that vertex to an adjacent vertex and continue the line past the adjacent vertex for a short distance.
- (3) Starting at the vertex in (2), draw a line from that vertex to the other adjacent vertex, and continue the line past the adjacent vertex for a short distance.
- (4) Connect the ends of the lines drawn in (2) and (3) with a straight line.

Step II. Mark the correct number of dots on each side.
- (1) On the innermost triangle mark a dot at each vertex.
- (2) On the second triangle, add dots such that each side has three dots on each side.

Step III. Decompose the array of dots into layers of dots.
- (1) The first layer is the dot on the vertex common to both triangles.
- (2) The second layer contains all dots in the innermost triangle except the dot in the first layer.
- (3) The third layer contains all dots in the second triangle except the dots in the first and second layers.

Step IV. Calculate the difference d between consecutive layer ($d = 1$).

Step V. Calculate the formula for the n^{th} term of the arithmetic series whose terms are the layers in III.
- (1) The general formula is $a + (n-1)d$, where a is number of dots in the first layer, and d is the difference calculated in step IV. No number is substituted for n. (The n^{th} term of the series is $1 + (n-1)1$ or n.)

Step VI. Calculate the formula for the sum of the arithmetic series.
- (1) The general formula is $\frac{(a + n^{th} \text{ term})}{2} \times n$ where a, n^{th} term, and n are defined in step V. (The sum of the series is $\frac{(1 + n)}{2} \times n$ or $\frac{n + n^2}{2}$.)

During the first phase, students learned how to draw the array, how to decompose the array into layers that form an arithmetic series, how to calculate the n^{th} term of any arithmetic series, and how to calculate the sum of any arithmetic series. Students learned these components as separate tasks. During the second phase, students had to put the components together to find the n^{th} *pentagonal* number, the n^{th} *hexagonal* number, and the n^{th} *octogonal* number. On the immediate test students were asked to find the n^{th} *triangular* number; on the delayed test they were asked to find the n^{th} *square* number.

Group Process Variables

The original coding system had fourteen categories of verbal interaction among group members: proposes idea, gives information, explains, criticizes, evaluates positively, detects/corrects omission, shows comprehension/insight, requests clarification/explanation, requests fill-in of memory, gives support, requests information, proposes strategy, negative interaction, other task-related interaction, non-task-related interaction. Transcripts of group sessions were coded using the fourteen-category system. The only category that related to achievement was "explains."

The coding system was modified to include general categories of group process that seemed to relate to achievement. The resulting categories were (1) gives explanation, (2) receives explanation, and (3) is ignored. "Gives explanation" included describing, elaborating, or clarifying something already said or done. Verbalizing a problem while solving it or repeating the solution to a problem without elaborating it or describing how to obtain it were not coded as explanations. "Receives explanation" was defined as being the target of the above behavior. "Is ignored" included all task-related questions that were not acknowledged by another group member.

Group Process and Achievement

Analysis of verbal interaction among members of the group shows that the high-ability student (HIGH) and one medium-ability student (MEDI) actively explained the components of the algorithm. The low-ability student (LOW) was usually the target of those explanations. The other medium-ability student (MED2) received fewer explanations. Participation was related to achievement; HIGH, MED1, and LOW showed excellent performance and MED2 performed less well. Closer examination of interaction reveals that this finding applies to specific components of the algorithm.

The first problem in the group phase was to determine the n^{th} *pentagonal* number. HIGH and MED1 assumed responsibility for solving the problem. Together they worked through the problem until they obtained the correct solution, while MED2 and LOW looked on. For most of this part of the session, HIGH and MED1 argued about whether the layers of dots in the array included dots in the previous layers (step III of the algorithm). HIGH eventually convinced MED1 (correctly) that the layers did not include dots in previous layers. No other errors were made during this episode of group work. Although LOW and MED2 asked questions of HIGH and MED1 (e.g., "What is that you're doing?", "How do you get that into the formula?" and "Is that the formula?"), HIGH and MED1 rarely acknowledged them. An acknowledgement, when it was made, was typically a curt, "Yeah."

Not until HIGH and MED1 obtained the correct solution did they offer any explanations to MED2 and LOW. As soon as HIGH and MED1 obtained the correct solution, they turned their attention to LOW. In response to LOW's question, "How do you get this [the correct answer]?", HIGH and MED1 correctly explained to her steps IV through VI of the algorithm. An excerpt of the protocol of this episode is given in Table II. LOW asked questions intermittently to ascertain whether she understood the explanations. These questions revealed incomplete or incorrect understanding of the algorithm for solving the problem. HIGH and MED1 responded to these questions with extensive explanations. HIGH and MED1 continued explaining until LOW indicated that she understood how to solve the problem. During this episode MED2 was silent, but seemed to be attending to the explanations given. At no time during work on the first problem did another group member ask MED2 whether she understood how to solve the problem.

TABLE II
Excerpt of Protocol of Interaction with LOW in Problem #1

Step V, VI	MED1:	Okay, this is the first number (points). This is the last number (points). We . . . just now found that last number by going *a* plus *n* minus 1 times *d*. *d* is the distance between. So the first number plus this number times the total number there are divided by 2. Okay, so that gives you . . . .
Step VI	HIGH:	If you put it 2*n* squared minus *n*, right. See, what we need is the first and last number. . . .
	LOW:	Okay, it's *a* plus . . . .
Step V, IV	MED1:	No, no . . . last number's going to be the first number you want to have. . . times the distance (*d*) between each number: 1, 5, 9. . . .
	LOW:	That's 4, then?
	MED1:	Right.
	LOW:	Isn't the last number we're going to have *n*?
Step V	HIGH:	No, the last number isn't *n*. *n* is the number of sides, and what we're finding here is the number in the whole figure when we have *n* dots on a side.
	LOW:	Okay.
Step V	MED1:	'Cause, like for 13, 13's gonna be the last number you substitute in for. So *n* is the number of different numbers in this series.
	LOW:	You'd have 13 minus 1.
	MED1:	No, you'd go 4 minus 1.
	LOW:	Oh, right.
	MED1:	So it's $n - 1$ times *d*, okay.
	LOW:	This is the last number, then, again.
	HIGH:	Right.
Step VI	MED1:	So you go to this first number, which is gonna be *a*, which is one, plus this thing times the number of terms there are divided by two.
	LOW:	Oh, right, Now I've got it.

For the second problem (the n^{th} *hexagonal* number), as for the first, HIGH and MED1 started to assume responsibility for solving it. After completing steps I, II, III in the algorithm, however, they turned the problem over to LOW: "Wait, do you understand how to do it? You write it." LOW made no errors when carrying out the steps of the algorithm but made several errors when reducing algebraic expressions. During work on the second problem, MED2 did offer comments (e.g., "So this formula will work for this one") but was not acknowledged.

Not until the start of the third problem (the n^{th} *octagonal* number) did the group act upon the instructions that they should help *all* group members. At this time, HIGH and MED1 instructed MED2 to work out the solution: "Do you understand it? Why don't you write it out this time?" Excerpts of the protocol during this episode are given in Table III. Although MED2 made errors, HIGH and MED1 exhibited less patience with her than during previous episodes with LOW. They tended to correct MED2 without offering explanations and to solve the problem for MED2 without allowing her to solve it.

TABLE III
Excerpts of Protocol of Interaction with MED2 in Problem #3

Step I		MED1:	No, there isn't one right here . . . . Wait, what are you doing? No, that's not it.
		MED2:	That's not it?
		MED1:	No it's not. You have to use two sides.
		HIGH:	What you just do is pick your two sides.
		MED1:	You only used *one* side.
		MED2:	Oh, that's what it was.
Step II, III		MED1:	Okay, here's your octogon (marks the dots in the array). So this is a point. The first number is this one (points). So the second one is 7 (points), the third one is 13 (points).
Step IV		HIGH:	So the number's six.
		.	
		.	
		.	
Step V		MED1:	*a* equals one.
		MED2:	So, it's one plus *n* − 1, six times . . . .
Step VI		MED1:	times *n* over 2.
		MED2:	Six *n* minus 6.
		MED1:	No, no, you forgot about these two.
		MED2:	Well, but you have to do those first.
		MED1:	No, that's wrong.
		LOW:	Plus 2.
	*	MED2:	Oh, it's *plus*. Well, how could you . . .?
		MED1:	You put the first two inside of this. Then it's 6*n* squared minus 4*n* over 2.
		HIGH:	Or 3*n* squared minus 2*n*.

The group's interaction with MED2 in Table III provides a striking contrast with the group's interaction with LOW in Table II. Whereas HIGH and MED1 had offered lengthy explanations to LOW, they offered few explanations to MED2 and usually cut her off during her calculations and questions. At least one group member presumed that lengthy explanations to MED2 were unnecessary because MED2 understood how to solve the problem: "Go ahead and do it. I'm sure she [MED2] understands it." Yet MED2 gave little evidence during group work that she understood how to solve the problem. Group members may have interpreted her infrequency of questions to mean that she understood the material. The group was more sensitive to questions than to other verbal behavior. Nearly all of the explanations were given in response to questions asked; a few were given in response to errors made. No other verbal behavior elicited explanations.

On the immediate test HIGH and MED1 made no errors. LOW made no errors when carrying out the algorithm, but made an error when simplifying an algebraic expression. Apparently, LOW did not thoroughly understand how to perform operations on algebraic symbols (e.g., $2n - n$ was simplified to 2); other group members had not addressed those difficulties in the group phase. MED2 made the same mistake on this test that she made in the final episode in group work. She substituted multiplication for addition in the formulas in steps V and VI of the algorithm. In group work MED2 had asked a question about the appropriate operation (at the asterisk in Table III), but her question was not acknowledged.

On the delayed test neither MED2 nor LOW could solve any part of the problem. MED1 made no errors. HIGH made the error that MED1 had made during group interaction related to decomposing the array into layers of dots; she included in the layers all dots in the previous layers.

TABLE IV
Frequency of Categories of Group Process and Errors on Test

	Group Work			Immediate test	Delayed test
Step in Algorithm	Gives Explanation	Receives Explanation	Is Ignored	Makes Error	Makes Error
HIGH					
I	1	0	0	0	0
II	1	0	0	0	0
III	1	0	0	0	1
IV	2	0	0	0	0
V	1	0	0	0	0
VI	2	0	0	0	0
MED1					
I	1	0	0	0	0
II	1	0	0	0	0
III	1	1	0	0	0
IV	1	0	0	0	0
V	3	0	0	0	0
VI	5	0	0	0	0
MED2					
I	0	2	0	0	-[a]
II	0	2	0	0	-
III	0	2	2	0	-
IV	0	0	0	0	-
V	0	0	3	1	-
VI	0	1	2	1	-
LOW					
I	0	0	0	0	-[a]
II	0	0	0	0	-
III	0	0	1	0	-
IV	0	3	1	0	-
V	0	4	2	0	-
VI	0	6	1	0[b]	-

[a]MED2 and LOW submitted blank sheets for the delayed test.

[b]LOW made and algebraic error, not an error in applying the algorithm.

Frequency of categories of interaction during group work and frequency of errors on the tests are summarized in Table IV for each student by step in the algorithm. Only the molar categories of group process shown to relate to achievement are included.

Analysis of the frequencies of observed behavior in Table IV reveals strong relationships between the *content* of group process and performance on the immediate test. HIGH and MED1 explained all six steps in the algorithm and made no mistakes on the test. (Demonstrations of step II were counted as explanations. Rather than describing the arrangements of dots in the array, which would have been awkward, students in all groups in the original study demonstrated the arrangements.) LOW received detailed explanations of steps IV, V, and VI in the algorithm and subsequently performed these steps correctly on the test. Most of the explanations that MED2 received were related to steps I, II, and III and she made no errors on these steps on the test. She received no explanations, however, on step V, and only a sketchy explanation of step VI, and made errors on both steps on the test. Moreover, the group ignored most of her questions about steps V and VI. The insufficient explanations, the fact that she was not given a chance to complete calculations, and the fact that the group did not acknowledge her difficulties seemed to have contributed to MED2's errors on steps V and VI on the test.

Delayed test performance was weakly related to events in group work. The lengthy explanations that LOW received aided her in immediate understanding but not in delayed recall. The sketchy explanations that MED2 received were inadequate for immediate performance and useless for delayed performance. HIGH's argument with MED1 about step III during group work interfered with her own delayed recall; she remembered the faulty arguments posed by MED1 rather than her own correct arguments.

It should be noted that difficulties during training, where all students worked individually, did not seem to relate to group interaction nor to test performance. LOW had trouble with step II during training (marking the correct number of dots on the array) but did not solicit explanations related to this difficulty, nor did she make errors on this step on the immediate test. MED2 misplaced the parentheses in step VI (the formula for the sum of the series) during training but did not ask for nor receive explanations about parentheses during group work, nor did she make this error on the immediate test. HIGH and MED1 made no errors during training.

The general roles that these students played in group interaction can be described simply: HIGH and MED1 were active explainers, LOW was an active solictor of explanations, and MED2 was not an active explainer nor an active solicitor of explanations. Role was in part related to a student's ability level. Although MED1 and MED2 had similar ability compared to the ability of the other group members, MED1 had somewhat higher ability than MED2. Thus, the two members of the group with highest ability explained and the low-ability member solicited (and received) explanations. The other medium-ability member was, by and large, ignored.

The preceding analysis of group process and achievement suggests that the roles of explaining and soliciting explanations overrode the expected effects of ability. Because HIGH had higher ability than MED1, one would expect HIGH to perform better than MED1. This was not the case. On the immediate test HIGH and MED1 obtained perfect scores but on the delayed test MED1 outperformed HIGH. (One might argue that their equal scores on the immediate test may be due to a ceiling effect of the test; on a more difficult test, HIGH may have outperformed MED1. Their performance on the delayed test, however, makes this counterhypothesis less tenable.) Although LOW had considerably lower ability than MED2, she outperformed MED2 on the immediate test.

IMPLICATIONS FOR FURTHER RESEARCH

To understand how the roles of explaining and receiving explanations influence learning, it is important to know the cognitive psychological mechanisms involved. Wittrock's (1974) model of learning as a generative process, in which "effective instruction causes the learner to generate a relationship between new information and previous experience" (p. 182), suggests a hypothesis. Explaining and receiving explanations may help learners understand new concepts and relate them to concepts already understood. Students discuss concepts in their own terms and, consequently, may understand each other's explanations better than explanations from books or even teachers. Receiving explanations from one's peer may help the learner to generate a description or definition of a relation in terms that he or she can understand. Giving explanations in familiar language may help the learner to solidify a definition in memory. The hypothesis predicts that the probability of making an error concerning a concept or relation on the test is negatively related to the incidence of explaining and receiving explanations concerning that concept or relation in group work. That the data from the group analyzed here is consistent with the predictions suggests that the hypothesis is plausible. However, the function that relates the probability of making errors to explaining and receiving explanations must be described more precisely and tested empirically.

The analysis here of group process in one interacting group is suggestive, but further analysis is needed to better relate interaction among group members to cognitive processes and achievement. Examination of processes and outcomes in other groups of heterogeneous ability and in groups of homogeneous ability may reveal other aspects of participation that relate to achievement. Longitudinal studies of small groups are needed to determine how students' roles in group process change