Abstract

There are many ways in which a GSS can be used to support group brainstorming. This paper reports the results of an experiment that manipulated task structure and time structure. Groups electronically brainstormed on intact tasks (where all parts of the task were presented simultaneously) or on partitioned tasks (where three subcategories of the task were presented to the groups). The time periods in which groups worked were either one 30-minute time period or three 10-minute periods separated by two-minute breaks. Groups in the partitioned task treatment generated 40% more ideas, but there were no time effects. These differences are attributed to the ability of the partitioned task to re-focus members’ attention more evenly across the entire solution space.

Keywords: Groupware, group support systems, electronic brainstorming, entrainment, problem structure, time

ISRL Categories: AA0903, HA11

Introduction

Group support systems (GSS) have been used in organizations to support brainstorming and idea generation for more than a decade (e.g., Dennis 1994; Nunamaker et al. 1987; Tyran et al. 1992). GSSs enable group members to interact electronically by exchanging anonymous, typed ideas, rather than taking turns speaking. In general, groups using GSS tend to generate more ideas than verbally interacting groups and as
many or more ideas than nominal groups (Dennis and Valacich 1993; Gallupe et al. 1991; 1992; Valacich et al. 1994). However, there have been important performance differences among the various studies (Dennis et al. 1996), so much so that some researchers have expressed concerns about the inconsistencies (e.g., Benbasat et al. 1993; Dennis and Gallupe 1993).

GSSs are flexible tools that can be used in a variety of ways (DeSanctis and Poole 1994; Nunamaker et al. 1991). In recent years, the issue of task/technology fit has been recognized as an important factor in understanding the effects of technology use (Goodhue and Thompson 1995; Vessey 1991). Technology use can produce different results depending upon its fit with the needs of the task (Goodhue and Thompson 1995) and the way in which groups choose to use it (DeSanctis and Poole 1994). Understanding how to use a GSS for a specific task is at least as important as recognizing the need to use one in the first place.

GSSs enable the use of new work processes that would be difficult or impossible without them (DeSanctis and Gallupe 1987; Nunamaker et al. 1991). However, groups usually have traditional processes for working on tasks and may have difficulty adapting the new GSS-based techniques into their work processes (DeSanctis and Poole 1994). Decision makers often shy away from complex techniques in favor of simple ones, even when complex ones are likely to lead to better performance (Kottemann and Davis 1991).

This study examines two rather uncomplicated approaches by which a GSS can be used to change the way in which groups work. One aspect of group brainstorming that has received little research attention is how the task is structured, yet it is known that how a task is presented can have significant effects of performance (Kahneman and Tversky 1972). Likewise, it has been argued that GSS use can change the pace of group work (Nunamaker et al. 1991), yet little research has examined how time is allocated in GSS use, despite considerable evidence that the time allotted for a task also can have important impacts on performance (McGrath et al. 1984).

This paper reports on the results of an experiment manipulating the task structure and time structure imposed on groups using a GSS to perform an idea generation task. First, previous research on these topics is reviewed and the hypotheses presented. Next, the research methodology is examined. Finally, the results are presented, and their implications for research and practice are discussed.

Theory and Research

Group brainstorming is influenced by both cognitive phenomena within individual group members and social phenomena as group members interact (Nagasundaram and Dennis 1993). Although social phenomena have tended to dominate previous research, cognitive factors are also important. Cognitive phenomena and social factors that arise from task structure are considered first followed by the issue of how they may interact with time.

Task Structure

One of the basic notions behind the scientific method is that a problem should be explicitly “decomposed” or broken down into a series of subproblems. Solutions are then obtained for each subproblem, and these solutions combined (Armstrong et al. 1975, p. 257).

It has been theorized that most tasks have a hierarchical or tree-like structure, enabling them to be divided and considered as a set of individual subtasks that focus on more specific subcategories within the overall task (Rosch 1978). Many other researchers have argued similarly (e.g., see Gettys et al. 1987; Pitz et al. 1980). For example, if the objective is to solve a campus parking task, one might pose it as one question: “How can we solve the parking problem?” Conversely, one might break it into several subcategories: “How could we solve the parking problem by providing more parking and by reducing demand for parking through alternate forms of transportation?” (see Gettys et al. 1987).

Theory suggests that when presented with a task structured as one all-encompassing question, individuals should produce a set of related ideas.
focusing on a small set of the task’s subcategories. Cognitive behavior is controlled by production rules—rules specifying the steps of cognition—that produce ideas when activated (Anderson 1983, 1987). Rules are activated automatically by input stimuli, without conscious control (Anderson 1992). For any given stimuli, there are often several rules that could be activated. Each rule has a certain strength (i.e., likelihood of being activated) based on past experiences.

The ideas generated by an individual member of the brainstorming group depend upon the stimuli and the relative strengths of the individual’s production rules. As activation spreads through memory, rules that are more closely related to the stimuli and to each other have the greatest strength and are most likely to be activated. These closely related rules will likely lead to the production of closely related ideas. Thus an uninterrupted stream of ideas will tend to follow a consistent train of thought in a set of semantically similar subcategories, because related stimuli activate related production rules. As a result, an individual may focus on a narrow subset of the task space, overlooking other promising subcategories.

There is no direct empirical evidence to support Anderson’s theory, but there is some related empirical evidence from individual problem solving research. Individuals presented with an all-encompassing task tend to focus only on a small fraction of the potential solution space, yet believe themselves to have produced a very complete set of solutions (Connolly et al. 1993; Gettys et al. 1987; Pitz et al. 1980). They explore a few related subcategories in greater depth rather than considering a broader range of subcategories in the overall task, usually missing many important ideas.

One solution is to adopt a structured technique that changes an individual’s focus so that effort is more evenly allocated across the task. There is some evidence that breaking the task into a set of subcategories improves performance by encouraging individuals to devote attention more evenly to the entire set of categories (e.g., Armstrong et al. 1975), although other studies have found few effects (e.g., Pitz et al. 1980). Some modest amount of structure—perhaps three or four categories—should lead to better performance (Samson 1988, p. 57).

Part of the reason for forming a group is to provide a diversity of approaches to the task so as to offset the potential narrowness of any one individual’s perception. Thus, what breaking the task into subcategories does for individual problem solving may be partially achieved by forming a group. Each member of a group has their own, different set of production rules available to be triggered by the brainstorming stimuli, which may lead to greater diversity.

Nonetheless, a similar phenomenon to that in individual research has also been observed in groups, where social factors within groups can constrain idea generation. Group brainstorming can suffer from cognitive inertia, the tendency of group discussion to focus on a few lines of thought in one subcategory (Lamm and Trommsdorff 1973). As group members interact, they may consciously or unconsciously adopt behavior norms. These norms or structures can constrain behavior (Giddens 1984). One of the structures typically found in group interaction is one of “not changing the subject” (Lamm and Trommsdorff 1973, p. 382). Repeatedly changing the subject to focus on new ideas becomes socially undesirable. While members may think of unrelated ideas, social structures may inhibit the contribution of these ideas to the discussion. Members are motivated to focus on one train of thought in a narrow set of related subcategories (Maier 1970). Thus groups also constrain their attention to a narrow set of subcategories when faced with one all-encompassing task statement. Such behavior has also been observed in GSS-based electronic brainstorming and has been associated with the generation of fewer ideas (Dennis and Valacich 1994).

In summary, cognitive factors encourage individuals to focus on only a few subcategories of a task, thus overlooking other areas. Social factors also encourage group members to work on the same subcategory rather than exploring many different parts simultaneously. Breaking the task into subcategories should overcome both of these shortcomings by encouraging groups to perform a more complete exploration of the task solution space, resulting in improved performance.
A review of the literature discovered only one previous study that has examined the impact of structuring the group brainstorming task in this way. That study (Dennis et al. 1996b) examined GSS groups working on a brainstorming task that was either presented in an intact form as one question or in a decomposed form as three separate questions. The groups in the three-question treatment were required to focus their attention on each part of the task separately in three independent brainstorming sessions and thus were forced to allocate their time equally to essentially three distinct tasks. This three-part approach resulted in about 60% more ideas being generated, with an even greater improvement in idea quality.

This approach of breaking the task into subparts and addressing each separately and in sequence could easily be used in traditional verbal brainstorming. However, GSSs also enable new forms of group work that are difficult or impossible in traditional verbal interaction. A GSS enables the task to be structured into subparts that are distinct and separate from each other (i.e., ideas from the separate subparts are not intermixed) and at the same time enables all subparts to be presented to the group members simultaneously so that each member can choose to work on same or different subpart from other members. In this way, members are not required to work on separate subparts for arbitrary lengths of time. They can allocate their time and effort to the subparts of the problem that are the most productive for them as individuals. Therefore, it is hypothesized:

\textbf{H1a:} Groups working on a multiple question task formulation will generate more ideas than groups working on a single question formulation.

\textbf{H1b:} Groups working on a multiple question task formulation will generate ideas of greater quality than groups working on a single question formulation.

\textbf{Time Structuring}

As discussed above, it is possible to partition the time allocated to a task into several distinct work periods. Group members could work on the task for a time period, interrupt their work to take a break or to perform some unrelated task, and then resume work on the task. This type of time structure may affect group performance in two distinct ways.

First, such an “incubation” period may have the same effect as the task structuring above. The goal of the task structuring is to encourage members to allocate their effort more evenly over the individual subcategories within the task. Forcing group members to break their work pattern and think about something unrelated may break cognitive inertia. When individuals resume work they may focus on a different part of the task, so that they redirect their efforts to different subparts of the task that were previously overlooked (Dominowski and Jenrick 1972; Woodworth and Schlosber 1954). Some studies of incubation have shown that an incubation period of only a few minutes (even just one minute) is sufficient to induce individuals to redirect their efforts and improve performance (Adamson and Taylor 1954; Dominowski and Jenrick 1972; Murray and Denny 1969).

Second, there is evidence that externally imposed time constraints affect the pace of group work (Gersick 1988). If groups perceive they have a short time period relative to the task, they work more quickly to accomplish the task (Kelly 1988; Kelly et al. 1990; Locke and Latham 1990; McGrath et al. 1984). The reverse is also true: if groups perceive that the time period is sufficient to complete a task they work more slowly; work

Idea quality should follow the same pattern as idea quantity: more ideas leads to more quality ideas (Osborn 1953). Previous research has found the number of ideas to be a reasonable predictor of overall quality (Dennis et al. 1996b; Diehl and Stroebe 1987; Valacich et al. 1994). Therefore, it is hypothesized:

\textbf{H1b:} Groups working on a multiple question task formulation will generate ideas of greater quality than groups working on a single question formulation.
expands to fill the time available for it (Parkinson 1957). Although short time periods may increase the rate of task performance and thus result in more ideas, quality may suffer (Kelly and Karau 1993). Time pressure tends to focus attention more narrowly so that some subparts of the task are overlooked, which leads to lower quality solutions (Karau and Kelly 1992; Kelly and Karau 1993; Kelly and McGrath 1985; Kelly et al. 1990).

To some extent, the perception of time available may be more important than the actual time available (Kelly 1988; Kelly et al. 1990; Locke and Latham 1990). Considerable research has shown that the amount of time allocated to the initial time period can establish the pace of work used in subsequent time periods, even if those time periods are of different lengths (Kelly and Karau 1993; Kelly and McGrath 1985; Kelly et al. 1990). This effect, called "entrainment," shows that a short initial work period can induce the group to work more quickly in subsequent time periods when more time is available and that a long initial time period induces slower work even when faced with time pressure in later periods.

In summary, a series of shorter, multiple time periods permits incubation and redirection of attention to occur, possibly increasing idea generation performance. The use of shorter work periods may increase the pace of idea generation, but reduce idea quality. Thus it is hypothesized:

**H2a:** Groups working in multiple time periods will generate more ideas than groups working in a single time period.

**H2b:** Groups working in multiple time periods will not generate ideas of greater quality than groups working in a single time period.

### Method

**Subjects**

Subjects were junior and senior undergraduate students enrolled in a general management course at a large state university. Four hundred students participated as members of 40 10-member groups that were randomly assigned to the treatments. About 42% were female. The average age was 21.0 years.

**Tasks**

The experimental task was the same as that used by Experiment 4 in Valacich et al. (1994) and was similar to tasks used in prior studies of electronic brainstorming. The task asked subjects to generate ideas to improve the environment. It asked them to "generate ideas to improve air quality, water quality, and land quality." Subjects were instructed to think broadly and to be creative in their ideas. The subjects generated more than 1,000 distinct solutions to the task, suggesting that the task is non-trivial, because complexity increases as the number of potential solutions increases (Campbell 1988).

**Independent Variables**

The two independent variables were time structure (one time period or multiple time periods) and task structure (single question or multiple question). Groups working in the single time period treatment were given one 30-minute time period to complete the task. Groups in the multiple time period treatment were given three 10-minute time periods (separated by two minute breaks) in which to complete the task.  

---

3It is difficult to make compelling arguments for choosing one specific number and length of time periods over another. Previous research on entrainment has tended to use three time periods, ranging from four minutes to 20 minutes (e.g., see Kelly and Karau 1993; Kelly and McGrath 1985; Kelly et al. 1990; Karau and Kelly 1992). Time periods of 10 minutes have been commonly used. In some cases, researchers have attempted to measure the "optimal" time needed to complete the task or to generate some arbitrary number of solutions (e.g., Karau and Kelly 1992), and then have arbitrarily varied the time periods around that "optimal." In other cases, no clear rationale has been given for the choice of time periods (e.g., Kelly and McGrath 1985; Kelly et al. 1990). In the absence of compelling theory, three 10-minute time periods were chosen to match prior research (e.g., Dennis et al. 1996) and so that the total time devoted to the task(s) was the same for all treatments.
Task structure was manipulated by presenting the task either as one question or as three questions. Three questions were used because this task had a "natural" decomposition into three areas. Groups working in the single question treatment were given the task as stated above as one question attached to one list in which to type their ideas. All ideas were typed into the one list. Groups in the multiple question treatment were given the task as stated above as three questions and three separate lists, each labeled with one of the three subcategories separately (air quality, water quality, land quality) in which to type their ideas (see Pitz et al. 1980).

**Software**

The software used in this study was GroupSystems (see Valacich et al. 1991). The software presented the user with a screen split horizontally, with the brainstorming question and the ideas of others displayed across the top and a window in which to type ideas on the bottom. In the single question treatment, ideas were placed in one list so that all group members could see all the ideas at the same time by scrolling up and down. In the multiple question treatments, the ideas were organized into three separate lists (one for each of the three subcategories), all of which were available to the subjects by selecting the list from which they wanted to read or into which they wanted to contribute. There were no restrictions on which list a subject could access.

**Measures**

Performance was first measured by counting the number of unique ideas. Output from each group was printed and coded by a rater. The rater was instructed to identify all unique (i.e., non-redundant) ideas proposed in the transcripts from each group. An idea was counted only once for each group even if it appeared multiple times in the same or several time periods or question lists. An idea was counted even if it was generated in the "wrong" task category (e.g., an air quality idea contributed in the land quality category). A second rater independently coded the output of 10 randomly selected groups (25%). Inter-rater agreement was adequate (alpha = .99).

Next, the quality of each idea was assessed by the first rater using a 1 = Very Poor Quality to 5 = Very Good Quality scale. The second rater independently assessed the quality of 25% of the ideas. Inter-rater agreement was adequate (alpha = .97). These quality assessments can be combined to produce several possible measures of quality. Three measures of idea quality from prior research were used: total idea quality, mean idea quality, and the quality of good ideas (see Diehl and Stroebe 1987).

Total quality, which has proven to be the most consistently reliable measure across studies (Diehl and Stroebe 1987), was calculated by summing the quality scores for each unique idea generated by the group. This measure rewards groups for all ideas they produce, even very poor ideas with low quality scores. For example, a group generating five very poor ideas (each rated a 1) would have the same score as a group generating one very good idea (rated a 5).

The second measure was mean quality, which has proven to be unreliable across studies (Diehl and Stroebe 1987). Mean quality was the average quality of ideas generated by each group (i.e., total quality divided by the number of ideas). This measure rewards groups that generate high quality ideas and penalizes those that produce any lower quality ideas. For example, a group generating one very good idea (rated a 5) would have a higher mean quality than a group generating three very good ideas (rated 5) and one neutral quality idea (rated a 3).

The third measure was the quality of good ideas. This was calculated in the same manner as total quality, but using only those ideas with a quality rating of 3 or higher. This measure attempts to strike a balance between total quality and mean quality. The 3rd measure of idea quality, which has proven to be the most consistently reliable measure across studies (Diehl and Stroebe 1987), was calculated by summing the quality scores for each unique idea generated by the group. This measure rewards groups for all ideas they produce, even very poor ideas with low quality scores. For example, a group generating five very poor ideas (each rated a 1) would have the same score as a group generating one very good idea (rated a 5).
quality so that groups are rewarded for all ideas they produce, except those that are poor (i.e., rated less than 3) (see Diehl and Stroebe 1987).

A post-session questionnaire was also used to assess subjects' satisfaction (five items, alpha = .83) and whether they had sufficient time (four items, alpha = .73). These measures were adapted from previous research (see the Appendix). All used seven-point Likert scales, with 1 meaning low (or insufficient time) and 7 meaning high (or sufficient time). All questionnaire data were analyzed using ANOVA with a group-nested-within-treatment term to capture any similarities in responses from members of the same groups (see Dennis 1996).

**Results**

The means and standard deviations for all measures are shown in Table 1. The results of the statistical analyses are shown in Table 2. There was a main effect for task structure on the number of unique ideas generated: groups in the multiple question treatments generated more unique ideas (F(1,36) = 8.98, p = .005). There were no effects due to time and no interaction effects. The pattern was the same for total quality and the quality of good ideas. Groups in the multiple question treatments had higher total quality (F(1,36) = 8.94, p = .005) and higher quality of good ideas (F(1,36) = 7.52, p = .009). There were no effects due to time and no interaction effects for either quality measure. There were no differences in mean quality for any treatments. Therefore, H1a and H1b are supported. H2a is not supported, but H2b (no quality effects) is supported.

Subjects in all treatments reported similar satisfaction, except for those in the single question/multiple time period treatment who reported significantly lower satisfaction (see Tables 1 and 2, Tukey, alpha = .05). Subjects in all treatments were equally likely to report that they had had sufficient time (see Tables 1 and 2).

**Discussion**

The results from this experiment indicate that the way in which the same GSS is used for the same task can have significant impacts on performance. Simply presenting participants with a different representation of the task and enabling them to address each part separately improved performance. Groups working on the multiple question task generated about 40% more ideas on average than those working on the single question task, with greater total quality and quality of good ideas.

These results have a clear, practical significance to GSS users. Using a GSS with task structure adds no measurable cost compared to using no task structure, but produces a significant improvement in performance. GSS users are strongly encouraged to seek opportunities to decompose idea generation tasks into subparts and enable participants to work on each part separately.

There is, perhaps, an even more important implication for GSS researchers in the magnitude of this improvement. Task structure has received relatively little attention compared to other factors in GSS use (e.g., anonymity). A recent meta-analysis found GSS use without this type of task structuring to increase the number of ideas produced compared to verbal brainstorming by a mean effect size of .80 (Dennis et al. 1996a). In the present study, the effect size due to task structure (compared to no task structure) was .95. In other words, the use of this simple task structure had about the same effect on performance as the mean effect of using GSS in the first place! This suggests that there may be considerable opportunities for even greater performance improvements from GSS use through more sophisticated applications of task structure. Numerous techniques have been developed to improve verbal brainstorming (e.g., VanGundy 1988). We believe that our results call for the systematic investigation, development, and testing of more advanced task structuring techniques designed specifically for GSS use.

A second fundamental implication has to do with understanding the impacts of GSS use. Researchers have speculated that explanations for the inconsistency in previous research may lie in the characteristics of the system, task and/or group (Dennis et al. 1991) or their fit with each other (Goodhue and Thompson 1995). The data in this study suggest that the source of the inconsistencies may be even more
Table 1. Means and Standard Deviations

<table>
<thead>
<tr>
<th>Questions</th>
<th>1 Time Period</th>
<th>3 Time Period</th>
<th>1 Time Period</th>
<th>3 Time Period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std.</td>
<td>Mean</td>
<td>Std.</td>
</tr>
<tr>
<td>Unique Ideas</td>
<td>122.4</td>
<td>64.4</td>
<td>141.6</td>
<td>29.6</td>
</tr>
<tr>
<td>Total Quality</td>
<td>203.0</td>
<td>105.1</td>
<td>261.3</td>
<td>81.2</td>
</tr>
<tr>
<td>Mean Quality</td>
<td>1.64</td>
<td>0.20</td>
<td>1.87</td>
<td>0.52</td>
</tr>
<tr>
<td>Good Idea Quality</td>
<td>67.9</td>
<td>42.3</td>
<td>127.1</td>
<td>96.8</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>5.40</td>
<td>1.10</td>
<td>5.42</td>
<td>1.02</td>
</tr>
<tr>
<td>Sufficient Time</td>
<td>6.15</td>
<td>1.07</td>
<td>6.08</td>
<td>1.21</td>
</tr>
</tbody>
</table>

of a "wicked problem" (Mason and Mitroff 1981) than first thought. The same GSS, task, and groups were used in all treatments. It was a simple change in the way in which the GSS was used that produced the significant impact on performance. Some of the source of inconsistency in previous research may therefore lie in the way in which the GSS is used—something that is more difficult to assess than GSS, task, or group characteristics because it requires a more detailed understanding of the use processes, not just the inputs to them. If this is the case, then a better understanding of use processes is needed—and researchers must be encouraged to document them much more carefully in writing their research for publication.

One important question for future research is understanding why this type of task structure produced the effects that occurred. It was theorized that task structure would encourage groups to allocate their effort more evenly across the various subcategories. This more even allocation of effort would reduce the chance that subcategories rich in ideas would be overlooked and thereby result in more ideas. If this occurred, the expectation is that the number of unique ideas in each category would be more evenly distributed; i.e., there would be less variance for the groups working in the multiple question and multiple time treatments.

This theory was tested by categorizing each idea as either primarily water quality, land quality, air quality, or other and calculating the variance for each group. A 2 x 2 ANOVA on the variances was then performed, and the main effects for task structure (F(1, 36) = 8.56, p = .006) were found, but no effects for time structure (F(1, 36) = 1.41, p = ns) or interaction (F(1, 36) = 0.10, p = ns) were found. Groups in the multiple question treatments generated ideas more evenly among the three categories (air: 36%; water: 30%; land: 34%; other: 0.2%) than did the groups in the single question treatments (air: 30%; water: 9%; land: 43%; other: 18%).

As hypothesized, groups in the multiple question treatments were more likely to allocate their ideas more evenly across the three subcategories. Their improved performance is attributed to this more even allocation of effort. In this study, the groups receiving the single question generated a far lower percentage of water quality ideas, preferring instead to focus on land and air quality. In fact, one group generated no water quality ideas at all, and another generated only one. The variance is defined as \( \sum (x_i - \mu)^2 \). In this case, one variance score was calculated for each group. Here, the \( x_i \) are number of land ideas, number of water ideas, and number of air ideas, and \( \mu = \sum x_i \) for each group. Higher variances would indicate a less even distribution of ideas among the three major categories and suggest that the groups allocated their efforts unevenly. When performance of the ANOVA was attempted, a Hartley test found these data to violate the ANOVA assumption of equal variance in each cell (i.e., the variance of the variance measure). The standard deviation was proportional to the mean, so a log transformation was performed prior to performing the ANOVA (see Neter et al. 1985, p. 615).
Table 2. Statistical Analyses

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Time Structure</th>
<th>Task Structure</th>
<th>Time × Task Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>F   p</td>
<td>F   p</td>
<td>F   p</td>
</tr>
<tr>
<td>Unique Ideas</td>
<td>40 Groups</td>
<td>0.96 ns</td>
<td>8.98 .005</td>
<td>2.52 ns</td>
</tr>
<tr>
<td>Total Quality</td>
<td>40 Groups</td>
<td>2.28 ns</td>
<td>8.94 .005</td>
<td>0.30 ns</td>
</tr>
<tr>
<td>Mean Quality</td>
<td>40 Groups</td>
<td>0.84 ns</td>
<td>0.31 ns</td>
<td>3.60 ns</td>
</tr>
<tr>
<td>Good Idea Quality</td>
<td>40 Groups</td>
<td>1.77 ns</td>
<td>7.52 .009</td>
<td>0.40 ns</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>400 Subjects</td>
<td>1.50 ns</td>
<td>24.48 .001</td>
<td>22.40 .001</td>
</tr>
<tr>
<td>Sufficient Time</td>
<td>400 Subjects</td>
<td>2.43 ns</td>
<td>1.68 ns</td>
<td>0.26 ns</td>
</tr>
</tbody>
</table>

The conclusion is that by de-emphasizing water quality, their total idea output suffered.

This analysis also points out one weakness in using task structure. Subjects in the single question treatments were more likely to generate ideas that did not fall into three main categories. By imposing task structure, subjects may have felt more constrained and avoided ideas outside these categories. One risk is that a structure may not completely cover the problem space; it may miss potentially important subcategories. By constraining ideas to predefined categories, the group may (intentionally or unintentionally) miss ideas outside of these categories. Another risk is that it may be more difficult to create good solutions without considering all subcategories of the task simultaneously. For example, ideas developed by focusing on one subcategory may violate a constraint imposed by some other aspect of the task and thus be infeasible (Campbell 1988). One question is whether this problem can be avoided by explicitly including an “other” category and encouraging participants to contribute ideas into it.

These results compare favorably to those of the previous study (Dennis et al. 1996b). The unstructured groups in that study worked on a task presented as one question in one long time period (30 minutes) that is similar to the single time period/single question treatment of the present study. However, the structured task in that study was broken into three parts and the group required to work on each part separately (part 1 for 10 minutes, part 2 for 10 minutes, part 3 for 10 minutes). The researchers found that the highly structured approach (combining both task and time structure in one treatment) resulted in the generation of about 60% more ideas.

We also tested this more restrictive approach hoping to induce a stronger entrainment effect. The multiple time period/multiple question treatment was redone, forcing groups to work on the first question in the first 10-minute time period, the second question in the second time period, and the third question in the third time period. The means from this treatment were essentially the same as the means in the original multiple time/multiple question treatment (unique ideas: 150.3; total quality: 234.7; mean quality: 1.56; good idea quality: 72.7; satisfaction: 5.42; sufficient time: 5.50). The statistical conclusions using these new values were identical to those in Table 2, with two exceptions. First, there were significant effects for the sufficient time questionnaire measure suggesting that subjects in the multiple time/multiple question treatment perceived that the time they had was less sufficient. Second, there was an interaction effect for mean idea quality, indicating that the average quality of their ideas was slightly lower (as predicted by Karau and Kelly [1992] for groups that experience time pressure). The conclusion is that this approach may have produced a greater entrainment effect, but still had little effect on the performance. Thus primary contribution to performance lies in task structure, not time structure. Merely separating the ideas into the three separate topic pools is sufficient to induce participants to allocate their effort more evenly; placing additional time restrictions does not improve performance.
These results, of course, are contingent on the choice of "appropriate" structures. In the interest of experimental control, only one form of structure was used: the subcategories and time periods selected by the researchers—three questions and three equal 10-minute time periods. It is not known how the results would have differed if each group (or each member individually) had been permitted to build their own subcategories or to select their own time periods. It may be that managers and team members (and perhaps even students) familiar with tasks can select appropriate decompositions for the task. Perhaps including more subcategories would have increased the number of ideas. Perhaps permitting each subject to individually choose their own structures might have improved performance by accentuating different parts of the problem for different group members to consider. The danger in permitting individuals to select their own structures is that they may unintentionally omit parts of the task (Connolly et al. 1993; Gettys et al. 1987; Pitz et al. 1980) or may expand the task to include irrelevant aspects. The ideal approach may be to have the group first generate ideas on possible decompositions for the task to ensure that no task elements are missed. Once this decomposition list is agreed on, the group can then generate ideas within each category.

One question for future research is how far structure can usefully go. Would the structure yield further gains if "air quality" was further subdivided into "automobile emissions," "factory emissions," and so on. Further structures are obviously possible, virtually indefinitely. Presumably there is some level of structure beyond which performance degrades with further structure. Exactly how much structure is useful clearly depends upon the group and the task, but the key point is that some moderate level of structure produced better results than none.

The anticipated effects from time structure did not occur. There were no significant effects from dividing the time available into separate work periods, except that subjects working on the single question task in multiple time periods were less satisfied. This suggests that manipulating time in this way has little effect. Unlike the original study (Dennis et al. 1996b), the subjects in this study were not forced to allocate their time to 10 minutes per question, and, unlike that study, there were no differences here in perceptions of sufficient time. Subjects in all treatments had similar perceptions about having enough time. We believe that our multiple time treatment was not strong enough to induce the anticipated entrainment effect. Perhaps shorter time periods might have strengthened the effect. ay be only noticeable when subjects are permitted only to work on one part of the task in a given time period.

This study suffers from the usual limitations of laboratory experiments (see McGrath 1982). The task was comparatively simple in that it required 30 minutes of work. The subjects here were ad hoc groups of undergraduate students with little intrinsic motivation and little prior GSS or business experience. It is possible that this form of task structuring only helps inexperienced users because those experienced with the task may be less likely to overlook parts of the task. However, the earlier study found the same effects whether studying undergraduates inexperienced with the task or managers working on a task with which they had prior experience. Several other issues relating to the use of laboratory experiments in GSS research have been noted (Dennis et al. 1991), but to date there is no empirical evidence to suggest that lab studies with undergraduates produce different conclusions than similar studies with managers (Dennis et al. 1996a).

Without presuming that all issues about the use of undergraduates as subjects have been addressed, it is believed that these results, which are similar to prior studies that have used both students and managers (Dennis et al. 1996b), have a message for managers and practitioners performing GSS-supported brainstorming. If the task offers an opportunity for modest structure, it should be taken. In this study, groups that used structure produced about 40% more ideas than those that did not. An effect of this size can be usefully employed in practical applications while research probes more deeply into the limits of its effectiveness and develops new techniques with even greater potential.
Acknowledgement
We would like to thank the associate editor for helpful comments on previous versions of this paper.

References


**About the Authors**

Alan R. Dennis is an associate professor of management in the Terry College of Business at the University of Georgia. He received a bachelor of computer science from Acadia University, an MBA from Queen's University, and a Ph.D. in MIS from the University of Arizona. His research interests include the use of groupware to support brainstorming and decision making. His past research has appeared in *Information Systems Research*, *Academy of Management Journal*, *Management Science*, *Communications of the ACM*, and *Journal of Applied Psychology*. His most recent book is *Business Data Communications and Networking* (6th edition) with Jerry FitzGerald.

Jay E. Aronson is an associate professor of management in the MIS group in the Terry College of Business at the University of Georgia. He earned his Ph.D. in industrial administration from Carnegie Mellon University. He has published in *Management Science*, *Decision Sciences*, *Information Systems Research*, *The Journal of Information Technology*, *Naval Research Logistic*, and the *Journal of Regional Economics* on group support systems, electronic brainstorming, large-scale modeling, network optimization, and cluster analysis. His most recent book is *Decision Support Systems and Intelligent Systems* (5th ed.) with Efraim Turban.

William G. Heninger is an assistant professor in the Department of Accounting & Law at the State University of New York at Buffalo. He received a bachelor of science in accounting, a master of accountancy from Brigham Young University and a Ph.D. in accounting from the University of Georgia. His research interests include group support systems, electronic brainstorming, auditor litigation, and earnings management.

Edward D. Walker II, CPIM, is an assistant professor in the College of Business Administration at Georgia Southern University. Prior to earning his Ph.D. in operation management at the University of Georgia, Dr. Walker worked in production planning and control, distribution, and plant management. His research interests include theory of constraints, project management, manufacturing planning and control systems, and performance measurement.
APPENDIX

Satisfaction

How do you feel about the process by which you generated ideas?

<table>
<thead>
<tr>
<th>Very Dissatisfied</th>
<th>Neutral</th>
<th>Very Satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

How do you feel about the ideas proposed?

<table>
<thead>
<tr>
<th>Very Dissatisfied</th>
<th>Neutral</th>
<th>Very Satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

How satisfied were you with the computer application your group used to discuss this problem?

<table>
<thead>
<tr>
<th>Very Dissatisfied</th>
<th>Neutral</th>
<th>Very Satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

All in all, how satisfied are you with being a member of this group?

<table>
<thead>
<tr>
<th>Very Dissatisfied</th>
<th>Neutral</th>
<th>Very Satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Overall, how enjoyable did you find your experience in this group?

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Neutral/Enjoyable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Time

For this idea generation session, did you:

<table>
<thead>
<tr>
<th>Have as much time as you needed</th>
<th>Neutral/Undecided</th>
<th>Want more time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Did you have enough time to review other comments and ideas?

<table>
<thead>
<tr>
<th>Ample Time</th>
<th>Neutral/Undecided</th>
<th>Inadequate Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Did you feel rushed in recording your ideas?

<table>
<thead>
<tr>
<th>Felt Rushed</th>
<th>Neutral/Undecided</th>
<th>No Rush</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Considering all the ideas you thought of, did you:

<table>
<thead>
<tr>
<th>Have time to express all your ideas</th>
<th>Neutral/Undecided</th>
<th>Not have time to express all ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

* Indicates reversed scored items.