# Subjective organization of U.S. presidents

NORMAN R. BROWN
IBM Thomas J. Watson Research Center, Yorktown Heights, NY

ROBERT S. SIEGLER
Carnegie-Mellon University

College students appear to subjectively organize U.S. presidents into three groups and to use this organization to help them learn new information. Results of a paired comparison task in Experiment 1 suggested that subjects organized the presidents into Founding Fathers (Washington through John Quincy Adams), post–World War II presidents (Truman through Reagan), and noncontemporary presidents (Jackson through Franklin Roosevelt). This hypothesized organization was tested in Experiment 2 by presenting subjects with several lists with alternative organizations of the presidents. Difficulty of learning classifications of particular presidents within each list was predicted on the basis of the relation between that president's classification in the list and in the hypothesized preexisting organization. Instances where the preexisting and new classifications were consistent both for the particular president and for all other members of the group were easiest to learn. Instances where the particular president's classification was consistent but where some other presidents in the group were classified differently were of intermediate difficulty. Instances where neither the individual president's classification nor that of the entire group were consistent were most difficult. The results provided converging evidence for the hypothesized subjective organization of the presidents and also demonstrated the usefulness of a method that may be applicable to testing subjective organizations in a wide variety of other areas.

Subjective organizations are prominent features of knowledge of many complex real-world domains, among them numbers (Shepard, Kirkpatric, & Cunningham, 1975), chess (Chase & Simon, 1973), algebra (Hinsley, Hayes, & Simon, 1977), dinosaurs (Chi & Koeske, 1983), and letters of the alphabet (Hamilton & Sanford, 1978). Such organizations serve several useful functions. They make material more memorable (Chase & Simon, 1973), more useful for drawing inferences (Spilich, Vesonder, Chiesi, & Voss, 1979), and easier to learn (Siegler & Robinson, 1982).

However, assessing the exact nature of subjective organizations is

AMERICAN JOURNAL OF PSYCHOLOGY
Spring 1991, Vol. 104, No. 1, pp. 1–33

difficult. People are extremely flexible in relying on different aspects of their knowledge to perform seemingly similar tasks (Sternberg & Powell, 1983). For this reason, knowledge assessments yielded by one approach often do not match those yielded by other approaches (Strauss & Ephron-Wertheim, 1986; Surber & Haines, 1987; Wilkening & Anderson, 1982). This creates a need for converging measures that can be used in tandem to validate hypotheses about subjective organizations and other types of knowledge.

This article demonstrates that convergent validation for existing organizations can be obtained by examining how new subjective organizations are learned. The influence of existing knowledge on learning is widely recognized (e.g., Anderson, 1983; Bransford, 1979; A. Brown & DeLoache, 1978; Collins & Stevens, 1982; Siegler & Klahr, 1982). Applied to the issue of subjective organization, the influence suggests that it should be easier to learn material that is more consistent with an existing organization than material that is less consistent with it. In particular, when presented with a new organization, it should be easiest to learn material that is both locally and globally consistent with the hypothesized existing organization; it should be somewhat more difficult to learn material that is locally but not globally consistent with it; and it should be most difficult to learn material that is neither locally nor globally consistent with it. As illustrated below, this approach can provide a basis for precise predictions concerning learning, and thus for rigorous tests of the initial knowledge assessment.

The domain to which this approach is applied is knowledge of U.S. presidents. Experiment 1 provides an assessment of subjective organization of the presidents. Experiment 2 tests predictions derived from the Experiment 1 knowledge assessment concerning ability to learn new organizations. The goal is to provide a generally useful approach for identifying and testing hypothesized organizations, as well as to add to understanding of how people organize the presidencies.

## EXPERIMENT 1

For most U.S. citizens, the president of the United States is a symbol of his times and a personification of the federal government. During his term of office, the public learns a great deal about the president's actions in the national and international spheres, as well as about his personality and social life (Kinder & Fiske, 1986). Because of the importance of the presidency, because there is always one and only one president, and because each presidency occupies a distinct time



period, it seems possible that people organize much of their knowledge of historical events around presidential administrations.

In the only published study that explicitly examined knowledge of U.S. presidents (Roediger & Crowder, 1976; also see, Crowder, 1976, p. 459), subjects were asked to recall the names and ordinal positions of all of the presidents. A pattern much like a serial position curve was found, with the earliest and most recent presidents being the most likely to be recalled. One explanation for this pattern of recall (though not the one presented by Roediger and Crowder) is that people organize the presidents into three distinct groups: a small group of early presidents, a small group of recent presidents, and a large group of presidents not included in the other two groups. When a subject knows that a particular president belongs to either the early group or the recent group, this knowledge can be used to considerably narrow the possible dates of the president's term in office. The positional constraints provided by the knowledge that a president is a member of the middle group are much weaker. Hence, such an organization would lead people to more accurately assign serial positions to the terms of early and recent presidents than to those of middle-period ones.

The hypothesis that people divide presidential terms into distinct subcategories is consistent with findings on how people represent other types of quantitative information. For example, Pliske and Smith (1979) presented subjects with a 12-term linear sequence, in which half of the quantities were associated with male names and half with female names. In the blocked condition, the quantities associated with one sex were all larger than those associated with the other sex; in the random condition, there was no such relation. In both cases, subjects were presented with pairs of names and had to decide as quickly as possible which of the two names was associated with the larger quantity. Subjects in the random group displayed the usual symbolic distance and serial position effects (Banks, 1977; Moyer & Dumals, 1978) on all items. Response time (RT) decreased as distance between items increased, and RT increased as items' distance from the ends of the list increased. In contrast, subjects in the blocked-presentation condition showed distance and serial position effects only for within-block trials. When items were drawn from different groups (e.g., Anne-Mike), responses were uniformly fast; neither interitem distance nor distance from an endpoint affected performance.

Maki (1981) obtained comparable results in a study of geographical knowledge. She asked students at a North Dakota University which of two midwestern cities was farther east or west (e.g., "Is Fargo east of Hibbing?"). When the cities were located in the same state (Hibbing-

Duluth), standard distance and serial position effects were obtained. However, when cities were drawn from different states (Fargo-Hibbing), there was little evidence for either effect.

Both Pliske and Smith (1979) and Maki (1981) suggested that differences in performance on within-category and between-category comparisons reflect the use of a dual-process approach. One process involves accessing information about each item's category membership (e.g., gender, state) and comparing the retrieved categories. If the categories differ, subjects can respond on the basis of well-known, relatively general information (e.g., Minnesota is east of North Dakota). The other process comes into play when both items are from the same category. Under such conditions, subjects must rely on more specific and less well-known information (e.g., whether Duluth is east of Hibbing). This process will often require an "analogue" computation (e.g., comparison of the two cities' locations on an imagined map). It is during this latter, analogue comparison that distance and serial position effects arise. Similar dual-process models have been proposed by Kosslyn, Murphy, Bemesderfer, and Feinstein (1977) to account for judgments of order among artificially grouped stimuli, by Hamilton and Sanford (1978) to account for order judgments among letters of the alphabet, by Siegler and Robinson (1982) to account for preschoolers' numerical magnitude comparisons, and by Wyer, Shoben, Furhman, and Bodenhausen (1985) to account for temporal judgments about events.

If people subjectively organize presidents into distinct categories, they may also use this type of dual-process approach to compare when different presidents served. Suppose that people organize U.S. presidents into an early group, a middle group, and a recent group, and that subjects have ready access to this group information. Such an organization and process would allow rapid categorical decisions on trials that involve presidents drawn from different groups. For example, a subject presented the Jefferson-Van Buren pair might realize that Jefferson was an early president and Van Buren was not, and conclude that Jefferson must have held office earlier than Van Buren.

When both presidents are from the same group, subjects would need to decide differently. One approach would be to retrieve temporally informative facts (N. Brown, 1990; N. Brown, Rips, & Shevell, 1985; Friedman & Wilkins, 1985) associated with the two presidents and use them to infer who served first. For example, a subject trying to decide whether Theodore Roosevelt held office before Hoover might recall that Roosevelt charged up San Juan Hill during the Spanish American War, and that Hoover was president when the stock market crashed in 1929. Given the additional knowledge that the Spanish-American War took place at the end of the nineteenth century, the subject could infer that Roosevelt held office before Hoover.



If people divide presidents into such groups, and use a dual-process approach to decide who served earlier, several empirical consequences should follow. Temporal proximity should influence difficulty of within-group temporal comparisons, but should exert little if any influence on between-group comparisons. Other things being equal, it should be more difficult to judge which presidency was earlier when the presidents are within the same subjectively defined group and served close in time (e.g., Wilson and Taft) than when they are within the same group but served farther apart in time (e.g., Polk and Wilson). In contrast, the difficulty of a between-group trial should not be affected by the distance separating the presidents in the comparison because such trials can be solved simply by retrieving and comparing group memberships.

The dual-process model also predicts that within-group comparisons will be more affected than between-group comparisons by "positional uncertainty." On within-group trials, the more sure a subject is of each item's position within the list, the less time should be needed for the comparison. On between-group comparisons, uncertainty about the particular position should have little effect on comparison times, because comparisons are made on the basis of category membership, rather than position within the overall set.

The concept of positional uncertainty has often been used to explain serial position effects obtained on quantitative comparison (Holyoak & Patterson, 1981) and serial recall (Estes, Allmeyer, & Reder, 1976) tasks. With most stimuli, serial position and positional uncertainty are completely confounded. This is not the case for the presidents, though; people can locate the terms of some central presidents (e.g., Lincoln) with much greater accuracy than they can locate the terms of some peripheral presidents (e.g., Monroe). Because serial position effects are assumed to reflect positional uncertainty, rather than serial position per se, and because these variables are partially unconfounded for this set of stimuli, we couch most of our discussion and analyses of serial-positionlike phenomena in terms of the more basic concept, positional uncertainty.

To summarize, we hypothesized that people organize U.S. presidents into three distinct groups—a relatively small Founding Fathers group, a large middle group, and a relatively small recent group—and that they use this organization to compare the order in which presidents served. This view implied that speed and accuracy should be discontinuously greater for presidents in the two small groups than for those in the larger middle-period group. Presidents in the two

small groups would participate in fewer of the difficult within-group comparisons, and more of the easy between-group comparisons. The hypothesized organization also implied that temporal distance and positional uncertainty would exert a greater influence on within-group comparisons than on between-group comparisons. If people divide presidents into temporally ordered groups, membership in different groups provides a straightforward criterion for making between-group comparisons, regardless of the distance between presidents or positional uncertainty of the individual presidents involved in the comparison. When presidents belong to the same group, however, distance between their terms and uncertainty about the exact position of each president should more strongly influence temporal comparisons.

## METHOD

### Subjects

Participants were 20 right-handed subjects between the ages of 18 to 30 years. All were native-born U.S. citizens. Subjects were recruited from the University of Chicago community through an advertisement placed in the school newspaper. Each subject was tested individually and was paid $3.00 for a session that lasted approximately 50 min.

### Materials

The 39 presidents were divided into two sets: an odd-term set [Washington (1), Jefferson (3), . . . , Reagan (39)] and an even-term set [Adams (2), Madison (4), . . . , Carter (38)]. Each president in the odd-term set was paired with each of the other 19 odd-term presidents, and each even-term president was paired with each of the other 18 even-term presidents. In this way, 361 comparison pairs were formed, 190 from the odd-term set, and 171 from the even-term set.

### Procedure

All stimuli were presented on a computer screen (CRT) controlled by a PDP-11/23. Subjects sat before the screen with their left index finger on a *ready* button on the left and their right index finger on the middle (*home*) button of a three-button group on the right. All responses were made with the right index finger. Subjects pressed the button immediately to the left of the home button when they believed the president whose name appeared on the left side of the screen held office earlier, and pressed the button immediately to the right of the home button when they believed that the president on the right side served earlier.

To initiate a trial, subjects pressed the ready button. Once they did so, a fixation word (*Ready*) appeared on the screen for 0.5 s. It then was replaced by the first and last names of the two presidents in the comparison. The names of the two presidents appeared on the same line of the screen; the first name of the leftmost president began where the fixation word had begun. The names remained in view until the subject responded. After the response, the screen was erased, and the subject was allowed to initiate the next trial. No feedback was given.



Before the experimental trials began, subjects received 28 nonspecific practice trials. The sequence of events for a practice trial was identical to the sequence that the subjects would see in the actual experiment, except for the specific words on the screen. Instead of the names of two presidents appearing, a single word, either *Left* or *Right* appeared in the center of the screen. Subjects were instructed to press the button to the left of the home button when the word *Left* appeared, and to press the button to the right of the home button when the word *Right* appeared.

Then the experimental phase was presented. The task was to decide which president in the comparison pair held office earlier. Instructions stressed the importance of accuracy. The presentation order of the 361 comparison pairs was randomized separately for each subject. On each trial, the presidential name that appeared on the left side of the screen was selected randomly.

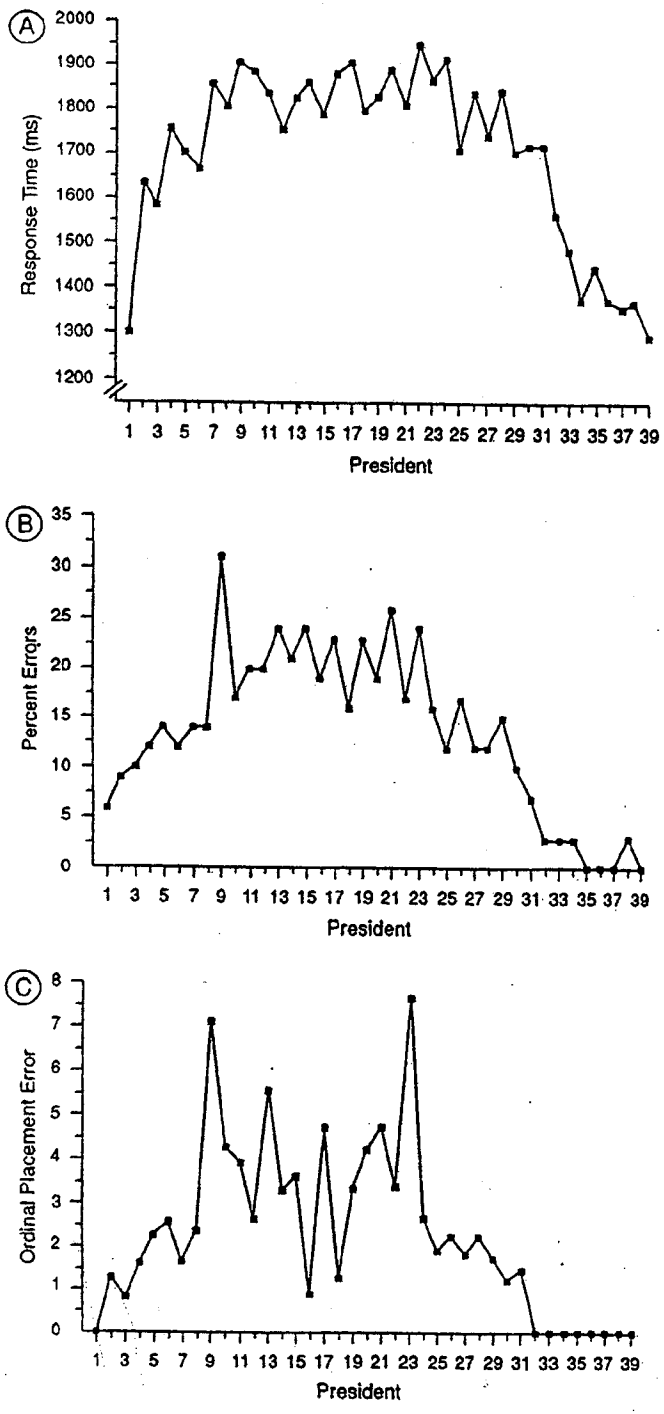
On completion of the paired-comparison task, subjects completed an ordinal placement booklet. This booklet contained a randomly ordered list of the 39 presidents. Subjects were instructed to rank order the terms of the presidents by assigning each president a number between 1 and 39, with each number to be used once.

## RESULTS AND DISCUSSION

The data analyses can be divided into two groups. One group of analyses involved "president mean" measures, that is, measures of each president's mean solution time, percent errors, and absolute ordinal placement error. The other group of analyses involved "comparison mean" measures, that is, measures of performance on each of the 361 paired comparisons.

### President means

Three measures were computed for each president. Two of them, mean RT and percent errors, were based on performance on the comparison task. Mean RT for each president was the mean of individual subjects' RTs for comparisons involving that president. The RTs used in this analysis excluded the 14% of trials where subjects answered incorrectly and the 2% where they answered correctly but took more than 6.0 s to do so. Mean percent errors for each president was the mean of individual subjects' percent errors on comparisons involving that president. The third measure, mean ordinal placement error, was the average absolute difference on the ordinal placement task between the numerical rank subjects assigned to a president's term and the actual rank of the term.






Mean RT, percent errors, and mean ordinal placement error for each president are displayed in Figures 1A, 1B, and 1C. Separate analyses of variance (ANOVAs) were computed for each of the three measures. In each analysis, the 39 presidents served as a fixed effect, and the 20 subjects as a random effect. The effect of president was highly significant for RT, $F(38, 722) = 17.75, p < .001$; percent errors, $F(38, 722) = 18.25, p < .001$; and ordinal placement error, $F(38, 722) = 10.43, p < .001$.

Earlier, it was argued that discontinuities between presidents on such measures could indicate subjectively defined category boundaries. In particular, because the Founding Fathers and recent president groups were expected to be smaller than the middle president group, mean speed and accuracy for presidents in those groups were predicted to be considerably greater than for middle period presidents. By this reasoning, the means in Figure 1 suggest that participants subjectively organized the presidents into three groups: a *Founding Fathers* group (the 6 presidents Washington through John Quincy Adams), a middle group of *noncontemporary* presidents (the 26 presidents Jackson through Franklin Roosevelt), and a group of *post–World War II* presidents (the 8 presidents Truman through Reagan).

The means shown in Figure 1 provide at least three reasons for believing that the post–World War II presidents form a distinct set. First, mean RTs for all 8 post–World War II presidents were faster than those for any of the other 31 presidents except Washington. Second, error rates for each of these 8 presidents were no more than half as high as those for any of the other 31 presidents. Third, there

Figure 1. Mean correct response time (A), percent errors (B), and ordinal placement error (C) for the presidents (Experiment 1 data). President names correspond to the following numbers:

| | | |
|---|---|---|
| 1. George Washington | 14. Franklin Pierce | 27. Woodrow Wilson |
| 2. John Adams | 15. James Buchanan | 28. Warren Harding |
| 3. Thomas Jefferson | 16. Abraham Lincoln | 29. Calvin Coolidge |
| 4. James Madison | 17. Andrew Johnson | 30. Herbert Hoover |
| 5. James Monroe | 18. Ulysses Grant | 31. Franklin Roosevelt |
| 6. John Q. Adams | 19. Rutherford Hayes | 32. Harry Truman |
| 7. Andrew Jackson | 20. James Garfield | 33. Dwight Eisenhower |
| 8. Martin Van Buren | 21. Chester Arthur | 34. John Kennedy |
| 9. William Harrison | 22. Grover Cleveland | 35. Lyndon Johnson |
| 10. John Tyler | 23. Benjamin Harrison | 36. Richard Nixon |
| 11. James Polk | 24. William McKinley | 37. Gerald Ford |
| 12. Zachary Taylor | 25. Theodore Roosevelt | 38. James Carter |
| 13. Millard Fillmore | 26. William Taft | 39. Ronald Reagan |

was a clear group boundary: Trials that involved the 32d president (Truman) were 155 ms faster than those that involved the 31st president (Franklin Roosevelt), and error rates were less than half as high.

A similar, though less striking, pattern led to the conclusion that subjects also distinguished between Founding Fathers and later presidents. First, RTs associated with all six Founding Fathers were faster than those associated with any president who served later in the nineteenth century; RTs for all Founding Fathers other than Madison were faster than those for any middle period president. Second, error rates associated with each of the Founding Fathers were as low as or lower than those obtained for any nineteenth-century president. Third, there seemed to be a discontinuity in RTs, with times for the 6th president (John Quincy Adams) being almost 200 ms faster than those for the 7th (Andrew Jackson).

## Comparison means

As noted above, the president means suggested that subjects partitioned the presidents into three groups: Founding Fathers (Washington through John Quincy Adams), noncontemporary presidents (Jackson through Franklin Roosevelt), and post–World War II presidents (Truman through Reagan). In this section, we first provide additional evidence that the boundaries between these groups were correctly located. We then demonstrate that a dual-process model that assumes this particular three-group organization accounts for more variance than either a single-process model, or a dual-process model that assumes an alternative three-group organization. Both analyses make use of the RT and error means computed for each of the 361 comparison pairs. Because all 20 subjects saw each comparison trial once, the maximum number of observations per mean is 20. Again, trials on which answers were incorrect or where the RT was greater than 6.0 s were excluded from the RT computations.

In the current analysis, we assume that the presidents are divided into three groups and that comparison judgments are made using a dual-decision process. Given these assumptions, it is possible to select the optimum three-group organization by comparing the difference between the average within-group trial and the average between-group trial across a wide variety of groupings. If discrete groups exist, boundaries between groups can be either correctly or incorrectly identified. If they are correctly identified, all trials that allowed subjects to use the fast categorical response process will be classified as between-group trials, and all trials that required subjects to use the slower analogue response mode will be classified as within-group trials. In contrast, if group boundaries are incorrectly assigned, some of the



trials that were performed using the fast categorical process will be misclassified as slow within-group trials, some of the trials performed using the slow analogue process will be misclassified as fast between-group trials, or both could occur. This means that other things being equal, the difference between the average within-group trial and the average between-group trial should be largest when group boundaries are correctly identified, and that this difference should diminish as assigned boundaries diverge from the true boundaries. Thus, if the discontinuities we located in the set of presidents' means accurately reflect breaks between subjective groups, this within- versus between-group difference measure should be greater for the 6-25-8 organization than for other three-group organizations.

To test this prediction, 120 three-group organizations were created. This was done by varying the locations of the top bound of the early group and the lower bound of the recent group in a factorial manner. The top bound of the early group ranged from Presidents 4 to 13, and the lower bound of the recent group ranged from Presidents 26 to 37. By selecting these particular limits for the two sets of group boundaries, we deliberately created end-groups that included at least 4 and no more than 13 presidents. Four was set as the minimum group size because anything less would have granted undue influence to the first and last presidents. Thirteen was set as the maximum size for an end-group because we were interested in isolating an early group boundary that we believed fell among the first third of the presidents and a recent group boundary that we believed fell among the most recent third of the presidents.

For a given organization, the difference between within-group trials and between-group trials was computed in the following manner: First, each of the 361 comparisons was classified as being either within-group or between-group relative to that organization's assigned boundaries. These meant that many comparison pairs were classified as being within-group trials for some organizations and between-group trials for others. For instance, the comparison between President 5 and President 7 was classified as a between-group trial when the top bound of the early group was 5 or 6, and as a within-group trial under all other circumstances. After all trials were assigned to trial type, within-group and between-group means were computed for that organization, and the mean of the between-group comparisons was subtracted from the mean of the within-group comparisons.

These within-group versus between-group differences are presented in Figure 2A and 2B. This figure identifies the ordinal position of the earliest president in the "recent presidents" group along the x-axis. Each curve within the figure connects points computed using

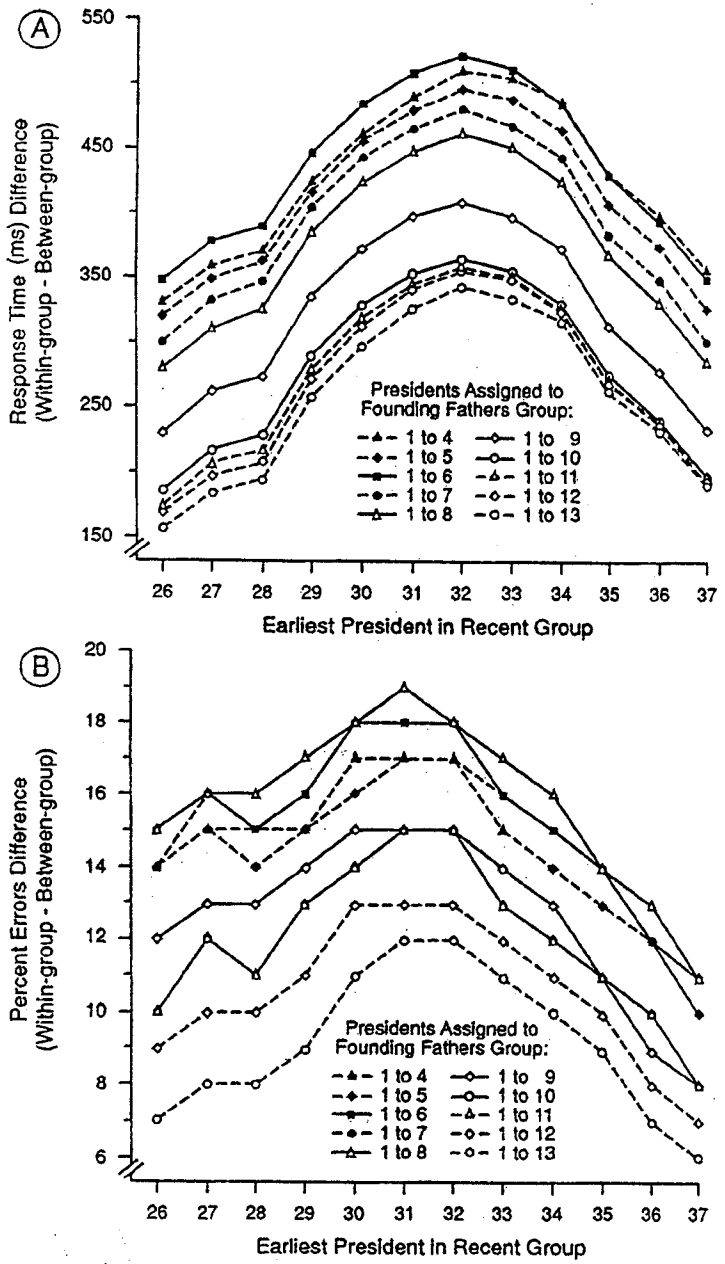



Figure 2. Difference between average within-group pair and average between-group pair for (A) correct response time and (B) percent errors (Experiment 1 data)



the same "Founding Fathers" group. For example, the leftmost point on the bottom curve of Figure 2A represents the difference when the Founding Fathers group was defined as the first 13 presidents and the recent presidents group extended from the 26th president onward. The next point on the same curve represents the difference computed when the Founding Fathers group was defined as the first 13 presidents and the earliest member of the recent presidents group was the 27th president.

The most important thing to notice about Figure 2 is that the largest RT difference occurs when the Founding Fathers group included the first 6 presidents and the recent presidents group extended from the 32d president onward (i.e., when trials were classified in accord with our hypothesized 6-25-8 organization). It should also be noted that only 1 of the 120 alternative groupings (Founding Fathers group includes first 8 presidents; recent presidents include most recent 9) displayed a greater difference on the percent errors measure. Although this fact might argue for an 8-22-9 organization of presidents, one must also consider that percent errors for the two groupings were very similar (18% and 19%). In contrast, the RT difference computed for this alternative organization was substantially smaller than for the hypothesized one (448 vs. 522 ms). Moreover, 20 organizations yielded greater RT differences than the one computed for the 8-22-9 organization. Thus, when errors and solution times are considered together, the data quite strongly supported our hypothesis that there are group boundaries between John Quincy Adams (President 6) and Jackson (President 7), and between Franklin Roosevelt (President 31) and Truman (President 32).

In the preceding analysis, it was assumed that a dual-decision process played a central role in our subjects' comparisons. Given this assumption, it was possible to obtain evidence that supported a particular organizational scheme. In the current analysis, we adopted the opposite strategy; we made assumptions about the organization of presidents and attempted to demonstrate that our dual-process model is correct. Specifically, we assumed that presidents were grouped into a 6-25-8 organization and divided the 361 comparison pairs into 162 within-group comparisons and 199 between-group comparisons. We then fit RT and percent errors means by a multiple regression model that included six variables:

1. a general intercept term
2. a general distance variable (with distance defined as number of terms separating the two presidencies in the comparison)
3. a general positional uncertainty variable (with positional uncer-

tainty defined as the average ordinal placement error for the two presidents in the comparison)

4. a group variable (defined as *1* for within-group trials and *0* for between-group trials) to reflect intercept differences separating within-group and between-group trials

5. a variable reflecting an additional effect of distance on within-group trials

6. a variable reflecting an additional effect of positional uncertainty on within-group trials.

Note that for within-group trials, the parameter estimate for distance is the sum of variables (2) and (5), and the estimate for positional uncertainty is the sum of variables (3) and (6).

The key predictions of the model involved variables (4), (5), and (6). If between-group and within-group comparisons are done differently, then whether the comparison is between or within groups should exert an effect beyond the general effect of distance and positional uncertainty. Within the regression equations, this means that adding the group variable (4) to an equation already including distance and positional uncertainty parameters should significantly increase the amount of variance that can be explained. The dual-process model also predicts that distance and positional uncertainty should have a greater effect on within-group trials than on between-group trials. This implies that the parameter estimate for within-group distance (5) should be greater than that for the general distance variable (2), and that the effect of within-group uncertainty (6) should be greater than that for the general positional uncertainty variable (3).

The regression analyses performed to test these predictions involved testing whether each variable could add significant variance to that which could be accounted for by the other five variables. Table 1 illustrates the results of the analyses. Within this table, the hypothesized subjective organization is labeled the 6-25-8 organization.

Overall, the regression based on this 6-25-8 organization accounted for 59% of the variance in both RT and error means for the 361 comparisons. As predicted, whether a comparison was between or within groups exerted a highly significant effect in both analyses, above and beyond that exerted by the general distance and uncertainty variables. Also as predicted, in both RT and error analyses, the within-group distance effect was considerably larger than the general distance effect, and the within-group positional uncertainty effect was considerably larger than the general uncertainty effect.

Table 1 also illustrates results of two other regression analyses that were conducted for comparison. The "one group" model differed from the original 6-25-8 model in both processing and representa-



Table 1. Response time (RT) and percent errors parameter estimates from 6-25-8, 13-13-13, and one-group-model regression analyses (Experiment 1 data)

| Variable | RT(ms) | | | Percent errors | | | |
|---|---|---|---|---|---|---|---|
| | 6-25-8 | 13-13-13 | One group | 6-25-8 | 13-13-13 | One group |
| General intercept | 1551*** | 1824*** | 1703*** | .051* | .029*** | .096*** |
| Distance | −9*** | −24*** | −19*** | −.003** | −.007*** | .007*** |
| Uncertainty | 25** | 45*** | 58*** | .013*** | .022*** | .027*** |
| Type of comparison (Within vs. between group) | 305*** | −51 | | .104*** | .000 | |
| Within-group distance | −19*** | −26** | | −.014*** | −.012** | |
| Within-group uncertainty | 51*** | 22 | | .032*** | .012** | |
| Overall $R^2$ | .59 | .48 | .46 | .59 | .51 | .48 |

* $p < .05$. ** $p < .01$. *** $p < .001$.

tional assumptions. This model took as its starting point the parsimonious assumption that people represent their knowledge of the presidents in a single list without subcategories and respond entirely on the basis of an analogue comparison process. The regression equation devised to express this one-group model included an intercept term and general distance and uncertainty terms [variables (1), (2), (3)].

The other alternative model, the 13-13-13 model, provided a control for the possibility that any superiority in percent variance accounted for by the hypothesized subjective organization relative to the one-group model was due to differences in number of parameters in the regression equations. In the 13-13-13 model, the first 13 presidents were assigned to the early presidents group, the next 13 to the middle presidents group, and the most recent 13 to the recent presidents group. The regression equation used to represent this model included the same six variables that were used to represent the hypothesized subjective organization. The two models differed only in the boundaries between groups, and therefore in the particular comparisons classified as within group and between group. This particular tripartite division was selected as a control because it seemed the simplest, least arbitrary point of comparison.

As shown in the last line of Table 1, the hypothesized subjective organization accounted for considerably more variance in both RT and error data than either of the alternative models. The fit obtained for the 13-13-13 model was little better than the one obtained for the one-group model. Therefore, the greater variance in the 361 comparison means accounted for by the model based on the hypothesized subjective organization cannot be attributed to the presence of three additional parameters. Rather, the greater variance accounted for is due to the model's particular classifications of comparisons as being within-groups or between-groups.

Taken together, the analyses presented above suggest that the subjects in this experiment used a dual-decision process when they made their comparison judgments and that the knowledge that supported this process was organized in a particular manner. However, one still might object that these results do not indicate the presence of specific preexisting groupings in long-term memory. Rather, it might be argued that subjects invented the 6-25-8 organization early in this experiment to make the comparison task more tractable. We address this objection in the next experiment by demonstrating that the preexisting organization can have either positive or negative effects on learning, depending on the fit between it and a new organization.



## EXPERIMENT 2

The view that existing knowledge greatly influences acquisition of new knowledge is widely accepted. Applied to the present context, the view suggests that the hypothesized organization of the presidents should provide a basis for predicting relative difficulty of acquiring new information about them. Accurate prediction of learning would also demonstrate that the hypothesized subjective organization is influential beyond the comparison task where it was originally hypothesized. Further, such a demonstration would illustrate the feasibility of a technique that seems potentially useful for testing a wide range of hypothesized organizations, not just ones involving the presidents. Experiment 2 demonstrates how this approach can be realized.

The experiment involved successive presentation of four lists of the presidents. Each list included the names of all 39 presidents, listed from top to bottom in chronological order and divided into groups of *early, middle,* and *recent* presidents. The experimental manipulation that differentiated the lists was the exact placement of boundaries between early and middle, and between middle and recent presidents. Subjects were asked to learn the list-defined classification of each president during a 1-min study period and then were tested on whether each president was classified as early, middle, or recent on the just-studied list.

If people organize presidential information as hypothesized, relative difficulty of acquiring new information should be predictable from the fit between this new information and existing knowledge. Three types of cases, defined in terms of relations between preexisting and list-defined classifications, seemed likely to be of differential difficulty. *Item-and-category-consistent* classifications seemed likely to be easiest to learn. These were cases where the list-defined category matched the hypothesized preexisting category both for the particular president and for all of the other presidents in that president's hypothesized preexisting category. *Item-consistent* classifications were expected to be of intermediate difficulty. These were cases where the list-defined category matched the hypothesized preexisting category for the particular president but not for all of the other members of that preexisting category. Finally, *inconsistent* classifications were expected to be the most difficult. These were cases where the list-defined category did not match the hypothesized preexisting category either of the particular president or (by necessity) of all members of the hypothesized preexisting group.

The relative difficulty of these three cases can be viewed in terms of three sources of information that subjects might use when asked

to recall a given president's classification on a given list. One cue is provided by episodic memory of the president's classification on the particular list; a second cue is provided by the long-term classification of the president in semantic memory; a third cue is provided by the knowledge that the first and second cues are redundant for presidents within a particular list.

Presumably, the more cues pointing to the correct answer, the easier it will be to state that answer. On item-and-category-consistent problems, all three cues point to the correct answer. For example, if Jefferson and all of the other first six presidents were classified as early presidents on a list, the memory of how Jefferson was classified on the list, the semantic memory representation of Jefferson as an early president, and the recognition that all presidents classified as early in semantic memory also were classified as early presidents on the list would all point to the correct classification of Jefferson as an early president.

The case is somewhat different with item-consistent classifications, such as would occur if Jefferson were classified as an early president, but Monroe and John Quincy Adams were not. On such an item, the memory of the item's classification on the list and the semantic memory representation would point to classifying Jefferson as an early president. However, there would be no general redundancy between list-defined and semantically defined classifications to indicate unambiguously the list-defined classification if the list-defined classification could not be directly retrieved.

Finally, if Jefferson were classified on a list as a middle-period president (an inconsistent classification), only the episodic memory of the list would support the classification. Semantic memory would suggest that Jefferson was an early president, and there would be no consistent treatment of early presidents to resolve the conflict. Thus, such inconsistent classifications were expected to be most difficult to learn.

Four lists were designed in such a way that the difficulty of learning classifications of particular presidents would depend on the interaction between the hypothesized preexisting organization and the classification within the particular list. Figure 3 illustrates this interaction in the context of Lists 1 and 4. In List 1, the first 6 and the last 8 presidents should be easiest, because their entire group is assigned to the same category in the list as in their preexisting organization. In contrast, in List 4, the 25 middle presidents should be easiest, because their entire group is assigned to the same category in the preexisting and list-defined organizations. The List × President interaction also was expected to influence which list-defined classifications would be most difficult to learn. In List 1, the hardest items to learn should be



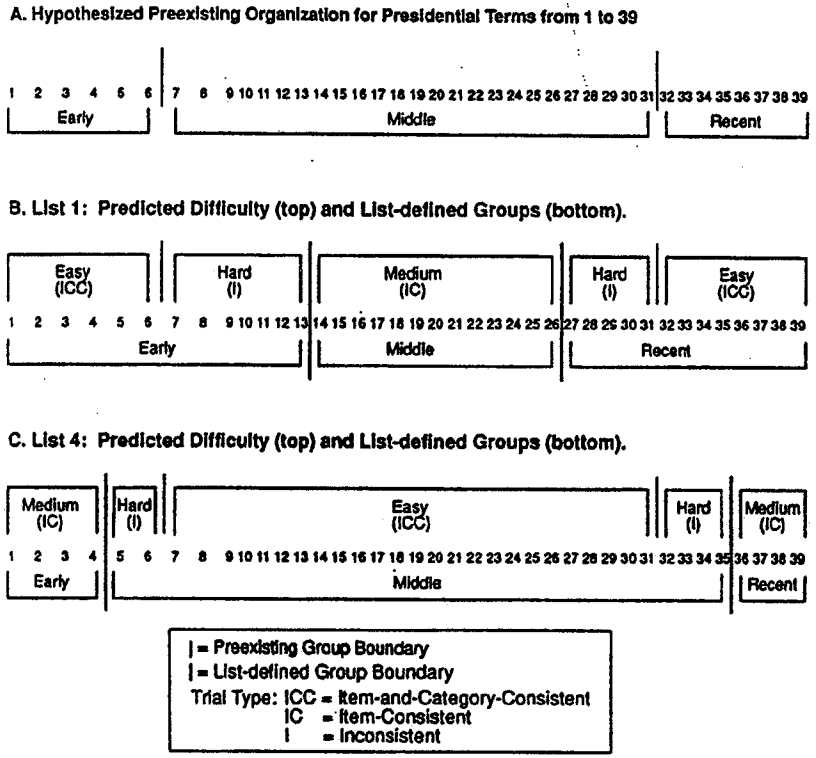



Figure 3. Hypothesized subjective organization of presidential categories (A) and predicted performance for List 1 (B) and List 4 (C) in Experiment 2

the 7th through 13th and the 27th through 31st presidents. Each of these presidents is assigned to a different category in the list than in the hypothesized preexisting organization. In contrast, in List 4, the hardest items should be the 5th and 6th and the 32d through 35th presidents, because these require inconsistent classifications on that list. Thus, just as there is no overlap in which items are predicted to be easiest within the two lists, there is no overlap in which items are predicted to be hardest.

Although these predictions follow logically from the hypothesized subjective organization of the presidents, there is little direct evidence that prior organization of information in long-term memory can have a negative effect on the acquisition of competing organizations of that information (e.g., Ekstrand & Underwood, 1965; Postman, 1963; Underwood & Postman, 1960). Since such an effect was presupposed

in the design of this experiment, results consistent with our predictions can be taken as evidence for the existence of extraexperimental transfer, and hence for the existence of a continuity between episodic and semantic memory (cf. Anderson & Ross, 1980; McKoon, Ratcliff, & Dell, 1986; Tulving, 1972, 1983).

## METHOD

### Subjects

Participants were 20 university students drawn from the same subject pool as those who participated in Experiment 1. They were tested individually and were paid $3.00 for the approximately 1-hr session.

### Lists

Each subject received the same four lists of presidents. Within each list names of the 39 presidents were printed in chronological order, one to a line, on the left side of the page. Two horizontal lines divided the 39 presidents into 3 groups; the upper horizontal line separated early from middle-period presidents, and the lower horizontal line separated middle-period from recent presidents. Names of the three groups, *early*, *middle*, and *recent*, were typed on the right side of the page, opposite the group they labeled.

The only difference between the four lists was the location of the group boundaries. List 1 divided the 39 presidents into 3 groups of 13. Thus, the break between early and middle groups fell between Fillmore and Pierce, and the break between middle and recent groups fell between Taft and Wilson. List 2 placed the first 10 presidents (Washington through Tyler) in the early group, the next 19 (Polk through Coolidge) in the middle group, and the most contemporary 10 (Hoover through Reagan) in the recent group. List 3 boundaries were selected to coincide with those of the preexisting categories inferred from Experiment 1. Thus, the first 6 presidents (Washington through John Quincy Adams) were in the early group, the next 25 presidents (Jackson through Franklin Roosevelt) were in the middle group, and the most recent 8 presidents (Truman through Reagan) were in the final group. Finally, in List 4, the first 4 presidents (Washington through Madison) were in the early group, the next 31 (Monroe through Lyndon Johnson) were in the middle group, and the latest 4 (Nixon through Reagan) were in the recent group. These lists were chosen for their differing relations to the hypothesized preexisting organization, as discussed for Lists 1 and 4 in the introduction to this experiment.

### Procedure

The same apparatus was used in this experiment as in Experiment 1. Stimuli were presented on a video display controlled by a PDP-11/23 computer. Responses were made on a four-button response panel. The button isolated on the left side of the panel was used by subjects to initiate each trial; the



three buttons on the right side were used to make responses. On each trial, subjects were to respond by pressing the button labeled *E* if the president was labeled *early* in the most recently studied list; by pressing *M* if the president was labeled *middle* in the list; or by pressing *R* if the president was labeled *recent*. The location of the labels on the buttons was counterbalanced over subjects, so that half of the subjects pressed the rightmost of the three buttons to indicate a recent president, and the other half pressed the leftmost of the three to do so.

At the beginning of the experiment, subjects were given a 33-trial familiarization period to become acquainted with the apparatus. Subjects initiated each trial by pressing the *ready* button with their left index finger. After the button press, a fixation word (*Ready*) appeared on the screen for 0.5 s, and then was replaced by *Early*, *Middle*, or *Recent*. The task was to press as quickly as possible the response button that corresponded to the word displayed on the screen. The word remained on the screen until the subject responded. After the word disappeared, the subject could initiate the next trial by pressing the ready button.

The experiment itself was organized into four trial blocks, each trial block involving a particular list. At the beginning of a trial block, the subject received one of the four lists, and was instructed to study the list carefully, concentrating on the group membership of each president rather than on the exact order. Then the subject was given a 1-min study period, followed by the test for that list.

The test for each list (each trial block) included three cycles through the 39 presidents. Within each cycle, each president was presented once. Order of presentation was randomized separately for each of the three cycles through the 39 presidents. No feedback was given after individual trials. However, between cycles, subjects received a message informing them of how many errors they had made in the just-completed cycle.

Each subject studied and was tested on each list. Five separate Latin squares were used to assign lists to the first, second, third, or fourth trial block. As a result, each list appeared an equal number of times in each of the four trial blocks.

## RESULTS AND DISCUSSION

Mean correct RT and percent errors were computed for each president within each list for each subject. Because a subject's knowledge of the classifications in a list was tested three times (the three cycles through the list within each trial block), these means were calculated from three or fewer observations. When a subject erred on all three trials, mean RT for that cell was estimated from the appropriate president, list, and subject means. In all, 3% of List × President × Subject cell means were estimated in this manner.

Separate 4 (List) × 39 (President) ANOVAs were calculated for the RT and error data, with list and president serving as fixed effects

and subjects serving as a random effect. The effect of president was significant at the .001 level for both RT, $F(38, 722) = 5.93$, and percent errors, $F(38, 722) = 6.44$, as was the effect of list [for RT, $F(3, 57) = 21.11$; for percent errors, $F(3, 57) = 28.93$], and the List × President interaction [for RT, $F(114, 2166) = 4.46$; for percent errors, $F(114, 2166) = 5.68$].

The major prediction of the experiment concerned the List × President interaction. The ANOVAs established that, as predicted, this interaction was highly significant for both RTs and errors. To test whether the interaction was of the anticipated form, we compared the three types of items that were predicted to differ in difficulty. Recall that item-and-category-consistent classifications were expected to be easiest; that item-consistent classifications were expected to be of intermediate difficulty; and that inconsistent classifications were expected to be the most difficult.

As shown in Table 2, the results consistently matched these predictions. Averaged across all lists and presidents, item-and-category-consistent classifications were 328 ms faster than item-consistent classifications, which in turn were 140 ms faster than inconsistent

Table 2. Performance on item-and-category-consistent (ICC), item-consistent (IC), and inconsistent (I) trials for each list (Experiment 2 data)

| | | Type of classification | | | |
|---|---|---|---|---|---|
| List | Measure | ICC | IC | I | Mean |
| 1 (13-13-13) | RT | 998 | 1439 | 1485 | 1293 |
| | Error | .02 | .18 | .32 | .16 |
| | *n* | 14 | 13 | 12 | 39 |
| 2 (10-19-10) | RT | 969 | 1260 | 1540 | 1198 |
| | Error | .02 | .10 | .24 | .09 |
| | *n* | 14 | 19 | 6 | 39 |
| 3 (6-25-8) | RT | 1010 | | | 1010 |
| | Error | .03 | | | .03 |
| | *n* | 39 | | | 39 |
| 4 (4-31-4) | RT | 833 | 1082 | 1180 | 938 |
| | Error | .00 | .08 | .10 | .03 |
| | *n* | 25 | 8 | 6 | 39 |
| Mean | RT | 954 | 1282 | 1422 | 1110 |
| | Error | .02 | .11 | .24 | .08 |
| | *n* | 92 | 40 | 24 | 156 |

*Note.* *n* = Number of presidents receiving each type of classification on each list.



classifications. Similarly, percent errors was lower on item-and-category-consistent classifications (2%) than on item-consistent classifications (11%), and lower on item-consistent classifications than on inconsistent classifications (24%). Planned comparisons indicated that significant differences ($p < .001$) separated performance on the 92 item-and-category-consistent classifications from that on the 40 item-consistent classifications [for RT, $F(1, 2166) = 279.96$; for percent errors, $F(1, 2144) = 27.46$]. Significant differences also separated performance on the 40 item-consistent classifications from that on the 24 inconsistent classifications [for RT, $F(1, 2144) = 136.45$; for percent errors, $F(1, 2144) = 153.67$].

These differences also were apparent when performance was examined within each list (Table 2). The three types of classifications could be compared on Lists 1, 2, and 4 (List 3 classifications were identical to the hypothesized preexisting organization, so all items were item-and-category-consistent classifications). Performance on item-and-category-consistent classifications was superior to that on item-consistent classifications on all three lists on both RT and accuracy measures [for RT: on List 1, $F(1, 722) = 72.17$; on List 2, $F(1, 722) = 39.54$; on List 4, $F(1, 722) = 68.71$; for percent errors: on List 1, $F(1, 722) = 53.22$; on List 2, $F(1, 722) = 22.92$; on List 4, $F(1, 722) = 64.60$; in all cases, $p < .001$]. Similarly, differences between item-consistent and inconsistent classifications within each list were all in the predicted direction, and four of the six differences were significant [for RT: on List 1, $F(1, 722) < 1$; on List 2, $F(1, 722) = 20.71$, $p < .001$; on List 4, $F(1, 722) = 6.04$, $p < .05$; for percent errors: on List 1, $F(1, 722) = 37.72$, $p < .001$; on List 2, $F(1, 722) = 39.73$, $p < .001$; on List 4, $F(1, 722) = 2.69$, $p > .1$].

These results followed the predicted pattern. Three alternative explanations needed to be addressed, however, before we could confidently conclude that the observed performance pattern reflected the influence of the hypothesized subjective organization. First, we consider the possibility that these results were produced by differences in the frequency with which particular presidents were assigned to the three types of classifications. Second, we consider the possibility that these results came about because the average distance between presidents and assigned group boundaries differed from classification to classification. Finally, we consider whether the requirement to learn a new and different organization before each block may have made trials classified in one way more difficult than trials classified in another.

The first alternative explanation was that differences among item-and-category-consistent, item-consistent, and inconsistent classifications stemmed from individual presidents not being assigned to the

three classifications equally often. If item-and-category consistent classifications involved better known presidents than inconsistent classifications, this alone could account for differences in performance between the classifications.

Two sources of evidence argued against this possibility. First, by the measure of presidential knowledge used in Experiment 1 (average ordinal placement error), inconsistent classifications, which were the most difficult classifications, involved better known presidents than the two easier classifications. Mean placement error of presidents involved in inconsistent classifications was 7.9, that for presidents involved in item-and-category-consistent classifications 8.3, and that for presidents involved in item-consistent classifications 12.2 (with lower placement error indicating greater knowledge of the president).

The view that ease of classification reflected the fame of the particular presidents involved in the classifications also failed to explain the significant List × President interactions that appeared on both solution time and accuracy measures. The president's fame would be the same regardless of the list in which he appeared. Moreover, differences among the three types of classifications were consistently evident within individual presidents. In general, responses to a given president were fastest and most accurate when the president's classification was item-and-category-consistent, slowest and least accurate when it was inconsistent, and of intermediate difficulty when it was item-consistent. On the RT measure, 46 of 51 within-president differences fit the prediction (Figure 4); on the percent errors measure, 47 or 51 met it (in both cases, $p < .001$, by a two-tailed sign test). Thus, this alternative interpretation could be rejected.

The second alternative explanation involved the distance of items from list boundaries. In previous experiments, items close to list-defined boundaries have tended to be more difficult than more distant items (Derosa & Morin, 1970; Taylor, Kim, & Sudevan, 1984). This was also true in the present experiment; *boundary distance,* defined as the number of presidents between the president being classified and the nearest list-defined group boundary, was significantly correlated with both RT ($r = -.58$, $p < .001$) and percent errors ($r = -.47$, $p < .001$). Boundary distance was also correlated with type of classification. For item-and-category-consistent classifications, average boundary distance was 6.7 presidents; for item-consistent classifications, 4.4 presidents; for inconsistent classifications, 2.4 presidents. This pattern is consistent with the view that differences in average distance from list-defined category boundaries were responsible for the differences among the three types of classifications.

There is one major difficulty with this alternative interpretation,



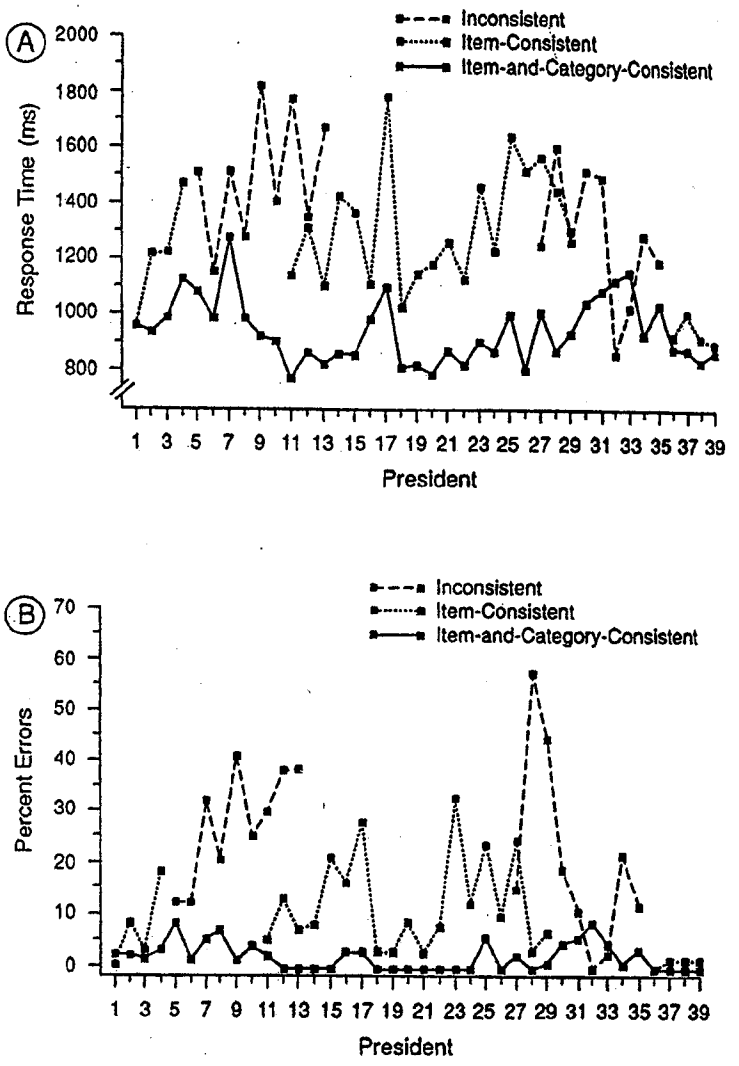



Figure 4. Mean correct response time (A) and percent errors (B) as a function of president and trial classification (Experiment 2 data)

though. It cannot account for the systematic differences among classifications that were present even when boundary distance was held constant. Table 3 illustrates the differences among item-and-category-consistent, item-consistent, and inconsistent trials when boundary dis-

Table 3. Performance on item-and-category-consistent (ICC), item-consistent (IC), and inconsistent (I) trials at each boundary distance (Experiment 2 data)

| Boundary distance | Measure | Type of classification ICC | IC | I | Mean |
|---|---|---|---|---|---|
| 0 | RT | 1228 | 1316 | 1325 | 1337 |
| | Error | .09 | .10 | .21 | .14 |
| | *n* | 4 | 6 | 6 | 16 |
| 1 | RT | 1199 | 1379 | 1472 | 1369 |
| | Error | .11 | .13 | .30 | .19 |
| | *n* | 4 | 6 | 6 | 16 |
| 2 | RT | 1062 | 1223 | 1361 | 1197 |
| | Error | .03 | .13 | .22 | .12 |
| | *n* | 6 | 6 | 4 | 16 |
| 3 | RT | 990 | 1366 | 1332 | 1217 |
| | Error | .03 | .12 | .16 | .10 |
| | *n* | 6 | 6 | 4 | 16 |
| 4 | RT | 938 | 1268 | 1622 | 1130 |
| | Error | .01 | .13 | .26 | .08 |
| | *n* | 8 | 4 | 2 | 14 |
| 5 | RT | 936 | 1243 | 1206 | 1043 |
| | Error | .02 | .08 | .28 | .05 |
| | *n* | 9 | 4 | 1 | 14 |
| 6 | RT | 929 | 1469 | 1493 | 1111 |
| | Error | .03 | .25 | .37 | .11 |
| | *n* | 8 | 3 | 1 | 12 |
| 7 | RT | 898 | 1019 | | 920 |
| | Error | .00 | .02 | | .01 |
| | *n* | 9 | 2 | | 11 |
| 8 | RT | 936 | 1034 | | 955 |
| | Error | .01 | .00 | | .01 |
| | *n* | 8 | 2 | | 10 |
| 9 | RT | 922 | 1023 | | 934 |
| | Error | .01 | .00 | | .01 |
| | *n* | 8 | 1 | | 9 |
| >9 | RT | 883 | | | 883 |
| | Error | .00 | | | .00 |
| | *n* | 22 | | | 22 |

*Note.* *n* = Number of presidents receiving each type of classification at each boundary distance.



tance is held constant. Illustratively, for presidents who were two presidents removed from a list boundary, item-and-category-consistent classifications took an average of 1,062 ms, item-consistent classifications an average of 1,223 ms, and inconsistent classifications 1,361 ms. Overall, within given levels of boundary distances, 22 of 24 differences among pairs of the three classifications were in the predicted direction, both for RTs and errors (in both cases, $p < .001$, by a two-tailed sign test). Thus, the hypothesized preexisting categories influenced learning of the list-defined categories even when distance of the president from the list-defined boundary was held constant.

Finally, we examined whether the results could be explained by interlist interference. A within-subjects design was used in the experiment; all subjects studied all four test lists. Presentation order of these lists was determined by five Latin squares. Given this design, one might argue that performance in this study reflected the difficulty subjects had in learning four different organizations of the same material in rapid sequence. Indeed, there is empirical support for the general notion that list learning is more difficult when list items are regrouped from trial to trial than when they are grouped consistently across trials (Bower, Lesgold, & Tieman, 1969). More concretely, consider a subject who received the 6-25-8 organization (List 3) during the first block and the 13-13-13 organization (List 1) during the second. In effect, this subject was first being asked to learn that Presidents 7 through 13 belong to the middle group, and then to learn that they do not belong to it. It is plausible that the requirement to unlearn the old group label and learn a new one could have resulted in the formation of a relatively weak link between "category-shifted" presidents (i.e., presidents that were listed as members of one category during one trial and as members of a different category during the next trial) and the new label. In addition, the presence of a preexisting, imperfectly unlearned, old group label may have interfered with the retrieval of the imperfectly learned, new category label during the test phase. For both of these reasons, subjects may have found trials involving category-shifted presidents to be more difficult than those involving "category-same" presidents (i.e., presidents who were assigned to the same category across all four experimental blocks), regardless of the preexperimental organization of presidents in memory.

The simplest way to address the interlist interference hypothesis is to focus on the category-same presidents (i.e., presidents who never appeared in inconsistent trials). In Figure 4, it can be seen that 21 of the 39 presidents were of this sort (Presidents 1–4, 14–26, and 36–39). Because these presidents were assigned to the same group across all four experimental blocks, the interlist interference hypoth-

esis provides no reason for predicting that the item-consistent trials involving these presidents should be more difficult than the item-and-category-consistent trials. In contrast, for reasons detailed above, our "interference from prior knowledge" hypothesis predicts that item-consistent trials should be more difficult than item-and-category-consistent trials, regardless of whether a particular president is assigned to one group or two over the course of the experiment. Figure 4 clearly demonstrates that the data are consistent with the "interference from prior knowledge" hypothesis and inconsistent with the interlist interference hypothesis. Item-and-category-consistent trials were faster than item-consistent trials for 20 of the 21 presidents, and they were less error prone for 19 of the 20 presidents showing a difference in percent error. Percent errors was the same in both conditions for one president (for both RT and percent errors, $p < .001$, by a two-tailed sign test). There is no straightforward way for the interlist interference hypothesis to explain this very consistent pattern.

## GENERAL DISCUSSION

The present results indicated that college students subjectively organize U.S. presidents into three groups—Founding Fathers, middle-period presidents, and recent presidents—and that they use these groups in comparing when the presidents served in office and in learning new organizations of the presidents. More generally, the study demonstrated how subjective organizations can be inferred from paired-comparison data, how these organizations can be used to predict learning patterns, and how learning patterns can provide converging evidence for the inferred organizations. We will now discuss questions raised by the study concerning the relation of presidential information to other historical knowledge, whether people divide the three main groupings of presidents into additional subcategories, and which conditions for learning new organizations are most useful for testing hypothesized subjective organizations.

Presidential information seems unlikely to be represented in isolation. Instead, it probably is integrated into a broader framework of historical knowledge. N. Brown (1990) found that recent presidents are closely linked to important public events that occurred during their tenure. Similar links between noncontemporary presidents and earlier events in American history also seem likely. For example, one can hardly think of the Civil War without thinking of Abraham Lincoln, and vice versa. This notion is also consistent with more general



findings that point to strong links in autobiographical memory between knowledge of individuals and knowledge of activities in which those individuals participated (Linton, 1986; Reiser, Black, & Kalamarides, 1986; Wagenaar, 1986; Williams & Hollan, 1981).

The view that knowledge of American presidents and of the historical events in which they participated are closely bound in memory introduces the possibility that the boundaries we have isolated reflect more than a partitioning imposed on the set of presidents. Instead, these divisions may reflect subjectively defined periods in American history that include presidential and numerous other kinds of information. That is, at a general level, people may organize their knowledge of American history into three periods: the post–Revolutionary War period, the post–World War II period, and a period that encompasses the nation's middle years. If this is the case, the present findings on the presidencies may be based on period information rather than on groupings of presidencies as such. This period information could be directly associated with individual presidents, with sets of presidents, or with events that are in turn directly associated with individual presidents or groups of presidents. Regardless of the exact details, this interpretation implies that a pattern of results much like that in Experiments 1 and 2 would arise if events from the three major periods, rather than names of presidents, were used as stimuli.

Perhaps the most surprising finding of the study involved the lack of differentiation among presidents in the middle period. Intuitively, it seems improbable that people consider Andrew Jackson and Franklin Roosevelt members of the same group. Similarly, it is hard to accept that students consider the years 1829–1945 to constitute a single period in American history.

One possibility is that college students retain little information about American history, and that the knowledge that they do retain is primarily of the recent past and the Revolutionary War period. Under this interpretation, knowledge of American history from 1829–1945 may be limited to unintegrated knowledge of a few isolated interludes, such as the Civil War, the Roaring 20s, and the Depression. Many presidents from the middle period would not be associated with any of these subperiods, and considerable uncertainty might exist concerning who was president during some of them, such as the Roaring 20s.

Even if people possessed accurate knowledge of the presidents during these relatively well-known subperiods, it would not always allow them to make accurate temporal comparisons involving these presidents. This is because the comparison task generally requires that temporal information be available for both of the presidents who are

being compared. A student might be sure that Lincoln was president from 1861–1865, but if he did not know when Hayes was president, the knowledge of Lincoln would be of little help in deciding whether Lincoln served before or after Hayes. The existence of large numbers of noncontemporary presidents who are only vaguely associated with the time when they were in office may have obscured the presence of cohesive subgroups within this extensive period.

The methods used in this study may render more tractable an important but difficult problem: how to provide converging evidence for empirically defined subjective organizations. Two types of subjective organizations have been studied: ones defined on the basis of a priori considerations, such as state boundaries and gender, and ones induced from patterns of empirical data. The categories defined on the basis of a priori considerations are easier to test, but the categories without such strong a priori bases raise questions that are at least as interesting. To cite just one illustration, why do people sharply distinguish Founding Fathers from other presidents, but not nineteenth- from twentieth-century presidents or pre–Civil War from post–Civil War ones? Precisely because alternative organizations could easily be imagined, these empirically defined subjective organizations seem especially revealing of the organizational framework subjects have imposed on the domain.

Although such empirically defined organizations can be both interesting and revealing, evidence for their existence is often problematic. The difficulty comes from inferring discontinuous groupings on the basis of continuous data, such as mean solution times and percent errors. Under such circumstances, discontinuities inevitably are to some degree in the eye of the beholder. Predicting learning of new organizations from assessments of existing ones, however, seems to provide a means for overcoming this problem. It allows tests of predictions that would be unlikely to be accurate if the hypothesized group boundaries were incorrect. Further, it seems widely applicable; the approach can be used whenever both existing and new organizations unambiguously assign individual instances to categories. Overall, then, the approach seems to provide a generally useful tool for learning more about subjective organizations.

## Notes

This research was supported in part by a National Institute of Mental Health postdoctoral fellowship to Norman R. Brown and in part by National Institute of Health Grant HD-19011 to Robert S. Siegler. We wish to thank Nancy Cooke, Fred Conrad, Robert Crowder, Ronald Fisher, Ruth Maki, Henry Roediger, and several anonymous reviewers for their comments on earlier versions of this paper. We also thank Judy Goodman, Sergio Canetti, and Jeff Kelley for their assistance.

Correspondence concerning this article should be addressed to Norman R. Brown, who is now at the Department of Psychology, Carnegie-Mellon University, Pittsburgh, PA 15213, or to Robert S. Siegler at this same address.
Received for publication September 13, 1989; revision received February 26, 1990.



## References

Anderson, J. R. (1983). *Cognitive psychology and its implications.* New York: Freeman.

Anderson, J. R., & Ross, B. H. (1980). Evidence against a semantic-episodic distinction. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 441–478.

Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgments. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 101–159). New York: Academic Press.

Bower, G. H., Lesgold, A. M., & Tieman, D. (1969). Grouping operations in free recall. *Journal of Verbal Learning and Verbal Behavior, 8,* 481–493.

Bransford, J. D. (1979). *Human cognition: Learning, understanding, and remembering.* Belmont, CA: Wadsworth.

Brown, A. L., & DeLoache, J. S. (1978). Skills, plans, and self-regulation. In R. S. Siegler (Ed.). *Children's thinking: What develops.* Hillsdale, NJ: Erlbaum.

Brown, N. R. (1990). Organization of public events in long-term memory. *Journal of Experimental Psychology: General, 119,* 297–314.

Brown, N. R., Rips, L. J., & Shevell, S. K. (1985). The subjective dates of natural events in very long-term memory. *Cognitive Psychology, 17,* 139–177.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York: Academic Press.

Chi, M. T. H., & Koeske, R. D. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology, 19,* 29–39.

Collins, A., & Stevens, A. L. (1982). Goals and strategies of inquiry teachers. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 65–120). Hillsdale, NJ: Erlbaum.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Derosa, D. V., & Morin, R. E. (1970). Recognition reaction time for digits in consecutive and nonconsecutive sets. *Journal of Experimental Psychology, 83,* 472–479.

Ekstrand, B. R., & Underwood, B. J. (1965). Free learning and recall as a function of unit-sequence and letter-sequence interference. *Journal of Verbal Learning and Verbal Behavior, 4,* 390–396.

Estes, W. K., Allmeyer, D. J., & Reder, S. M. (1976). Serial position functions

for letter identification at brief and extended exposure durations. *Perception & Psychophysics, 19,* 1–15.

Friedman, W. J., & Wilkins, A. J. (1985). Scale effects in memory for the time of events. *Memory & Cognition, 13,* 168–175.

Hamilton, J. M. E., & Sanford, A. J. (1978). The symbolic distance effect for alphabetic order judgments: A subjective report and reaction time analysis. *Quarterly Journal of Experimental Psychology, 30,* 33–43.

Hinsley, D. A., Hayes, J. R., & Simon, H. A. (1977). From words to equations: Meaning and representation in algebra word problems. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 89–108). Hillsdale, NJ: Erlbaum.

Holyoak, K. J., & Patterson, K. K. (1981). A positional discriminability model of linear-order judgments. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1283–1302.

Kinder, D. R., & Fiske, S. T. (1986). Presidents in the public mind. In M. Hermann (Ed.), *Political psychology* (pp. 193–218). San Francisco: Jossey-Bass.

Kosslyn, S. M., Murphy, G. L., Bemesderfer, M. E., & Feinstein, K. J. (1977). Category and continuum in mental comparisons. *Journal of Experimental Psychology: General, 106,* 341–375.

Linton, M. (1986). Ways of searching and the contents of memory. In D. Rubin (Ed.), *Autobiographical memory* (pp. 50–67). New York: Cambridge University Press.

Maki, R. H. (1981). Categorization and distance effects with spatial linear orders. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 7,* 15–32.

McKoon, G., Ratcliff, R., & Dell, G. S. (1986). A critical evaluation of the semantic-episodic distinction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 295–306.

Moyer, R. S., & Dumais, S. T. (1978). Mental comparison. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 12, pp. 117–155). New York: Academic Press.

Pliske, R. M., & Smith, K. H. (1979). Semantic categorization in a linear order problem. *Memory & Cognition, 7,* 297–302.

Postman, L. (1963). Extra-experimental interference and the retention of words. *Journal of Experimental Psychology, 61,* 97–110.

Reiser, R. J., Black, J. B., & Kalamarides, P. (1986). Strategic memory search process. In D. Rubin (Ed.), *Autobiographical memory* (pp. 100–121). New York: Cambridge University Press.

Roediger, H. L., III, & Crowder, R. G. (1976). A serial position effect in recall of United States presidents. *Bulletin of the Psychonomic Society, 8,* 275–278.

Shepard, R. N., Kirkpatric, D. W., & Cunningham, J. P. (1975). The internal representation of numbers. *Cognitive Psychology, 6,* 82–138.

Siegler, R. S., & Klahr, D. (1982). When do children learn? The relationship between existing knowledge and the acquisition of new knowledge. In



R. Glaser (Ed.), *Advances in instructional psychology* (pp. 121–206). Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Robinson, M. (1982). The development of numerical understanding. In H. Reese & L. Lipsitt (Eds.), *Advances in child development and behavior* (pp. 242–312). New York: Academic Press.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18,* 275–290.

Sternberg, R. J., & Powell, J. S. (1983). The development of intelligence. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology* (Vol. 111, pp. 341–419). New York: Wiley.

Strauss, S., & Ephron-Wertheim, T. (1986). Structure and process: Developmental psychology as looking in the mirror. In I. Levin (Ed.), *Stage and structure: Reopening the debate* (pp. 59–76). Norwood, NJ: Ablex.

Surber, C. F., & Haines, B. A. (1987). The growth of proportional reasoning: Methodological issues. *Annals of Child Development, 4,* 35–87.

Taylor, D. A., Kim, J., & Sudevan, P. (1984). Representation of linear orders. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 61–71.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford University Press.

Underwood, B. J., & Postman, L. (1960). Extra-experimental sources of interference in forgetting. *Psychological Review, 67,* 73–95.

Wagenaar, W. A. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology, 18,* 225–252.

Wilkening, F., & Anderson, N. (1982). Comparison of two rule-assessment methodologies for studying cognitive development and knowledge structure. *Psychological Bulletin, 92,* 215–237.

Williams, M. D., & Hollan, J. D. (1981). The process of retrieval from very long-term memory. *Cognitive Science, 5,* 87–119.

Wyer, R. S., Shoben, E. J., Furhman, R. W., & Bodenhausen, B. V. (1985). Event memory: The temporal organization of social action sequences. *Journal of Personality and Social Psychology, 49,* 857–877.