

# Spatially Distributed Instructions Improve Learning Outcomes and Efficiency

Jooyoung Jang, Christian D. Schunn, and Timothy J. Nokes  
University of Pittsburgh

Learning requires applying limited working memory and attentional resources to intrinsic, germane, and extraneous aspects of the learning task. To reduce the especially undesirable extraneous load aspects of learning environments, cognitive load theorists suggest that spatially integrated learning materials should be used instead of spatially separated materials, thereby reducing the split-attention effect (Sweller & Chandler, 1994). Recent work, however, has suggested a new distinction between two common formats of spatially separated displays: spatially distributed versus spatially stacked (Jang & Schunn, 2010). Moreover, a distinction between instructions and learning task materials has rarely been made. Across two studies with 106 college students (56 in Study 1 and 50 in Study 2), we compared spatially distributed (multiple sources of information are placed side by side) versus spatially stacked (only one at the top is visible) instructions, without changing the learning task materials, on both task performance and learning. With materials more typical of practice, Study 1 showed that the distributed-display instructions led learners to more efficient learning; learners finished the task faster and scored higher in the overall learning test. With materials more tightly controlled for spatial format per se, Study 2 replicated the effect and found that the benefit of the distributed instructions appeared to be associated with changes in cognitive load. Implications for educational practice are discussed.

*Keywords:* cognitive load theory, split-attention effect, instruction design

Learning usually takes place through interactions with multiple sources of information. Moreover, learning materials commonly involve a mixture of textual and pictorial information, which has spurred research on various types of graphs and tables and their effectiveness under different circumstances (Cleveland & McGill, 1985; Larkin & Simon, 1987; Peebles & Cheng, 2003; Shah & Hoeffner, 2002; Shah, Mayer, & Hegarty, 1999; Zacks & Tversky, 1999). Determining the optimal type of representation (i.e., tables, graphs, and diagrams) for a given learning content is an important strand of research in instructional design. But because of the volume of information to process, improving learning often also requires optimizing the spatial arrangement of information, that is, spatially organizing multiple sources of information to support learnability and problem solving. In this article, we investigate relations between spatial organization of materials and learning from the perspective of cognitive load theory.

According to cognitive load theory (Sweller & Chandler, 1994; Sweller, van Merriënboer, & Paas, 1998; van Merriënboer & Sweller, 2005), learning is defined as a process of schema construction and automation, which takes place within the human

memory system. Because working memory capacity is severely limited in the number of items that can be stored and processed simultaneously (Baddeley, 1992; Miller, 1956), learning materials that present too much new information overload working memory, thereby reducing learning.

Building on the notion of limited working memory resources, cognitive load theorists posit three types of cognitive load that are influential in the process of learning: intrinsic, germane, and extraneous loads (van Merriënboer & Sweller, 2005). An intrinsic load is the type of load that is required for learning itself: Conscious processing of information for learning demands more cognitive resources than does automatic processing, and simultaneous processing of multiple items (i.e., complex tasks that require integration of information to solve the problem) consumes more resources than does independent, sequential processing with smaller amounts of information. In this conceptualization, intrinsic load is caused by the nature of the learning tasks and cannot be alleviated by changes in material design.

By contrast, germane and extraneous loads are the types of load that are not required for task performance and learning but instead involve additional load caused by the way materials are presented. When this type of load facilitates learning by encouraging learners to actively engage in the process of schema construction, it is called germane (Sweller et al., 1998; van Gog, Paas, & van Merriënboer, 2008). For example, giving a fully worked out example may not be as effective as having students fill in the blanks of the worked example. Incomplete examples draw learners' attention and help them actively participate in the cognitive processes of learning.

By contrast, extraneous load hinders learning by imposing an unnecessary load that is not related to learning or a necessary component of the core task being performed. Extraneous load is

---

This article was published Online First January 24, 2011.

Jooyoung Jang, Christian D. Schunn, and Timothy J. Nokes, Department of Psychology, University of Pittsburgh.

This research was supported by National Science Foundation Grant SBE-0823628. We thank the lab teaching fellows—Daniel Belenky, Soniya Gadgil, Melissa Patchan, Nikole Patson, and Elisabeth Ploran—for their assistance in these studies.

Correspondence concerning this article should be addressed to Jooyoung Jang, Department of Psychology, University of Pittsburgh, 823 Learning Research and Development Center, 3939 O'Hara Street, Pittsburgh, PA 15260. E-mail: joj15@pitt.edu

thought to be especially harmful when intrinsic and/or germane load is high because the different load types consume the same cognitive resources, and problems occur when total cognitive capacity is exceeded. Several factors were previously found to be relevant to extraneous load (Sweller et al., 1998). Among these factors, the split-attention effect is the factor that is directly related to the issue of spatial organization of learning materials. In particular, the theory suggests that materials should be prepared in a spatially integrated format, so that learners will not incur unnecessary processing costs from splitting and coordinating attention across spatial locations (Chandler & Sweller, 1991, 1992, 1996). As illustrated in the middle-left cell of Table 1, learning materials are commonly presented in a separated format. When the learning components in the multiple sources are independent from one another (i.e., are nonintegrative tasks), a separated display may suffice or perhaps may even be the best format. In contrast, when learning components in the multiple sources are closely related to one another and meaningful only when taken together (i.e., are integrative tasks), the separated display hinders learning (Chandler & Sweller, 1992). Therefore, when two learning materials that should be combined are split over space, mental integration of the two sources is required for learning. With separated displays, the learner's attention is split across two different locations, requiring him or her to search and match information, thereby consuming cognitive resources to coordinate this information (i.e., the split-attention effect). A meta-analysis by Ginns (2006) confirmed that task interactivity moderated the split-attention effect: Low element interactivity materials (i.e., nonintegrative tasks) were less susceptible to the cognitive load of split attention than were high element interactivity materials.

A simple solution for reducing the split-attention effect is to provide a single integrated display (see Table 1, the top-left cell); presenting multiple sources of information in an integrated format removes the burdens associated with search and matching corresponding elements. The advantage of spatially integrated displays for learning has been demonstrated in many studies (Chandler & Sweller, 1991, 1992, 1996; Mayer, 1989; Mayer & Gallini, 1990; Sweller, Chandler, Tierney, & Cooper, 1990; Tarmizi & Sweller, 1988; Ward & Sweller, 1990). For example, studies in several different domains (e.g., geometry, algebra, kinematics) found that worked examples were effective only when the text and accompanying diagrams were presented in an integrated format (Tarmizi & Sweller, 1988; Ward & Sweller, 1990). Ward and Sweller (1990) compared conventional problems, conventional worked examples, and integrated worked examples in the domain of kinematics (e.g., linear motion, projectile motion, and collision problems). As expected, students who were presented with homework in the integrated worked examples format outperformed the other two groups in the mean number of test problem solved and median error rate, but there was no difference in student performance between the conventional problem and conventional worked example groups. Although replacing conventional problem-solving activities with worked examples is often found to be beneficial for learning, the advantage has only been observed when an integrated format of materials was provided. This small difference in the display (separated vs. integrated) was powerful enough to eliminate the advantage of carefully studied worked examples on learning.

### Considering Another Variation of Spatial Layout

Although the benefit of a single integrated display over separated displays has been demonstrated in many different settings, a single integrative display may not be possible or even ideal in many situations. First, learners do not choose integrated versus distributed formats of display; learners can only work with the materials provided to them. Thus, this factor is typically not actionable by learners themselves. By contrast, we consider a factor that is actionable by the learners themselves. Second, as the number of sources needing to be combined increases, a single integrative display can become overly cluttered, which in turn will produce longer searching and parsing time for critical pieces of information (Wickens & Carswell, 1995). When the size of the task is especially large, which requires monitoring and conducting complex multiple steps (e.g., aviation and weather forecasting), a single integrated display simply cannot be constructed for the task (e.g., showing temperature, wind velocity, humidity, and pressure variables distributed over three spatial dimensions and over time). Third, there are many tasks in both formal and informal learning environments that have the instructions separate from the learning task (e.g., learning how to use a statistics program or learning how to assemble a chair from IKEA). This kind of separation is a widely applicable kind of complication to consider.

Thus, there is a need to find methods that both learners and instructors can use to optimize the organization of disparate sources using simple manipulations and without producing overly cluttered displays. Such methods would be of practical value (e.g., would enable learners to realize considerable benefits from simple but effective manipulations) as well as of theoretical value (e.g., what is the implication of such formats in terms of information processing, perception, and cognitive load theory?). Previous studies (Jang, 2009; Jang & Schunn, 2010) pointed toward one simple feature of separated displays (i.e., spatially distributed vs. stacked) as a promising candidate that can satisfy both needs. The purpose of the current study is to compare two types of separated display (i.e., spatially distributed vs. spatially stacked) and examine their effectiveness in task performance and learning.

As discussed above, separated displays are often considered negative contrasts to integrative displays: Separated displays, in contrast to integrative displays, have been observed to hinder learning. However, separated displays can take different forms for the learner, and these different forms have not been systematically studied. As shown in Table 1, even when multiple sources of information are presented separately, the sources can be displayed either side by side (i.e., spatially distributed; see the middle cells of Table 1) or sitting on top of one another with only the top source fully visible (i.e., spatially stacked; see the bottom cells of Table 1). Although the difference between the two displays may seem trivial, previous studies have found a sizable benefit of spatially distributed displays on integrative problem solving. A series of studies of integrative problem-solving tasks (Jang, 2009; Jang & Schunn, 2010) demonstrated that spatially distributed task materials, compared with a spatially stacked display, produced far shorter task completion times with no loss of accuracy. But the effect only occurred for tasks that required integration of information, consistent with cognitive load theory. For example, when participants solved a nonintegrative task (symbol arithmetic), the spatial organization of task materials had no effect on task perfor-



mance. The effect has been replicated using different media (computer and paper), which suggests that the effect is quite general (Jang & Schunn, 2010). Note that increasing use of computer displays (and increasing use of laptops) tends to push complex tasks into stacked formats because the available screen real estate tends to be smaller than what a problem solver has available in the physical world (e.g., he or she can spread out a stack of paper on a desk or pin multiple papers up on a wall; even compare the space available on a small laptop screen, iPhone, or iPad relative to a larger desktop monitor).

The benefit of spatially distributed task materials over stacked displays has also been reported in studies of multimedia learning (Olive, Rouet, François, & Zampa, 2008; Wiley, 2001). Olive et al. (2008) tested the effect of display format on a summarizing activity and found that the alternative display—corresponding to the stacked display in our framework—of source documents and notepad increased the cognitive load of the summarizing activity. While participants wrote a summary, they were also asked to perform a secondary task (i.e., responding to randomly spaced auditory tones). Participants in the stacked display condition took longer to respond to the tones than did those in the distributed display condition, suggesting that a heavier cognitive load is associated with the stacked display. Also, longer summarizing task completion time was observed in the stacked display condition. As an explanation of the effect, the authors suggested that people in the stacked display condition spent more time and effort memorizing the source documents because they knew that the source documents would no longer be available when they were writing on a notepad.

Similar to Olive et al. (2008) but more closely related to learning, Wiley (2001) investigated how well people understand provided learning materials using either a two-window browser design (the distributed display) or a single browser design (the stacked display). Wiley found that the two-window browser design led students to a better conceptual understanding of historical texts. For example, a larger number of inferential questions were answered correctly when the students used the two-window browser. Wiley hypothesized that the two-window browser design might have promoted integration across sources. However, we explore other ways in which spatial distribution becomes important in learning tasks.

### **Toward an Explanation for a Distributed Versus Stacked Display Effect**

Despite two recent demonstrations of the benefit of the distributed display relative to the stacked display, its generality, theoretical value, and explanations for the effect are currently incomplete. Starting from the perceptual level with vision and information processing research, the advantage of having a spatially distributed display over a stacked display at first inspection disagrees with what is currently known: The distributed and stacked displays are functionally equivalent at the perceptual level because people can only see information in a very small area at a time, about 1° of their field of vision. For detailed information to be perceived and accurately processed, it must fall within 1° of the center of the fovea. People can barely detect detailed visual content in the region outside of the parafovea, that is, farther than 5° from the center of the fovea (De Valois & De Valois, 1988; Hirsch &

Curcio, 1989). As a common example that demonstrates how small the area is, when you concentrate hard on the *o* in the middle of the nonsense word *cbneuoavaeuacs* and do not let your eyes wander, you will notice that reading the fifth letter to the right of *o* is impossible. Therefore, if one considers this feature of visual perception, a spatially distributed display does not provide an immediate benefit to information processing because only a small piece of information in only one source can be processed at a time. Thus, explanations for the effect must be sought at another level: working memory and cognitive load theory.

One could argue that the interplay of rapid eye movements and working memory are used to overcome the limitations of visual perception (Kong, Schunn, & Wallstrom, 2010). However, working memory is also very limited, which is the primary assumption of cognitive load theory. In situating the benefit of spatially distributed displays in cognitive load theory, the most plausible hypothesis would be that the degree of split attention is relatively higher in the spatially stacked display than in the spatially distributed one as a result of greater search costs and attention management issues. But there have been no empirical investigations of whether the split-attention effect is binary in nature (as the integrated vs. separated framework suggests) or whether there are degrees of effects as access costs increase (e.g., from displays that are side by side or on top of each other). Further, it is possible that effects of spatial format are due to factors other than the direct effects of cognitive load per se (e.g., due to changes in strategies in response to the load rather than purely reduced accuracy and slowed reaction times that comes from load increases directly). To clarify this theoretical piece, it is necessary to measure the cognitive load across distributed or stacked displays (explored in Study 2).

For both practical and theoretical purposes, it is useful to map the extent of the benefit across variations of information being distributed and dependent measures. Whereas Jang and Schunn (2010) measured the effect of display layout on task performance (i.e., task completion time and accuracy), Wiley (2001) measured degree of learning (i.e., proportion of sentences in student essays showing deeper understanding of the presented information, performance in inference verification, and analogical thinking) as outcomes. Thus, spatial layout appears to affect both problem-solving efficiency and learning outcomes. In settings in which learners spend significant amounts of time solving problems, effects on problem-solving efficiency can be thought to improve learning efficiency as well, because learners can then do more problems per unit of time. However, to properly claim a facilitated learning effect, the combination of improved learning outcomes and improved learning efficiency has to be verified within a single learning study, as several previous studies have done while studying other effects (Cerpa, Chandler, & Sweller, 1996; Chandler & Sweller, 1991; Kalyuga, Chandler, & Sweller, 1998; Martin-Michiellot & Mendelsohn, 2000; Sweller & Chandler, 1994; Sweller et al., 1990). It is possible that learners with spatially distributed rather than stacked displays learn more in a fixed amount of time, learn the same amount in less time, or learn more in less time. We attempt to clarify this ambiguity in the current study.

Another issue is the content of what is spatially distributed or stacked. Both the Wiley (2001) study and the Jang and Schunn (2010) study manipulated the spatial layout of task materials (i.e.,

materials that are used to solve given problems). However, in a wide array of educational settings, the instructions for a task are often given separately from the learning task materials, especially for complex procedural tasks (e.g., conducting statistical analysis using a computer program). Cognitive load theorists commonly describe all learning materials as instructional materials, but we think the spatial layout issues with respect to information integration warrant separate consideration for instructions and learning task materials. Cognitive load theorists have examined the effectiveness of the integrated display across different learning phases: initial instructional period, acquisition phase (studying worked examples and/or practice problems), and test phase (Sweller et al., 1990). But the comparison was made in a temporal manner (i.e., during which phase an integrated display was used), not in terms of content. In some tasks in domains such as geometry or algebra, there is not much difference between instructions in the form of worked example problems and learning task materials (i.e., new problems to solve). However, when a task involves acquiring procedural skills as well as learning-related concepts, such as with learning statistics, the distinction between instructions and learning task materials can be more clearly made.

Statistics is an ideal example of an integrative task with separate instructions that is widely taught across college classrooms. It is clearly an integrative learning task: Acquiring knowledge and skills of data analysis and interpretation often involves “learning many different ‘pieces’ of knowledge and integrating them in some unified whole” (Lovett & Greenhouse, 2000, p. 197). At the same time, instructions to the learner are presented separately from the learning task environment—for example, instructions are on paper, and the learning task is on the computer in Excel or a statistical package. Therefore, statistics was chosen as the domain of learning to study. As depicted in Figure 1, in statistics learning, instructions are commonly step-by-step textual guidelines and relevant screen shots for conducting a series of statistical testing, and learning task materials are data files and spreadsheet or statistical packages such as Excel and SPSS.

Relative to the problem-solving task per se, these instructions are separate and thus task performance might not be influenced by

the relative spatial layout of instructions. However, the real task of the learner is a learning task (i.e., learning to solve the given problem), and for this task, the spatial layout of the instructions is relevant. In the current study, we examine the format of instructions, a factor that should also matter for learning but has not been previously examined. This is a somewhat curious state of affairs given how ubiquitous complex procedural tasks are taught through instructions separate from the learning task itself. We examine its effect on both task performance and learning.

## Study 1

### Method

**Participants.** Participants consisted of 56 undergraduates (38 women) enrolled across three lab sections of an introductory cognitive psychology course at the University of Pittsburgh. All students had previously completed prerequisite courses in psychology research methods and introductory statistics. However, these prior courses were at a relatively general and basic level and did not cover the more detailed data analysis practices associated with cognitive psychology experiments. The cognitive psychology labs allowed students to develop an extended and more integrated schema for experimental design, data harvesting, data organization, exploratory data analysis, and inferential statistics as they commonly apply to cognitive psychology lab studies.

**Design.** The experiment consisted of three phases—background, practice, and testing—with students randomly assigned to one of two conditions during the practice phase. The background and practice phases took place during the normally scheduled lab activities. Students learned and practiced how to analyze independent two-sample data using common data organization and statistic tools (i.e., Excel and SPSS). The testing phase took place during a normally scheduled data analysis review session.

Each student within each of the three lab sections was randomly assigned to one of two different conditions applied during the practice phase: distributed format instructions or stacked format instructions. Dependent variables included data analysis task performance (i.e., time and accuracy during the practice phase, as indices of learning efficiency) and learning outcomes (i.e., test scores during the testing phases, as indices of acquired procedures).

**Materials.** Students worked individually at computer workstations in the lab for the background and practice phases. During the background phase, detailed handouts were provided for four main steps of the taught data analysis procedure: (a) organize the data, (b) create a pivot table, (c) create a graph, and (d) run a *t* test. The core skills to learn in the first step were to count the number of men and women in the data set, add a new column for the derived dependent variable, and calculate the derived dependent variable (in this case, total number of correct responses). In the second step, students learned how to use the pivot table function in Excel: where to enter independent and dependent variables and how to calculate means, standard deviations, and sample sizes within the pivot table. In the third step, the skills to acquire were calculating standard errors and effect sizes using Excel formulas and creating a graph with standard error bars. Finally, in the fourth step, students learned how to export data from Excel to SPSS, run an independent-samples *t* test, and interpret outputs. Text instructions and accompanying screen shots were tailor-made to fit the

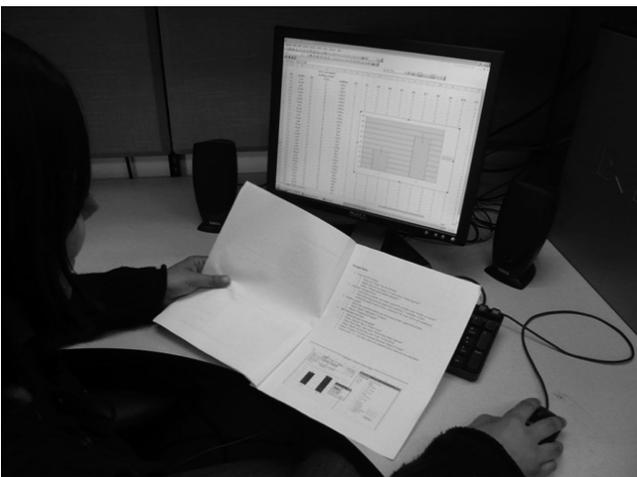


Figure 1. Distinction between instructions and learning task materials in statistics education.

practice data provided in the background phase and distributed to students in the form of stapled letter-size documents. Two practice data sets were provided to the students, one for the in-class learning activity and the other for homework. The data sets were essentially the same in terms of structure but had different numbers and cover stories.

During the practice phase, half of the students received the relevant instructional handout in distributed format and half received the handout in stacked format. The instructions were less detailed than those provided in the background phase. However, core information—for example, abstract instructions on what to do next and all complex equations—was included. For the distributed-format condition, instructional text and accompanying screenshots were presented on an 11- × 17-in. paper, with four sources of information laid side by side on a single side of the paper (i.e., four separate panels of information). For the stacked-format condition, the same information was presented in four pages of letter-size paper stapled together (see Table 1). The selection of particular document sizes for the two conditions was based on ecological validity. U.S. letter size, 8 1/2 × 11 in. (or the slightly larger A4 size used in many other countries), is the typical format in which lab instructions are printed for students. By contrast, U.S. ledger size, 11 × 17 in. (or the closely matched metric A3 size), is often the largest paper size on which instructors could conveniently print out instructions, and it conveniently folds to U.S. letter size for inclusion in binders. As is also likely to be typical practice, if an instructor were to use ledger paper to include all instructions for a lab on one sheet, there would be some compression of images or text. Thus, our separate format condition included the content of four letter-sized pages on one ledger-size page through some compression of images and text and some deletion of white space.

To measure task time and accuracy during the practice phase, we provided a task worksheet, which (a) asked students to record the start time of each step, (b) briefly mentioned which instructional step to complete next, (c) asked students to report a few requested results (e.g., numbers of men and women who provided the data, means and standard deviations of a condition, the effect size, and *t* test results), and (d) asked students to record the end time of each step. The task worksheet consisted of four pages, a page for each main analysis step.

To keep track of help requests from students, which may account for effects of condition on task time and accuracy, we asked lab instructors to record the frequency of help requests from each student during the practice phase. A table was provided to instructors to record the name of the students asking for help and the number of help requests. During the testing phase, to measure overall learning outcomes, we had participants complete a 14-item closed-book test (10 multiple-choice and four short-answer questions; Cronbach's  $\alpha = .41$ ), which contained three types of questions—factual, conceptual, and integrative questions—that were developed to closely match the broad instructional goals of this unit. There were four factual questions for which the answers could be found directly in the given instruction sheets, seven conceptual questions for which the answers should be inferred with some critical thinking about a given step in the instruction sheets, and three integrative questions for which the answers could only be obtained through integration across multiple steps. Sample items for each type of question are presented in Figure 2. The modest overall alpha reflects the diversity of concepts that were

<b>Factual question</b>
What is the equation (excel expression) to calculate standard error? a. =([StdDev of a condition]/[Count of the condition]) b. =([StdDev of a condition]/sqrt([Count of the condition])) * c. =[Average of a condition]/[StdDev of the condition] d. =[Average of a condition]/sqrt([StdDev of the condition])
<b>Conceptual question</b>
A bar graph should be used when the a. Independent variable is a discrete variable * b. Independent variable is a continuous variable c. Dependent variable is a discrete variable d. Dependent variable is a continuous variable
<b>Integrative question</b>
You created a descriptive statistics table (i.e. a pivot table) on a data using excel. However, you just realize that there was one more participant that you have not coded into the excel sheet by mistake. After you fix the problem, a new table has been created. What would be the easiest way to check whether the new data point is now counted in or not? a. check the Count for a condition b. check the Average for a condition c. check the Count for the Grand Total * d. check the StdDev for a condition

Figure 2. Sample items for each of three test question types. The asterisk indicates the correct answer.

being tested. To fit within the constraints of the lab, the overall test was relatively short and certainly too short to produce reliable submeasures (e.g., of factual vs. conceptual learning).

**Procedure.** For the background phase, over two consecutive labs that were one week apart (1.5 hr each, 3 hr total), students learned how to analyze data from a one-factor study. They learned four main steps of data analysis (i.e., organize data, create a pivot table, draw a graph, and run a *t* test), two per lab. The first three steps were done in Excel and the last step was done in SPSS. Detailed handouts for each step were provided and lab instructors walked through each step with students during the lab. Also, homework was assigned to allow students to begin to practice each step on their own. Students could ask questions or request help from instructors at any time during the two weeks of the background phase.

The practice phase occurred in the third lab. Students were randomly assigned (within each lab section) to one of the two conditions: distributed- versus stacked-format instructions. Students were asked to analyze new data from a one-factor study, but within a maximum time of 40 min and on their own. Also, they were not allowed to refer to the detailed handouts they had previously received. In other words, they were only allowed to use the less detailed handout given that day, in either distributed or stacked format. While completing this data analysis task, they filled out a task worksheet, which asked students to self-report start time, results, and end time. Students who finished the task early turned in handouts and task worksheets and were allowed to quietly engage in any other activity while the other students completed the task. After the 40 min of data analysis activity, all students took the overall learning test for 25 min.

From the practice phase, the main dependent measures were task time (i.e., time duration for the sum of the four time steps), task accuracy (i.e., the percentage of correctly reported analysis results across the four steps), and the frequency of help requests from instructors. From the testing phase, dependent variables were the learning outcomes (i.e., the percentage of correctly answered items on the overall test).

## Results and Discussion

Six students were excluded from the analysis: One did not follow experimental instructions, three were absent during either or both of the first two lab sessions (i.e., background phase), and an additional two students scored more than two standard deviations below the mean in the overall learning test (one from the distributed-format condition and the other from the stacked-format condition). This left 50 participants in total (25 in the distributed-format condition and 25 in the stacked-format condition).

**Task performance: Measures from the practice phase.** The full set of dependent measures consisted of the time to complete the task, the percentage of task accuracy, overall learning outcomes, and the frequency of help requests. First, a one-way multivariate analysis of variance (MANOVA) was conducted to determine the effect of the two types of instructions (distributed vs. stacked) on the four dependent variables. A nonsignificant Box's  $M$  ( $p = .06$ ) indicated no violation of the homogeneity of variance-covariance matrix assumption. Significant differences were found between the two formats of instruction on the dependent measures, Wilks's  $\Lambda = .64$ ,  $F(4, 45) = 6.27$ ,  $p < .001$ . Univariate tests indicated significant differences on task time, help requests, and test scores<sup>1</sup> (the first two are discussed in this section and test score is described in the next section as it is a measure from testing phase; see Figure 3): Students in the distributed-format condition, compared with those in the stacked-format condition, completed the overall data analysis task 6.6 min faster,  $F(1, 48) = 14.24$ , mean square error ( $MSE$ ) = 544.5,  $p < .001$ ,  $d = 1.07$ , at the same high level of overall accuracy,  $F(1, 48) = 0.04$ ,  $MSE = 6.23$ ,  $p = .85$ ; that is, there was no sign of a time versus accuracy trade-off. Also, compared with the distributed-format condition, students in the stacked-format condition asked for help significantly more often,  $F(1, 48) = 11.95$ ,  $MSE = 4.74$ ,  $p < .001$ ,  $d = 1.03$ . Means and standard deviations of task time and task accuracy on each step can be found in Appendix A.

**Learning outcome effects: Measures from the testing phase.** Even though students in the distributed-format condition spent less time than those in the stacked-format condition completing the task, they scored significantly higher in the overall learning outcome,  $F(1, 48) = 4.84$ ,  $MSE = 976.8$ ,  $p = .03$ ,  $d = 0.63$ . This difference was sizable (almost 10%), with students moving from a mid-C to a mid-B in terms of grades (see Figure 3). All three question types showed a similar tendency of better learning with distributed versus stacked formats. However, a formal statistical analysis was considered inappropriate because there were too few items in each section to produce reliable submeasures. Means and standard deviations can be found in Appendix A.

In sum, a distributed format of instructions led to much shorter practice task completion times with no loss of accuracy, which is an index of efficient learning. Moreover, the benefit of distributed-format instructions appeared in high learning outcome test scores and a smaller number of getting stuck events while completing the practice task (i.e., low help request frequency), which are additional indices of efficient learning.

**Correlational analyses.** To examine whether there were independent, counteracting, or mediated effects of instructional format on each of the dependent variables, we examined correlations among the measures. As shown in Table 2, two significant correlations were found. First, the task completion times were nega-

tively correlated with accuracy, which means that the longer it took students to complete the task, the less accurate they became. Thus, there was no evidence of a time and accuracy trade-off. Rather, the relationship was in the opposite direction, which is consistent with an overall learning efficiency effect of distributed-format instructions. Second, the task completion time positively correlated with the help request frequency, which is plausible, given that it will take some time to ask questions and get help.

Considering the significant difference in the frequency of help requests by the format of instructions and the significant correlation between the task time and the help request frequency, we conducted an analysis of covariance (ANCOVA) to test whether the effect of time difference holds up when the time to ask questions and get help is taken into account. The test of homogeneity of regression was found to be nonsignificant,  $F(1, 46) = 0.583$ ,  $MSE = 21.76$ ,  $p = .45$ , showing that the effects of help requests on total time are similar and linear in the two instructional conditions. The result showed that even considering the difference in help requests, the effect of time difference remained significant,  $F(1, 47) = 7.34$ ,  $MSE = 271.71$ ,  $p = .01$ . Adjusted means indicated that the 5 min of the unadjusted 7-min time effect remained when the help request factor was taken into account. Thus, the help request factor accounts for only a small fraction of the overall effect on time.

## Study 2

The results of Study 1 provided clear support for the benefit of spatially distributed instructions on both task performance and learning outcomes over a spatially stacked format. However, it is possible that the novelty per se of the format of instructions used in the spatially distributed condition may have caused the effect. Specifically, the size of the paper used in the distributed-format condition was larger than the conventional handouts students typically received in the lab. Thus, the unfamiliar distributed format could have engaged and motivated students in a way that familiar letter-size instructions would not. Also, because of the difference of physically available space in four pages of U.S. letter size 8-1/2- × 11-in. paper (i.e., the stacked-format condition) and one page of U.S. ledger size 11- × 17-in. paper (i.e., the distributed-format condition), the size of instruction screen shots were not exactly matched between conditions. To address these issues, we modified the materials of the stacked-format condition to take an unfamiliar format as well as match content exactly by cutting the distributed-format instructions into four pages of narrow paper. In this way, both formats were equivalently novel to students.

Study 2 also directly measures one factor typically associated with the split-attention effect, cognitive load. During the practice phase, cognitive load was measured by using a subjective rating scale. Among the three common cognitive load measures (i.e., subjective rating scale, dual-task paradigm, and physiological test), a subjective rating scale was used because of the constraints of the classroom setting and its well-documented validity and reliability

<sup>1</sup> Levene's tests of equality of error variances for each of dependent variables were, for overall task time,  $F(1, 48) = 0.32$ ,  $p = .57$ ; for overall task accuracy,  $F(1, 48) = 0.05$ ,  $p = .83$ ; for help request frequency,  $F(1, 48) = 0.04$ ,  $p = .84$ ; and, for overall test score,  $F(1, 48) = 0.36$ ,  $p = .55$ .

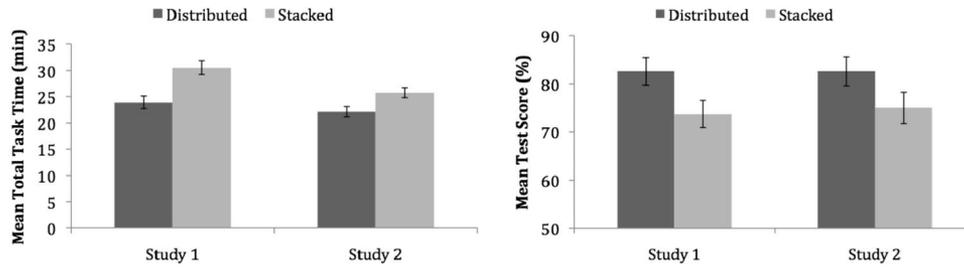


Figure 3. Mean total task completion time with standard error bars (left) and mean total test score with standard error bars (right) in Study 1 and 2 as a function of condition.

(Ayres, 2006; Marcus, Cooper, & Sweller, 1996; Paas, 1992; Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Paas & van Merriënboer, 1994).

## Method

**Participants.** Participants consisted of 50 undergraduates (34 women) enrolled across three lab sections of an introductory cognitive psychology course at the University of Pittsburgh.

**Design.** The overall design was the same as in Study 1, with a single-factor of format of instruction (i.e., distributed vs. stacked). In addition to the dependent measures from Study 1 (i.e., task time and accuracy, learning outcomes, and frequency of help requests), cognitive load measures were added in Study 2.

**Material.** The exact same materials were used during the background phase. During the practice phase, to eliminate the possibility of a novelty confound in one of the formats, we designed both formats of instruction to be unfamiliar and more precisely the same across conditions. The distributed format was left in the same novel format used in Study 1. The stacked format was changed from Study 1: Here it was created by cutting the distributed-format paper into four smaller pages ( $4 \times 11$  in.), and the four pages were stapled at the top (see Table 1).

The task worksheet was modified to include commonly used subjective rating scales of cognitive load, with the scales placed below instructions for each data analysis step. The scales consisted of two rating questions regarding perceived cognitive load (i.e., perceived difficulty and invested mental effort). Specifically, students were asked to rate how easy or difficult the step was and how much mental effort the step took on 9-point scales, which are commonly used for measuring cognitive load, ranging from 1 = *very, very easy* to 9 = *very, very difficult* and from 1 = *very, very low mental effort* to 9 = *very, very high mental effort* (Paas & van Merriënboer, 1994).

Table 2  
Intercorrelations Between Variables in Study 1

Variable	1	2	3	4
1. Task completion time (min)	—	-.29*	-.21	.40**
2. Task accuracy (%)		—	.04	.08
3. Overall learning outcome (%)			—	-.05
4. Help request frequency				—

Note.  $N = 50$ .

\*  $p < .05$ . \*\*  $p < .01$ .

As in Study 1, a short test was used to measure overall learning outcomes. However, a few floor, ceiling, and low item-to-total correlation questions were deleted from the original test and more questions were added to increase overall test reliability and validity, which led to a 22-item test (Cronbach's  $\alpha = .46$ ), composed of 18 multiple-choice and four short-answer questions: eight factual, eight conceptual, and six integrative questions. The low alphas across two tests in Study 1 and 2 presumably reflect the relatively short length and a diversity of independent knowledge and skills components that are being measured.

**Procedure.** The overall procedure was the same as in Study 1. However, because of the classroom time constraints, students were given 35 min to complete the practice task in Study 2, instead of 40 min. For the learning outcomes test, 25 min were given, the same as in Study 1.

## Results and Discussion

Two students scoring two standard deviations below the mean in the task accuracy (both from the stacked-format condition) and two students scoring two standard deviations below the mean in the task time (one from each condition) were excluded as outliers, which left 46 participants in total (23 in the distributed-format and 23 in the stacked-format conditions).

**Task performance: Measures from the practice phase.** As in Study 1, overall task time, overall task accuracy percentage, help request frequency, and overall learning outcomes were submitted to a MANOVA. The test of homogeneity of variance-covariance matrices was found to be nonsignificant ( $p = .51$ ). The overall effect of the format of instruction (i.e., distributed vs. stacked) on task time, accuracy, and learning outcomes was replicated, this time with equivalently novel formats.<sup>2</sup> As shown in Figure 3, students in the distributed-format condition completed the overall data analysis task 3.6 min faster,  $F(1, 44) = 6.62$ ,  $MSE = 149.76$ ,  $p = .01$ ,  $d = 0.76$ , at the same high level of overall accuracy,  $F(1, 44) = 0.35$ ,  $MSE = 48.02$ ,  $p = .56$  (no sign of a time vs. accuracy trade-off). Unlike the result of Study 1, there was no significant difference in the help request frequency by the format of instructions,  $F(1, 44) = 1.11$ ,  $MSE = 1.07$ ,  $p = .30$ , although there was a small difference in the means in the same

<sup>2</sup> Levene's tests of equality of error variances for each of dependent variables were, for overall task time,  $F(1, 44) = 0.01$ ,  $p = .95$ ; for overall task accuracy,  $F(1, 44) = 0.61$ ,  $p = .44$ ; for help request frequency,  $F(1, 44) = 1.96$ ,  $p = .17$ ; and, for overall test score,  $F(1, 44) = 4.18$ ,  $p = .05$ .

direction as Study 1. This pattern of results is consistent with the ANCOVA analyses of Study 1 that suggested the effects of condition on task time were not entirely explained by the effect of condition on help requests. A table of means and standard deviations for the full set of dependent measures can be found in Appendix B.

**Cognitive load: Measures from practice phase.** The average cognitive load rating (Cronbach's  $\alpha = .91$ ) was submitted to an analysis of variance (ANOVA). One additional student in the stacked-format condition who answered 1 on all of the 9-point scale questions was removed, which left 45 participants in total (23 in the distributed-format condition and 22 in the stacked-format condition). As predicted by cognitive load theory, there was a moderate effect of format on the cognitive load at the overall level,  $F(1, 43) = 4.06$ ,  $MSE = 5.51$ ,  $p = .05$ ,  $d = 0.61$ . Further, as shown in Figure 4, this general pattern of condition differences was found for all eight ratings: Students in the stacked-format condition generally reported a higher cognitive load.

**Learning outcome effects: Measures from the testing phase.** As in Study 1, students in the distributed-format condition, who spent less time completing the task, also scored significantly higher in the overall learning outcome,  $F(1, 44) = 5.16$ ,  $MSE = 720.09$ ,  $p = .03$ ,  $d = 0.67$ . Again, because of the small number of items per question type, analyses focusing on the effect of format on particular question types were unlikely to be informative.

In sum, the results of Study 2 were generally consistent with those of Study 1 in that the distributed format of instructions facilitated completing an integrative task (i.e., the data analysis task) more quickly while maintaining high accuracy and producing better learning outcomes across question types. The results of the cognitive load measures, new in Study 2, suggest that distributed format instructions might produce small differences in cognitive load compared with stacked instructions but that other underlying explanations may be required.

**Correlational analyses.** To unpack the relationship among outcomes, we calculated correlations between dependent measures. As shown in Table 3, several statistically significant correlations were found. First, overall task completion time was negatively correlated with overall learning outcome score, presumably because students who understood the instructions were better able to complete the task quickly and do well on the posttest. Although this relationship was not found to be statistically significant in Study 1, the correlation between the task completion time and the

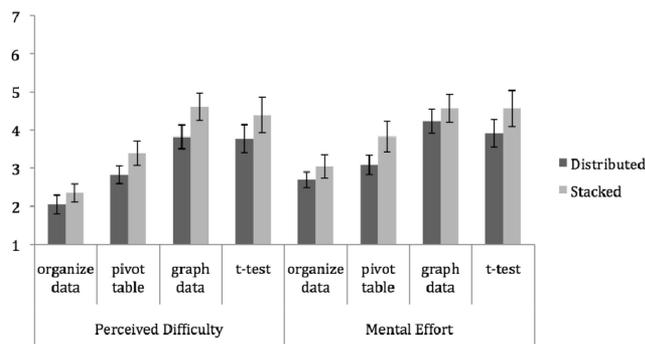


Figure 4. Mean perceived difficulty and mental effort across data analysis steps (on 1–9 scales) as a function of condition.

Table 3  
Intercorrelations Between Variables in Study 2

Variable	1	2	3	4	5
1. Task completion time (min)	—	.12	-.42**	.38**	.22
2. Task accuracy (%)		—	.01	.02	.01
3. Overall learning outcome (%)			—	-.51**	.01
4. Overall cognitive load				—	.12
5. Help request frequency					—

Note.  $N = 45$ .

\* $p < .05$ . \*\* $p < .01$ .

learning outcome was consistently negative across the studies. Second, across the two studies, there was no evidence of a time and accuracy trade-off that would have yielded a positive correlation between the task completion time and the task accuracy. In Study 1, a negative relationship was observed, which offers a strong negation of the trade-off effect, and in Study 2, no statistically significant relationship was observed, which consistently suggests that longer task completion time did not produce high accuracy. Third, overall task completion time correlated positively with the cognitive load measure of perceived difficulty. This is a plausible relationship: When a student understands a data analysis step, the step would not be perceived as difficult, nor would it take long to complete. Fourth, the two cognitive load measures correlated with each other, confirming that they were measuring the same construct.

Finally, the overall learning outcome score was negatively correlated with the cognitive load measures of perceived difficulty and mental effort. This is an interesting relationship that can be explained from the perspective of cognitive load theory. In a heavy extraneous load situation, students have to devote much of their limited working memory on the extraneous load (i.e., there is a loss of resources for learning) rather than purely on the intrinsic and germane load (i.e., actual learning). Thus, high extraneous cognitive load leads to high interference in working memory, which eventually produces negative effects on learning. To test the relationship further, we conducted an ANCOVA on learning outcomes as a function of condition (i.e., the format of instructions) with the cognitive load measures taken into account as a covariate. The test of the homogeneity of regression was found to be nonsignificant ( $p = .45$ ). The analysis found that the effect of the format of instructions on the test score became nonsignificant when the effect of the cognitive load measures was included,  $F(1, 43) = 3.08$ ,  $MSE = 382.08$ ,  $p = .09$ . Thus, we can conclude that the effect of the format of instructions appears to be at least partially mediated by the effect of cognitive load—the stacked format of instructions taxes the learner's working memory with a relatively high extraneous cognitive load, thereby hindering the process of learning.

## General Discussion

Studies 1 and 2 consistently demonstrated a large benefit of a spatially distributed display of instructions on task performance (time and accuracy) and learning outcomes. During the practice phase, students who received the distributed format of instructions finished analyzing data faster than did those who received the

stacked format of instructions, and there was no sign of a speed–accuracy trade-off, which suggests that students achieved the same amount in less time. Furthermore, students in the distributed-format condition scored higher on the overall test than did those in the stacked-format condition. These benefits appear to be at least partially attributable to the difference across conditions in the amount of cognitive load: Students in the stacked-format condition reported a higher level of cognitive load, and taking cognitive load into account reduced the effects of condition on learning outcomes. These data patterns are consistent with a cognitive load–based extension of the split-attention effect. Overall, the spatially distributed display of instructions appears to partially reduce load and considerably increase learning efficiency. But it may also be that factors other than cognitive load are relevant here.

### Theoretical and Practical Implications

As noted in the introduction, the effects of spatial distribution can be examined across different objects being distributed (most notably task vs. instructions) and on different dependent measures (most notably task performance, learning levels, and learning efficiency). The current research suggests that simply manipulating the spatial format of instructions, leaving learning task materials alone, may suffice to significantly improve both task performance and learning. Whether the different possible combinations across the formats of instructions and task materials (i.e., distributed instructions–distributed task materials, distributed instructions–stacked task materials, stacked instructions–distributed task materials, stacked instructions–stacked task materials) also lead to different effects should be addressed in future studies. Presumably, the combination of distributed instructions with distributed task materials would be the optimal pair relative to distributed representations of either instructions or task materials. Relevant to this hypothesis, Cerpa et al. (1996) compared conventional manual plus computer software to be learned (i.e., task materials and instructions were separated across screen and paper) against integrated computer-based training software (i.e., task materials and instructions were integrated on the screen). They manipulated both task materials and instructions and found that people with integrated onscreen presentation of task materials and instructions outperformed individuals using the conventional manual plus computer software as well as individuals using the conventional manual plus integrated training software. The current study adds to that work by showing that just the act of distributing instructions while holding task material format constant is enough to improve learning significantly. This result is very useful because there are many settings in which instruction materials can be easily changed but task materials cannot be easily changed.

In terms of theoretical implications for the split-attention effect, the results suggest that the effect appears not to be binary in nature (as the integrated vs. separated framework would suggest); rather, there appear to be degrees of effects as access costs increase (e.g., integrated displays < distributed displays < stacked displays along a dimension of access cost). Further, each step along the continuum appears to be associated with an increase in cognitive load.

But before accepting cognitive load as the underlying variable, it is important to consider other possible explanations of the distributed versus stacked effect. For example, perhaps some of

the time advantage is simply due to the time spent turning pages in the stacked-format condition, which is not a possible factor in the comparison of distributed versus integrated displays. Or, said another way, the stacked instructions are not only spatially separated but also temporally separated, which potentially introduces other factors like physical page flipping time. A previous study (Jang & Schunn, 2010) examined this issue directly (at least in terms of effects on problem-solving time) by counting page flips from video. They found that the time spent turning pages could explain only 30 s out of 6 min of total time difference between conditions. Applying this scheme to the current study, we would estimate that the time to turn pages would account for at most tens of seconds (4 pages  $\times$  less than 4 returns to prior pages  $\times$  1 s per page flip) out of the 7-min total time difference across conditions found in the current Study 1.

Another interesting theoretical issue is the relationship between learning from stacked displays (which provide information in a temporally separated manner) and learning from animations. Stacked displays and animations have some features in common in that learners can only see partial information at any given time and prior information sometimes disappears as new information appears. But they are also very different in several ways. First, learners with stacked displays can go back and forth between information pieces more efficiently than they can with animations. It is often hard to search and go back to a particular screen of information in animations; the information accessing cost may take several effortful clicks and time to wait and watch as the information unfolds. Second, stacked displays provide information indices that are harder to implement in animations. For example, in a stacked display, a group of information occupies a physical space, and this page may act as an index for retrieving information later as well as clearly bracketing separable packages of information. In the current study, the task consisted of four independent steps that each had complete instructions in a single page.

On the practical side, in environments that require extensive integration across multiple sources of information that could not be sensibly integrated into single displays (e.g., analyzing multidimensional statistical data, interpreting many slices of brain images, analyzing temporal changes in weather, integrating and deriving a coherent story for historical events), the results of this study and those of Olive et al. (2008) and Wiley (2001) suggested that the sources (instructions and task materials) should be placed side by side rather than stacked. For electronic materials, side-by-side placement may not be possible within typically available monitor sizes, but printouts frequently can be generated. For instructions, the medium is frequently paper, and paper sizes larger than 8.5  $\times$  11 in. are generally available. For example, in a study by Mayer, Steinhoff, Bower, and Mars (1995), students in the integrated material condition had text describing a lightning mechanism and corresponding illustrations on a facing pages, whereas the other group received text and illustrations in separate booklets. Thus, various forms of spatial distribution of instructions allow easy and handy manipulations that are applicable to many settings. When appropriate technology is available and a group of people is engaged in a learning or problem-solving task, having a large computer display (e.g., a projector or SMART Board) would also be effective.

## Caveats

Although the cognitive load measures indicated relatively higher load in the stacked format of instructions, this was but one measure of cognitive load that is a complex, multidimensional construct and a measurement taken at only one point in time for a multiphased procedure. A subjective measure of cognitive load was used because of the constraints of an experiment run in a classroom setting, but an objective measure (e.g., a dual-task paradigm) may have brought important converging evidence. Also, multiple measurements of cognitive load across learning and testing phases (van Gog & Paas, 2008) could have provided additional insights into the large effect of instruction format on problem solving and learning. Future studies should explore using the objective and/or more broadly distributed in time measures of workload.

Second, it is possible that differences in perceived difficulty and mental effort only partially explain the spatial distribution effect. To unpack this issue further, we begin with the notion of cognitive load and what exactly spatial distribution does to cognitive load. Having to find information across pages does not necessarily imply greater reliance on working memory. But the general theoretical treatment of cognitive load includes the loads associated with processing information and search is a kind of processing. For example, Ayres and Sweller (2005) argued that “it is that act of search that imposes an extraneous cognitive load” (p. 143). However, we argue that factors other than cognitive load may play a role in explaining effects of spatial distribution, such as (a) people may prefer to divert time and learning resources to memorizing information in displays when the information access cost of returning to that information is high (Fu & Gray, 2000), (b) it takes time to turn over pages or manipulate windows (as discussed earlier), and (c) keeping track of information pages may act as an additional task (Jang & Schunn, 2010; Wickens & McCarley, 2008). In particular, the first point, choosing an information encoding strategy that fits to information access cost, can provide an interesting interpretation for the phenomenon at hand. When information access cost is high, people may invest more time in encoding external information into long-term memory because information in their internal memory system is more accessible than information outside of it under those circumstances (Fu & Gray, 2000, 2006). Presumably, the information access cost of stacked displays would be relatively higher than that of distributed displays. In the stacked-format condition, a physical action of turning pages is required to access certain information, but in the distributed-format condition, only an eye movement or a head turn is asked for. Thus, with stacked instructions, learners may (implicitly) spend more time memorizing information to avoid frequent page-turning actions, which would produce longer task times and a greater load on memory, which could affect later memory retrieval for more conceptually important information that is buried with temporary task details. These alternative factors should be examined in greater depth in future studies.

## Future Work

Our work suggests that spatially distributed instructions lead to efficient and effective learning at least in part because they reduce cognitive load, specifically, by alleviating the burden of the split-

attention effect. But the current study did not take individual differences in skill level into account. Several studies investigating the split-attention and redundancy effects (Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga et al., 1998; Kalyuga, Chandler, & Sweller, 1999) suggest that what benefits novices may not do the same or may even do the opposite for experts. Calling it the *expertise reversal effect*, one study found that advanced learners (i.e., intermediate electricians) learned less with redundant information provided in integrated texts and diagrams and learned more from a diagram-only format (Kalyuga et al., 1998). That is, separated displays produce a split-attention effect for novices, but for experts, the integrated displays provide redundant information that cannot be avoided, labeled a redundancy effect. Because the expertise reversal effect was observed for the issue of integrated versus separated displays, a study of the effect for distributed versus stacked displays across levels of expertise is needed. But it is unclear whether the redundancy effect would also be observed in our case. Unlike integrated displays, all forms of distributed displays keep the multiple sources as separated entities; thus, even though the sources may be placed close to each other, each source can be considered as a chunk and more easily disregarded when necessary.

## References

- Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction, 16*, 389–400. doi:10.1016/j.learninstruc.2006.09.001
- Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 135–146). New York, NY: Cambridge University Press.
- Baddeley, A. (1992, January 31). Working memory. *Science, 255*, 556–559. doi:10.1126/science.1736359
- Cerpa, N., Chandler, P., & Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. *Journal of Educational Computing Research, 15*, 345–367. doi:10.2190/mg7x-4j8n-ckyr-p06t
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332. doi:10.1207/s1532690xci0804\_2
- Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62*, 233–246.
- Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology, 10*, 151–170. doi:10.1002/(SICI)1099-0720(199604)10:2<151::AID-ACP380>3.0.CO;2-U
- Cleveland, W. S., & McGill, R. (1985, August 30). Graphical perception and graphical methods for analyzing scientific data. *Science, 229*, 828–833. doi:10.1126/science.229.4716.828
- De Valois, R. L., & De Valois, K. K. (1988). *Spatial vision*. New York, NY: Oxford University Press.
- Fu, W. T., & Gray, W. D. (2000). *Memory versus perceptual motor tradeoffs in a blocks world task*. Paper presented at the 22nd Annual Conference of the Cognitive Science Society, Philadelphia, PA.
- Fu, W. T., & Gray, W. D. (2006). Suboptimal tradeoffs in information seeking. *Cognitive Psychology, 52*, 195–242. doi:10.1016/j.cogpsych.2005.08.002
- Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16*, 511–525. doi:10.1016/j.learninstruc.2006.10.001
- Hirsch, J., & Curcio, C. A. (1989). The spatial resolution capacity of

- human foveal retina. *Vision Research*, 29, 1095–1101. doi:10.1016/0042-6989(89)90058-8
- Jang, J. (2009). *Lost in space: The surprising role of information spatial layout*. Unpublished master's thesis, University of Pittsburgh, Pittsburgh, PA.
- Jang, J., & Schunn, C. (2010). *The performance benefits of spatially separated vs. stacked information on information integration tasks*. Manuscript submitted for publication.
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23–31. doi:10.1207/S15326985EP3801\_4
- Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design (theory on learning). *Human Factors*, 40, 1–17. doi:10.1518/001872098779480587
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351–371. doi:10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6
- Kong, X., Schunn, C. D., & Wallstrom, G. L. (2010). High regularities in eye-movement patterns reveal the dynamics of the visual working memory allocation mechanism. *Cognitive Science*, 34, 322–337. doi:10.1111/j.1551-6709.2009.01075.x
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–100. doi:10.1111/j.1551-6708.1987.tb00863.x
- Lovett, M., & Greenhouse, J. (2000). Applying cognitive theory to statistics instruction. *American Statistician*, 54, 196–206. doi:10.2307/2685590
- Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, 88, 49–63. doi:10.1037/0022-0663.88.1.49
- Martin-Michiellot, S., & Mendelsohn, P. (2000). Cognitive load while learning with a graphical computer interface. *Journal of Computer Assisted Learning*, 16, 284–293. doi:10.1046/j.1365-2729.2000.00141.x
- Mayer, R. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology*, 81, 240–246. doi:10.1037/0022-0663.81.2.240
- Mayer, R., & Gallini, J. (1990). When is an illustration worth a thousand words? *Journal of Educational Psychology*, 82, 715–726. doi:10.1037/0022-0663.82.4.715
- Mayer, R., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development*, 43, 31–41. doi:10.1007/BF02300480
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97. doi:10.1037/h0043158
- Olive, T., Rouet, J.-F., François, E., & Zampa, V. (2008). Summarizing digital documents: Effects of alternate or simultaneous window display. *Applied Cognitive Psychology*, 22, 541–558. doi:10.1002/acp.1380
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84, 429–434. doi:10.1037/0022-0663.84.4.429
- Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38, 63–71. doi:10.1207/S15326985EP3801\_8
- Paas, F., & van Merriënboer, J. J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, 86, 122–133. doi:10.1037/0022-0663.86.1.122
- Peebles, D., & Cheng, P. C.-H. (2003). Modeling the effect of task and graphical representation on response latency in a graph reading task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 45, 28–46. doi:10.1518/hfes.45.1.28.27225
- Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, 14, 47–69. doi:10.1023/A:1013180410169
- Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology*, 91, 690–702.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12, 185–233. doi:10.1207/s1532690xci1203\_1
- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, 119, 176–192. doi:10.1037/0096-3445.119.2.176
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251–296. doi:10.1023/A:1022193728205
- Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, 80, 424–436. doi:10.1037/0022-0663.80.4.424
- van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*, 43, 16–26. doi:10.1080/00461520701756248
- van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2008). Effects of studying sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. *Learning and Instruction*, 18, 211–222. doi:10.1016/j.learninstruc.2007.03.003
- van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17, 147–177. doi:10.1007/s10648-005-3951-0
- Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, 7, 1–39. doi:10.1207/s1532690xci0701\_1
- Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37, 473–494. doi:10.1518/001872095779049408
- Wickens, C. D., & McCarley, J. S. (2008). *Applied attention theory*. Boca Raton, FL: Taylor & Francis.
- Wiley, J. (2001). *Supporting understanding through task and browser design*. Paper presented at the 23rd Annual Conference of the Cognitive Science Society, Edinburgh, Scotland.
- Zacks, J., & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory & Cognition*, 27, 1073–1079.

(Appendices follow)

### Appendix A

#### Means and Standard Deviations for Task Time, Accuracy, Overall Learning Outcome, and Help Request Frequency in Study 1 as a Function of Condition

Dependent measure	Distributed format ( <i>N</i> = 25)		Stacked format ( <i>N</i> = 25)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Task completion time (min)	23.9	5.8	30.5	6.5
Organize data	4.7	2.1	6.0	3.4
Create pivot table	4.1	2.4	5.8	3.4
Graph data	6.0	2.6	8.6	2.3
Run a <i>t</i> test	9.1	4.2	10.2	5.5
Task accuracy (%)	94.1	14.1	93.4	12.4
Organize data	96.8	9.5	95.2	13.3
Create pivot table	94.4	16.9	95.2	16.6
Graph data	86.7	25.5	85.3	25.6
Run a <i>t</i> test	96.0	15.6	95.0	16.1
Overall learning test (%)	82.7	14.4	73.8	14.0
Factual	94.0	10.9	84.0	19.0
Conceptual	77.0	19.8	71.8	17.0
Integrative	83.2	20.6	71.6	25.7
Help request frequency	0.5	0.7	1.6	1.4

### Appendix B

#### Means and Standard Deviations for Task Time, Accuracy, Overall Learning Outcome, and Help Request Frequency in Study 2 as a Function of Condition

Dependent measure	Distributed format ( <i>N</i> = 23)		Stacked format ( <i>N</i> = 23)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Task completion time (min)	22.1	4.9	25.7	4.7
Organize data	5.3	1.8	6.0	2.5
Create pivot table	4.7	1.7	5.3	2.3
Graph data	6.1	1.6	8.0	2.7
Run a <i>t</i> test	6.0	3.4	6.3	3.0
Task accuracy (%)	90.2	10.4	88.1	12.8
Organize data	92.2	15.7	94.8	10.8
Create pivot table	93.0	9.7	92.2	13.1
Graph data	72.4	33.0	71.0	35.3
Run a <i>t</i> test	97.8	10.4	88.0	26.0
Overall learning test (%)	76.0	9.9	68.1	13.5
Factual	82.8	14.4	75.2	15.6
Conceptual	74.7	17.7	67.0	21.1
Integrative	71.0	18.7	62.7	20.0
Help request frequency	0.6	0.8	0.9	1.1

Received May 5, 2010

Revision received October 25, 2010

Accepted October 27, 2010 ■