The Development of Children's Knowledge of Temporal Structure

William J. Friedman
Oberlin College

Adults in Western cultures are immersed in a variety of time structures that pattern their lives on daily, weekly, and annual scales. Knowledge of these structures is involved in important forms of thought, including orientation, planning, and recollection, but little is known about the mental representations and processes underlying this knowledge. In recent studies of adults' knowledge of days of the week and months of the year (Friedman, 1983, 1984; Koriat & Fischhoff, 1974; Koriat, Fischhoff, & Raziel, 1976; Seymour, 1980a, 1980b; Shanon, 1979), researchers have begun to apply cognitive methods and theories to these issues, but much remains to be learned.

Even less is known about the development of children's knowledge of time structures (Friedman, 1978, 1982; Jahoda, 1963). Since the 1920s a number of psychologists and educational researchers have studied children's ability to give the present date, recite the months, and so forth. Unfortunately, these studies seldom employ either systematic task variation or error analysis and thus provide little more than an indication of the ages at which children become familiar with the clock and calendar. In a more recent study, Friedman (1977) examined the operations that 3–10-year-olds could perform on natural and conventional systems (e.g., parts of a day, seasons of the year, days of the week) using a series of card-arrangement tasks. The main findings were that by 4–6 years most children could construct a correct order of daily activities or seasons, and that by 8 years (the first age group in which most children could read the cards) they could order day-of-the-week and month cards. However, there were important limits to the initial representations of order—for example, the inability to judge the correctness of veridical orders starting with a different element—that were not overcome until later ages. Apparently, representations of temporal structure, even for a particular content, develop gradually. The initial encoding and processing of order preserves part of the formal structure but shows certain rigidities that are not present at later ages.

The present research concerned how representations of two temporal structures, the days of the week and months of the year, change with age. To ask questions about representations, it is helpful to begin by considering both the formal properties of the information and the ways in which it might be coded and processed. Like the number system, these contents can be conceptualized as a set of elements and the specific relations between the elements. The elements are the days and months themselves, and they are formally related in a variety of ways, for example, by sequential forward and backward order, recurrence, and relative proximity to other elements. There are a number of psychologically plausible representational models of these systems, models that vary some-

This work was supported by Research and Development grants from Oberlin College. I am grateful to Anita Belofsky, Lee Fertig, and Jeanine Spelt for testing subjects and to the staff and children of the Oberlin City Schools.

[Child Development, 1986, 57, 1386–1400. © 1986 by the Society for Research in Child Development, Inc. All rights reserved. 0009-3920/86/5706-0016$01.00]
what in the particular structural relations captured. One model that has been proposed to account for adults' knowledge of days of the week and months of the year distinguishes two representational-processing systems, a verbal-list system and an image system (Friedman, 1983). The verbal-list system codes the links between each element and its successor, with much weaker reciprocal links. Processing involves the sequential verbal-articulatory activation of individual elements, and this activation is a discrete event that can be counted. The image system codes spatial-like relations between the elements, and an image, once constructed, allows direct access to position-like information about the elements activated.

Other possible models of month and week knowledge include semantic models (Seymour, 1980a, 1980b), exclusively serial models, and models using operations on numerical equivalents of the elements. In studies with adults, only the Friedman (1983) model has consistent support, and the others are all inadequate for explaining performance on at least some of the tasks that adults can solve (Friedman, 1983, 1984). However, none of the models have been tested developmentally, and it remains to be determined which, if any, of them can capture the performance of children between the time when the systems are first learned and the time when adult competence is achieved.

A plausible prediction, related to the first model, is that the systems are first represented solely as a verbal list, and only later is this form supplemented by image representations. Most children are taught the systems as a list to be recited, like counting numbers and the alphabet, and it would thus not be surprising if their first representations were a series of links between sequential elements (see Fuson & Hall, 1983, and Siegler & Robinson, 1982, for discussions of early counting and Klahr, Chase, & Lovelace, 1983, for a discussion of the alphabet representation in adults). However, just as numerical representations (Siegler & Robinson, 1982) and spatial representations (Liben, Patterson, & Newcombe, 1981; Siegel & White, 1975) change to forms that allow a more flexible extraction of relations between elements, so temporal representations may change to include a form like the image system. The image model of the second stage has the advantage of being consistent with the end state, that is, the performance of adults. However, it is difficult to predict the ages at which the developmental transition might occur. Certainly children are capable of using imagery on some tasks by the time they usually learn the days of the week and months (e.g., Dean & Harvey, 1979; Marmor, 1975).

I conducted three experiments to evaluate the prediction that representations develop from a verbal-list stage to a stage at which image representations are present. These experiments also allow one to consider the other models developmentally and to determine at approximately what ages adult competence is reached. In Experiments 1 and 2, I used the ability to judge "backward" order as the criterion for the second representational stage. In Experiment 3 another criterion was used—the ability to directly judge the approximate "distance" of remote times.

In the backward-order tasks, subjects judged the relative backward cyclic order of noncontiguous elements (e.g., if one imagines going backward in time from Thursday, would Tuesday or Sunday be encountered first?). The problems are cyclic in the sense that they assume continuity between adjacent weeks or years. I excluded contiguous elements (e.g., Wednesday when going backward from Thursday) so that the problems could not be solved simply by knowing that two elements are adjacent. This type of task, which is readily solved by adults (Friedman, 1983), should be quite difficult if one's representation codes mainly links between elements and their successors. However, image processing need not have strong directional properties because position-like information is directly accessible. Performance on the backward tasks was contrasted with performance on similar forward-order tasks. These should be amenable to solution by a sequential process, as long as multiple steps through the order can be made.

**Experiment 1**

Second, fourth, sixth, eighth, and tenth graders and undergraduates judged which of 2 days of the week would come next after a third, reference event going either forward or backward in time. In pilot testing, second graders were the youngest group in which the majority of subjects claimed to know the days of the week and could recite them. Stimuli were presented orally so that reading was not required.

**Method**

**Subjects.**—Fourteen children from each of grades 2, 4, 6, 8, and 10 and 16 undergraduates served as subjects. Each group was evenly divided by sex except the tenth-grade...
group, which had six girls and eight boys. The school-age subjects came from public schools in a small-town community. The eighth and tenth graders were tested outside of school hours and paid $82.00 for their participation. The undergraduates were volunteers from an introductory psychology class. The ages for school-age subjects are available only in whole years, so the following means provide a slight underestimate of the mean age. The means for grade 2 through grade 10 are 7.8, 9.9, 11.8, 13.1, and 15.2, respectively.

Stimuli.—School-age subjects received eight sets of seven problems each, and college subjects received 12 sets of seven problems. Only the four sets relevant to the school-age subjects received 12 sets of seven problems. The order of problems within a set varied randomly across subjects. In addition, the order of the eight or 12 sets was randomized across subjects. The order of mention of the two choice days was determined randomly and held constant across subjects.

Procedure.—Subjects were tested individually in a single session. At the start of the procedure the second graders only were asked if they knew the days of the week, and all successfully recited them. Before the first set of problems, subjects were given the following instructions, with the direction (forward or backward) appropriate for the first task. “Think about going forward [backward] through the days of the week. In the question, I’m going to say 2 days that could go next going forward [backward]. Then I’m going to give you another day to start with. When I say the day that we start with, I want you to think which of the 2 other days comes next going forward [backward].” Next, subjects were given a sample problem. The sample problem for the forward instruction was, “Does Saturday or Sunday come next after Monday?” The sample problem for the backward instruction was, “Does Monday or Tuesday come next after Thursday?” The instructions and examples were repeated with paraphrasing until the subject appeared to understand the task. If the subject seemed confused after the first few items, the instructions were repeated. The instructions for the other direction were given just before the first set in that direction was presented. The experimenter emphasized the direction of each problem set before it was presented.

The experimenter read each problem aloud and recorded the subject’s first response. Subjects were told to say “I forgot” if they forgot the question or lost their concentration. In such cases the problem was scored as missing. All other answers were scored as correct or incorrect.

Results and Discussion
Performance on the forward sets was evaluated by finding a conjoint score, the number correct out of 14 (or the number of nonmissing problems) in the two forward sets. No distinction was made between the two forward sets because set differences in accuracy were small and inconsistent from age group to age group. A single score was obtained for the two backward sets for the same reason. Subjects who solved 12 or more problems out of 14 were classified as passing the task, \( p < .02 \) by the binomial test. In the case of missing problems, the criterion was 10 or more out of 12, \( p < .04 \), and 11 or more out of 13, \( p < .03 \). Missing problems were rare, and no subject was missing more than two of the 14 for a direction.

Table 1 shows the number of subjects who met the criterion in each age group. More than half of the subjects at each age from fourth grade onward were able to accurately judge the relative forward order of 2 days from a third, though there was an unexpected decrease in accuracy from sixth grade to eighth grade. For the backward tasks it was not until about the tenth grade that half or
The results of Experiment 1 indicate that more subjects met the criterion. Performance on both the forward and backward task types improved with age, \( \chi^2(5, N = 86) = 42.00, p < .001 \), and \( \chi^2(5, N = 86) = 34.93, p < .001 \), respectively. There appears to be little change on either task type from fourth to eighth grade, and fewer than a third of the subjects in these groups were able to consistently judge backward relative order.

Order of mastery of the two types of problems was tested by cross-classifying subjects by their pass or fail status on each task. Of the 24 subjects who met the criterion for one but not both task types, all met the forward criterion, \( p < .0001 \), by the binomial test, two-tailed, showing that the forward task was mastered at an earlier age. Twenty-six subjects failed both tasks and 36 passed both tasks.

The results of Experiment 1 indicate that early representations and processes capture only part of the temporal structure that is formally present in time systems. In fact, the first ability measured here, naming the days of the week in order, does not even imply the ability to solve the forward problems. Only one of the 14 second graders could judge relative forward order even though all could recite the days of the week. The former task requires adopting different reference points, crossing the conventional boundary between one week and another, and extracting the order through more than one step before producing a response. The representations and processes underlying the early ability to recite the days of the week may not be sufficient to meet these requirements. Of course, superficial task features, such as the necessity of holding the two choice days and the reference day in memory, may also limit the performance of the second graders.

### TABLE 1

<table>
<thead>
<tr>
<th>Grade</th>
<th>Forward Pass</th>
<th>Forward Fail</th>
<th>Backward Pass</th>
<th>Backward Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Fourth</td>
<td>9</td>
<td>5</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Sixth</td>
<td>12</td>
<td>2</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Eighth</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Tenth</td>
<td>14</td>
<td>0</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>College</td>
<td>16</td>
<td>0</td>
<td>16</td>
<td>0</td>
</tr>
</tbody>
</table>

What is clearer is the developmental sequence between judging forward relative order and backward relative order. Here, the two tasks are formally similar, but there is a considerable delay between the ability to solve forward problems and the ability to solve backward problems. This disparity suggests that representations and processes that are adequate for extracting relative forward order need not be adequate for judging a novel relation such as relative backward order. Another possibility, however, is that the failure of the younger children on the backward tasks was due to an inability to understand what it means to go backward through the days of the week and not to limitations of their representations. However, error analysis showed that only one subject, a second grader, consistently answered according to forward order in the backward conditions. Thus, it seems unlikely that many subjects who failed the backward task did so because they construed it as requiring forward judgments. Furthermore, in an unpublished study of similar backward- and forward-order judgments of familiar daily events, I found that 12 of the 20 third graders were significantly accurate, \( p < .04 \), on the backward task. This would tend to indicate that the limiting factor in the present study, at least for the fourth grade and older groups, is not the failure to understand the idea of going backward in time. It may instead be the gradual development of representations for this content during middle childhood and adolescence. Experiment 3 provides convergent support for this conclusion by using a task that is solved by fourth grade but on which preadolescents appear to use different processes.

The results of this study are consistent with the model predicting a transition from exclusively verbal-list representations to a later stage at which image representations are also present. The former should be adequate for the forward tasks, which are solved by fourth grade, whereas the latter may be required for the backward tasks, which are not solved until adolescence. Also consistent with this model are subjects’ introspective reports. In response to a probe question, randomly inserted after one of the four tasks, most subjects from grades 2 through 10 who provided ratable responses claimed to be saying the days to themselves while solving the problems. From sixth through tenth grade, nearly a third of the subjects reported some sort of spatial image of the days of the week, and the adult subjects, who filled out a method report sheet consisting of Likert-type scales, reported "picture the days in your..."
mind” far more frequently than “saying the intervening days to yourself.”

Each of the other models has difficulty accounting for subjects’ performance. The semantic model fails to explain how subjects could solve either task, and is therefore inadequate for explaining the competence of the fourth grade and older groups. This is because semantic models rely on absolute position codes that are insufficient for solving relational problems with a changing reference point (see Friedman, 1984). A serial activation model alone does not easily explain the high levels of accuracy eventually reached on the backward tasks or the shift from forward only to forward and backward competence. The main difficulty for the numerical model is that it, like the semantic model, relies on absolute position codes in the form of numbers. These are sufficient for problems in which all of the days fall within the same conventional week (e.g., does Wednesday [4] or Friday [6] come next after Monday [2]?), but to solve cross-boundary problems (e.g., does Sunday [1] or Tuesday [3] come next after Friday [6]? additional operations are needed to alter the postboundary number, along with rules about when to do so. Even such an elaborated model does not fare well in predicting errors. An analysis of errors as a function of individual problems showed very little difference in accuracy between problems that cross and do not cross the boundary. The numerical model also fails to explain subjects’ reports of reciting and imagery.

Experiment 2

This experiment concerned the ability of 9–20-year-olds to judge forward and backward relative order of the months of the year. The youngest subjects were from the third grade, the first grade in which, according to pilot testing, most children could recite the months in order. The tasks were similar to those used in Experiment 1.

Method

Subjects.—Sixteen children from each of grades 3 and 5 and 16 undergraduates were initially tested. Subsequently, tenth graders were added to the sample to provide additional information about performance changes between fifth grade and college age. Data from the college subjects have been published previously in Friedman (1983, Experiment 3) and are presented here for the purpose of developmental comparison. Half of the subjects in each group were female and half male, except for the tenth-grade group, which consisted of nine females and seven males. The mean ages for the four groups are 8.9, 10.7, 15.9, and 20.0 years. The school-age subjects were from public schools in the same community as the subjects of Experiment 1. The college subjects were volunteers from psychology classes.

Stimuli.—Eight problems from each of four sets were used as stimuli. Only the two sets relevant to the present study, the forward and backward sets, will be described in detail and reported. The remaining sets required subjects to judge which of two choice months was a particular distance from a reference month.

Problems in both the forward and backward sets were of the same form, but instructions at the start of each set emphasized the direction for that set. A sample sentence is, “Does January or September come next after May?” The choice months were always 4 and 8 months away from the reference month in the appropriate direction. The eight reference months were randomly selected for each set with the restriction that two come from each quarter of the year. Problems were free to cross the December–January boundary.

Order of problems within a set and order of mention of the two choices in each problem were randomized and held constant across subjects. The four problem sets were given in one of eight orders to each subject. Four of the orders were cyclic permutations of one another, and the remaining four were left-right inversions of the first four.

Procedure.—Subjects were tested individually in a single session. Before the forward set they were given the instruction, “If you go through the year, tell me the month that will come next after these.” If the backward task preceded the forward task, the word “forward” was inserted before the word “through.” For the backward set, the instruction was, “If you went backward through the months of the year, tell me the month that will come next after these. Remember, you will be going to earlier months.” After each of the first few problems in each set, the experimenter reminded the subject of the forward or backward direction. At the end of each task, subjects were asked, “How did you figure out the answer to these questions?” Next they were asked if they “said the months or thought about a picture” or did it some other way.” The order of mention of “said the months” and “thought about a picture” varied across subjects.

Responses to each sentence were recorded and scored as correct or incorrect. Re-
Responses to the method probes were subsequently assigned blind by task, sex, and (for all but the tenth-grade group) by age into the categories “image,” “reciting,” “both,” or “other.” Responses categorized as “image” included referring to a continuum, cycle, or calendar descriptively or by gesture. The reciting category included saying the intervening months to one’s self, “counting,” or “reading through the months.”

Results and Discussion
Performance on the forward and backward sets was evaluated by computing the number of problems correct out of eight. Subjects who judged seven or eight correctly were classified as passing the task. This criterion corresponds to a minimum proportion correct of .87, a figure close to the .86 proportion used in Experiment 1 but somewhat less conservative against the null hypothesis of chance responding, binomial $p < .08$, two-tailed.

Shown in Table 2 is the number of subjects in each age group who met this criterion in the forward and backward sets. By the fifth grade, more than half of the subjects were able to consistently judge relative forward order, but apparently not until several years later are comparable levels reached on the backward task. Both the tenth-grade and college groups perform about equally well on the two task types. Age improvement was substantial and significant on both the forward and backward sets, $x^2 (3, N = 64) = 24.00, p < .001$, and $x^2 (3, N = 64) = 35.50, p < .001$, respectively.

The developmental sequence of mastering the two task types was tested as in the previous experiment. A binomial test showed that the forward task was a prior achievement, $p < .01$, two-tailed, with 13 of the 15 subjects who passed only one task passing the forward task. Twenty-two subjects failed both and 27 passed both.

Results from the method probes are summarized in Table 3. Only responses that fell in the imagery and reciting categories are tabulated. At the third and fifth grades, most subjects reported reciting in both the forward and backward tasks, and few referred to images. In contrast, half or more of the subjects in the older two age groups were assigned to the imagery category in each task. The age changes for both the forward and backward tasks were significant, $x^2 (3, N = 56) = 21.50, p < .001$, and $x^2 (3, N = 54) = 25.58, p < .001$, respectively.

As for the previous experiment, the results of Experiment 2 reveal a substantial difference between the ages at which subjects can judge forward relative order and backward relative order. The developmental priority of forward order and the age trends in reported methods are consistent with the two-stage model, and, for the same reasons as in Experiment 1, the results do not support the alternative models. Among the other difficulties for the alternative models, a comparison of problems that crossed and did not cross the December–January boundary showed very small differences, thus failing to support numerical models.

Additional support for the verbal-list characterization of the third and fifth graders comes from observations of these children while they solved the problems. Lip movement was detected on an average of about one-quarter of the problems in these age groups but almost never among the college subjects. (This information was not collected for the tenth graders.) It is also notable that reports of imagery were frequent in the tenth-grade and college groups, both of which showed considerable success with the backward task.

Experiment 3
Fourth, seventh, and tenth graders and undergraduates were tested on day-of-the-week and month tasks requiring categorical judgments of the forward distance between one element and another. In one task they judged whether or not a given day of the week was one of the next 3 days after a reference day. In another task they judged whether or not a particular month was one of the next 4 months after a reference month. It was assumed that subjects who have an image representation of days of the week or months can rapidly judge remote items (e.g., those 6
units away) to be beyond the category boundary. Subjects whose representation only allows sequential activation of elements would have to step through the three or four items after the reference item in order to determine, by default, that a remote item belonged to the postboundary category.

The main difference between the performance of subjects with the two types of representations would be seen in response-time (RT) curves plotted as a function of number of units between the reference element and the other element. An optimal sequential approach should lead to a substantial linear increase in RT as a function of distance up to the boundary and a flat curve thereafter. This follows from the assumption that stepwise activation requires approximately equal increments in RT for each additional element and from the assumption that the search will be terminated when either the second stimulus element or the preboundary element is encountered. According to the model, an image representation directly codes the approximate relative positions of items. Therefore, distant items should have RTs approximately equal to, say, the next item or two after the reference element. The strongest effect might be expected to be longer RTs for items adjacent to the category boundary (see Friedman, 1984; Seymour, 1980a, 1980b, for supportive findings in adults) because in some image processes the elements' "locations" may be difficult to discriminate from the boundary "location." Because both models predict relatively long RTs near the category boundary, the clearest differences come from comparing "remote" items beyond, but not adjacent to, the category boundary with some point close to the reference element, such as the next item. This difference should be small for the image model, assuming direct access to position information, but large for the sequential activation model, assuming stepwise activation of each element up to the boundary. The size of the latter can be predicted from previous studies of the months and week to be on the order of ½ sec per item (Friedman, 1983, and unpublished studies). I expected that the performance of children who had only recently learned the days of the week and months of the year would be consistent with the sequential-activation model and that older children would shift toward the pattern associated with the image model.

The performance comparisons of this experiment do not depend on differential success and thus guard against a possible criticism of Experiments 1 and 2 that age differences could be attributed to a failure to understand the task. Another change in Experiment 3 was the use of a story format and feedback on performance to increase subjects' engagement in the tasks.

**Method**

**Subjects.**—Sixteen subjects from each of grades 4, 7, and 10 participated in the study, as did 16 undergraduates. Ten of the fourth graders, seven of the seventh graders, six of the tenth graders, and eight of the undergraduates were male. The mean ages for the four groups were 10.3, 13.5, 16.6, and 19.0 years. The school-age subjects were from public schools in the same community as the subjects in Experiments 1 and 2. They were paid $3.00 for their participation. The undergraduates received credit toward a requirement for an introductory psychology course.

**Stimuli.**—The experiment consisted of two main tasks—a week task and a month task. On each trial a pair of days or months was presented. In the week task there were 42 trials, with each day paired with each other day. Thus there were seven problems for each forward distance of one to six. In the month task there were 48 problems with eight forward distances, 1, 2, 3, 4, 6, 7, 8, and 9. The 6 starting months for each distance were from each sixth of the year, with the particular month being chosen randomly from the pair in the sixth (e.g., January or February). Problems in each task were presented in a constant random order with the restriction that...
one problem from each distance appear in each seventh of the order in the week task and each sixth of the order in the month task. Order of presentation of the week and month tasks was counterbalanced.

**Procedure.**—Subjects were tested individually in single sessions that took place in a laboratory on a college campus. They were seated facing an Apple II Plus computer with an A2M2010 monitor. The experimenter was seated beside the subject. After the experimenter entered the subject's name and age, the instructions for the first task appeared on the screen. The experimenter read them aloud as the subject read along silently. Subjects were given an opportunity to review the instructions if they wished before testing began. The instructions for the week task are given below.

Vorn is a creature from another planet. He is trying to escape from his enemies. When Vorn lands on earth, he can make himself invisible. But he can only stay invisible for the next 3 days after he lands. For example, let's pretend Vorn lands on Friday. He will stay invisible on Saturday and Sunday and Monday. But on Tuesday and Wednesday and Thursday he will be visible.

In this game we will tell you what day Vorn lands on earth. We will also tell you what day his enemies land. Your job is to figure out if they will be able to see him.

If they would be able to see him, press the button showing that he could be seen.

If he would be invisible, press the button showing that he would be invisible.

Here is an example.

**Vorn lands Friday**
**Enemies land Monday**

Which button would you press? You would press the "invisible" button because Vorn would still be invisible.

Here is another example.

**Vorn lands Friday**
**Enemies land Tuesday**

Which button would you press? You would press the "seen" button because Vorn could be seen.

After each problem comes on the screen, try to press the correct button as quickly as you can. But also try to be correct. As soon as you have pressed a button, the computer will tell you whether your choice was correct or incorrect. Everyone is expected to make some mistakes, so do not worry if you get some wrong.

A metal plate covered all but the two response buttons ("Z" and "/"). Each of the two had a picture over it (a line figure with antennae and a square with wavy diagonal lines) to remind subjects of its meaning. On each trial, two sentences of the form illustrated in the instructions appeared centered on the screen. Subjects initiated each trial by pressing either key. Following the subject's response, the display indicated "Correct" or "Incorrect," or "Too Late!" and "Incorrect" if the subject had not yet responded by the time 20 sec had elapsed. Response and RT were automatically recorded.

The experimenter observed subjects as they performed the task and made a trial-by-trial record of the overt behaviors: moving lips, reciting elements aloud, counting on fingers, and tapping periodically. Following the task subjects were asked, "How did you think about the days to get the answer?" Next they were asked, "Can you tell me the 3 days that come after ———?" for the days Saturday, Wednesday, and Sunday.

The procedure and instructions were similar for the month task. However, in this task subjects were to pretend to route messages to an astronaut who would be in orbit for the 4 months after the month of the launch and back at the home base thereafter. A sample stimulus is "Launch in November—Message in March," for which the "orbit" button would be the correct response. For a November launch and an April message, the "home base" button would be correct. The two pictures above the "Z" and "/" keys showed a building with antennae on the roof and a space ship orbiting the earth. After the method probe, subjects were asked, "Can you tell me the 4 months that come after ———?" for the months October, August, and January.

**Results**

**Week task.**—Response time and proportion correct were averaged across the different problems for each distance. Performance data for two seventh graders and one tenth grader were lost due to a programming error. Shown in Figure 1 is RT as a function of distance, separately for each age group. In this figure and the subsequent analyses for the week and month tasks, both correct and incorrect trials were used. Separate analyses using only correct trials and only correct trials in which the RT was less than 10 sec produced highly similar patterns. At each age RT increases with distance up to the boundary between 3 and 4 days, with no apparent subsequent increases. However, the slope from distance 1 to the remote times, beyond but not adjacent to the distance boundary (distances 5 and 6), becomes substantially smaller in older groups.
Response time was subjected to an ANOVA with group, sex, and order (week task first or second) as between-subject factors and distance as a within-subject factor. Sex was not a significant factor alone or in interaction, and order was marginally significant, $F(1,45) = 4.09, p < .05$, with subjects receiving the week task second responding more rapidly. Group, $F(3,45) = 31.29, p < .0001$, distance, $F(5,225) = 24.12, p < .0001$, and the group x distance interaction, $F(15,225) = 2.05, p < .02$, were all significant. Post-hoc Tukey comparisons for the group effect showed that fourth graders were slower than tenth graders, $p < .05$, and college students, $p < .05$.

Trend analyses conducted separately for each age group showed significant linear and quadratic trends for distance, $p's < .01$, for each group, but the strength of the linear effect decreased substantially with age. The age change in the linear effect is largely captured in another analysis designed to test for group differences in the time needed to access remote relative to proximal items. This analysis, which is equivalent to a planned comparison for the group x distance interaction, was run as a between-groups ANOVA. The difference between the distance 1 mean and the average of the distance 5 and 6 means was computed for each subject. A one-way ANOVA with age group as a between-subject factor showed a significant group effect, $F(3,57) = 4.90, p < .05$. The linear component was highly significant, $F(1,57) = 13.60, p < .001$, whereas the deviation from linearity was not significant. This shows a linear decrease with age in the additional time needed to categorize remote as opposed to proximal days. Post-hoc Tukey comparisons supported the difference between the fourth-grade and college groups, $p < .01$.

The magnitude of the additional time needed to access remote items can also be compared to an estimate from a sequential-access model. Using the $\frac{1}{2}$ sec per step estimate from previous studies, and assuming that subjects activate each day up to the boundary distance, one would expect a 1-sec difference between the RTs of distance 1 and remote items. For each age group, 95% confidence intervals were placed around the mean difference between distance 1 and distances 5 and 6. The intervals are $1155 \pm 564$, $677 \pm 263$, $506 \pm 216$, and $345 \pm 220$ msec for fourth-grade through college groups, respectively. Only the fourth graders' confidence interval overlapped the 1-sec value.

Shown in Figure 2 is the proportion of problems correct as a function of the distance after the first day. For each age group, distance 3 problems have the greatest error rates. An ANOVA similar to that for RT showed distance to be the only significant factor, $F(5,225) = 30.51, p < .0001$. Thus it appears that subjects in different groups perform with approximately equal accuracy, and all have the greatest difficulty with problems just before the distance category boundary.

Overt behaviors were analyzed as a single category, on the assumption that all were indicative of verbal list or other serial processing. Fourth graders frequently engaged in overt behaviors before responding (on a mean of .41 of trials), whereas older subjects seldom did (means of .04, .03, and .03 for the three groups, respectively). For each age group, overt behaviors were more common at dis-
distances near the category boundary than at distances far from it.

Subjects reported a variety of methods for solving the problems, the most common being counting or counting plus some other method. Many other subjects claimed to directly know the answer in some way. Unfortunately, a large proportion of the reports were vague and difficult to categorize, and attempts to achieve acceptable interobserver agreement were unsuccessful.

Month task.—Shown in Figure 3 is RT as a function of distance for each age group. Response time and error data for four seventh graders and eight tenth graders were lost due to a programming error. As for the week task, the RTs generally increase up to the boundary with no apparent subsequent increases. In addition, the slope from distance 1 to the remote items (distances 6–9) decreases from younger to older age groups. An ANOVA similar to that for the week task showed significant effects for group, $F(3,36) = 22.15$, $p < .0001$, distance, $F(7,252) = 10.53$, $p < .0001$, and the group $\times$ distance interaction, $F(21,252) = 2.30$, $p < .01$. The interaction between sex and distance was also significant, $F(7,252) = 3.09$, $p < .01$, with males showing a somewhat flatter distance curve and females a more curvilinear distance curve. Post-hoc comparisons for the group effect showed that college students responded significantly more rapidly than fourth graders, $p < .05$. As for the week task, trend analyses showed significant linear and quadratic trends for distance for each age group, $p$'s $< .02$, with substantial decreases in the linear effect with increasing distance.
To compare proximal and remote items, the difference between the distance 1 mean and the average of the distance 6, 7, 8, and 9 means was computed. The one-way ANOVA of this variable by age group showed a strong group effect, \( F(3,48) = 10.52, p < .0001 \), with a strong linear effect, \( F(1,48) = 28.73, p < .0001 \), and a nonsignificant residual. As for the week task, this shows a linear decrease with age in the time needed to judge remote as compared to proximal months. Post-hoc Tukey comparisons, weighted for unequal group sizes, showed that fourth graders differed from each of the other groups, \( p's < .05 \).

Ninety-five percent confidence intervals placed around the differences between distance 1 and remote items showed that the fourth and seventh graders' intervals included the 1.5-sec difference predicted by the counting model (i.e., \( \frac{1}{2} \) sec per month from distance 1 to distance 4), whereas the older groups had significantly faster access times. The intervals for the four groups in order of increasing age are 2111 ± 776, 997 ± 565, 681 ± 423, and 340 ± 180 msec.

Proportion correct is shown in Figure 4. Again, performance is least accurate immediately before the category boundary. The ANOVA revealed significant effects for group, \( F(3,36) = 5.79, p < .01 \), and distance, \( F(7,252) = 16.73, p < .0001 \), but their interaction was not significant. Sex, \( F(1,36) = 6.26, p < .02 \), and order, \( F(1,36) = 4.62, p < .04 \), were also significant. Females were more accurate than males, and subjects who received the month task second were more accurate than those who received it first.

Overt behaviors were tabulated as for the week task. Again, there was a sharp drop in the frequency of overt behaviors after fourth grade, with mean frequencies of .37, .08, .05, and .04 for the four groups. Subjects reported counting, knowing directly, and a combination of the two as the most common methods for solving the problems. However, too many reports were ambiguous to achieve acceptable interobserver agreement on method categories.

**Reciting tasks.**—After the main week task, subjects were asked to give the 3 days after Saturday, Wednesday, and Sunday, and after the main month task they were asked to give the 4 months after October, August, and January. Errors on these tasks were tabulated as a function of problem, group, and error type. Only five errors occurred on the week task. These were made by two fourth graders and one tenth grader. However, errors were common on the month task, with 10 errors on the October problem, 16 on the August problem, and five on the January problem. The number of subjects making at least one error on the month task was 11, 6, 3, and 0 for the four age groups, respectively. Combining the week and month tasks, the most common error types were omissions (e.g., September, November, December, January)—12, inserting an inappropriate segment in the list—6, reversing the order of two elements—5, and running on beyond the last element required—4. There was little indication of difficulty crossing the boundary between December and January in the October problem, with only two errors and one noticeable pause at this point. It is notable that the most common types of errors would have led to classification errors near the boundary on the main tasks, with fewer errors on the first item or two after the starting element and on re-

![Figure 4](image-url)

**FIG. 4.**—Mean proportion correct for each distance in the month task of Experiment 3.
mote items. It is also notable that accuracy seems to vary between the month problems, with January being a particularly easy starting point and August being a particularly difficult one (see Klahr et al., 1983, for a similar effect in adults’ alphabet order judgments).

Discussion

The findings for both the week and month tasks are consistent with a developmental shift from verbal-list representations to image representations. The fourth graders showed a substantial difference between the response times for remote and proximal problems, a difference on the order of ½ sec or more for each additional step up to the category boundary. This difference is of a magnitude comparable to that found in previous tasks in which subjects appeared to use a verbal-list process. The difference decreased progressively with age, with the older two age groups showing differences that are probably too small to be attributed to a verbal-list process.

The pattern of overt behaviors during the interval between the stimulus and response is also consistent with the two-stage model. For both the week and month tasks, overt behaviors, such as lip movement, reciting aloud, and rhythmic tapping, were common among fourth graders, including on remote problems. For the older age groups, these behaviors were uncommon, especially on remote problems. Since the few overt behaviors shown by the oldest subjects mainly occurred near the distance category boundary, it seems that these subjects sometimes engaged in a two-stage process, first determining if the target was near the boundary and then reciting if necessary.

The accuracy data for both tasks showed the highest error rates at the preboundary distance. The high error rate for these distances itself can be explained by either the verbal list or image models. Under the former, one would only need assume that subjects made recitation errors of the sort detected on the posttasks. Under an image model one would expect boundary proximity effects of the sort seen in previous studies (e.g., Friedman, 1984). However, it is somewhat surprising that the week distance 4 problems should not show similarly high error rates, because they are also adjacent to the category boundary. The special difficulty of the distance 3 problems on this task may be attributable, in part, to confusion, articulated by several subjects, over whether that distance was to be included in the “visible” or “invisible” category. Another possible explanation is that subjects sometimes erroneously included the “landing day” in the 3-day interval.

General Discussion

Children’s representations of the temporal structure of the days of the week and months change in the years after the systems are first learned. A number of findings can be interpreted in terms of a two-stage model in which order information is first stored and processed in a verbal-list system and at a later stage is also stored and processed in an image system. In Experiments 1 and 2, forward order could be accurately judged by fourth or fifth grade, within about 2 years of the time when the days or months can be recited in order, but backward order judgments could not be made with comparable levels of accuracy until about the tenth grade. There was also a shift in introspective reports from methods involving reciting to the use of imagery. In Experiment 3, fourth graders showed a large difference between the categorization time for proximal and remote items, with diminishing differences in older groups. Many fourth graders showed overt behaviors indicative of a reciting strategy even on remote items, whereas seventh and tenth graders and adults rarely showed overt behaviors.

None of the other models seem to provide a good account of these findings. Semantic models cannot explain how the tasks in Experiments 1 and 2 are solved and fail to predict the strong category boundary effects in Experiment 3, because the boundaries are defined by changing reference points (see Friedman, 1984). Pure serial-activation models do not explain the ability of older subjects to judge backward order or rapidly access information about the position of remote items. Numerical models would have difficulty explaining the relative ease with which subjects solve problems that cross from one cycle to another in Experiments 1 and 2 and do not predict the category boundary effects in Experiment 3.

In addition to supporting the two-stage model, this study shows that the development of children’s representations of these contents is protracted. Children can usually recite the days of the week and months in a conventional order by about third grade, but, by the two performance criteria used in this study, adult-like competence is not achieved until early to mid-adolescence. Part of this interval is spent mastering the list structure itself. The reciting task in Experiment 3 revealed that most of the fourth graders were able to give the 4 months after the conventional start but
that errors were very common at both fourth and seventh grade when reciting began with an interior month. Failures were uncommon in the Experiment 3 week posttask, but Experiment 1 showed a delay between the age at which children could recite the days in conventional order (grade 2) and the age at which they could judge forward relative order (grade 4). Introspective reports also indicated that reciting was the predominant approach to the latter, and most problems began with interior days, so it seems reasonable to suppose that part of the problem at grade 2 lay in beginning the recitation with an interior day. These steps in the mastery of the list structures appear to parallel those of much younger children who are learning to count (Fuson & Hall, 1983).

Because the second stage is reached at about the same ages for the week and month representations, it is tempting to speculate that a second limiting factor is the absence of certain operations later brought about by a general cognitive transition. While this possibility cannot be ruled out, there are a number of reasons to seek an explanation for the ages of mastery that is specific to these contents. In a study mentioned previously, I found that backward relative order of daily events could be accurately judged by about 9 years, well before similar judgments can be made for the days of the week and months of the year. In fact, the wider literature on the ability to reverse temporal sequences (e.g., Brown, 1976; Fivush & Mandler, 1985) shows accurate judgments at even earlier ages when briefer sequences are used. In a similar vein, it is clear that reciting processes, such as those apparently used in the week and month tasks, are applied by substantially younger children to the number series (e.g., Fuson & Hall, 1983; Siegler & Robinson, 1982). This theme of content-to-content variation in the ages at which particular representational properties are present seems to characterize other domains as well (see, e.g., Mandler’s 1983, discussion of the development of spatial representations, and Keil, 1984, 1985).

To understand the development of representations of a particular content, it seems that three sets of factors must be considered: (1) the formal structure of the domain, (2) natural representational and processing constraints, and (3) the history of operations that are performed on the elements. In the case of the time systems considered here, formal properties like forward order, recurrence, and relative time of occurrence of nonadjacent elements clearly impose limits on the forms that representations can take. Second, cognitive psychologists have suggested a number of functional constraints on information processing, constraints that are probably the products of human evolution. Of relevance here is the distinction in certain theories between verbal-articulatory and image-processing systems (e.g., Baddeley & Lieberman, 1980; Paivio, 1971, 1978). The two systems are presumably responsible for a substantial proportion of the operations that can be performed in working memory. As Friedman’s (1983) model suggests, both systems allow the “online” extraction of relationships between the elements of certain time systems. Of course, the cognitive system allows still other ways of extracting relationships, for example, the direct retrieval of associates to a problem (e.g., see Siegler & Shrager, 1984). From a developmental perspective, one can view these representations and processes as offering potential ways of extracting structural information. But the third factor, the history of operations performed on the content, seems necessary to account for the representations and processes that are actually used at particular points in development.

Children are probably exposed to the names of days and months as early as they hear numbers. But for number, the direct teaching and subsequent rehearsal of counting operations lead to serial representations, with the later addition of cyclic rules, by the preschool years (Siegler & Robinson, 1982). When parents and teachers later teach addition during the early school years, number also acquires representations in the form of problem-answer associations (Siegler & Shrager, 1984). Day and month representations are probably limited to simple element-content associations (e.g., a particular month is linked to the concept “my birthday”) until parents and teachers begin teaching the days and months as series. Rehearsal then leads to development of verbal-list representations and processing, usually by about second or third grade. The subsequent developmental course does not parallel number’s because children are not taught and do not perform addition- or subtraction-like operations. If they were, adults would probably be able to use retrieval to answer problems like “what month is 3 after March” instead of having to use slow verbal-list processes (Friedman, 1983). However, the verbal-list system is not the only one available to adults for reasoning about day and month order, and the questions remain why image representations eventually develop and why they do so around early to mid-adolescence.
It seems likely that in the first several years after the days and months are learned, overt and covert recitation constitute a large proportion of the total operations that are performed on these contents. This is because parents and teachers usually regard the ability to recite the elements in order as the appropriate criterion for "knowing" the days or months. In addition to prompted recitation, preadolescents probably use verbal-list processes for much of their thinking about real-life temporal problems, for example, determining how long it is until some event of interest such as the weekend or a holiday. However, as prompted recitation diminishes, real-life problem solving assumes a greater proportion of the total number of operations performed on these contents. In addition to determining the amount of time until some event, these problems probably include planning the order of future activities (e.g., when different school assignments should be completed) and remembering the order of past events. The verbal-list system is cumbersome for extracting the relative order of nonadjacent elements either forward or backward from the present. Thus, it seems plausible that children will begin to use the image system, initially in conjunction with the verbal-list system, to solve particular problems. For example, in deciding on the best order in which to complete several school assignments, a student might extract the order of the days in question using the verbal-list system, while constructing an image to hold in working memory the relative times at which the different assignments are due. If successful, images may be constructed for more and more problems, and more and more relative times of occurrence may be stored in a permanent image representation. This representation could then be used directly to construct images appropriate for a particular task. Obviously, this account raises numerous questions about the detailed mechanisms involved, and it cannot be empirically evaluated at present. However, without this sort of "natural selection" process, it may be difficult to understand how representations develop.

References


