# Cognitive Load and Learning Effects of Having Students Organize Pictures and Words in Multimedia Environments: The Role of Student Interactivity and Feedback

☐ Roxana Moreno
Alfred Valdez

*The cognitive load and learning effects of dual-code and interactivity—two multimedia methods intended to promote meaningful learning—were examined. In Experiment 1, college students learned about the causal chain of events leading to the process of lightning formation with a set of words and corresponding pictures (Group WP), pictures (Group P), or words (Group W). Some students were presented with the organized causal chain of events to study, whereas others were given a self-organization task. Consistent with a cognitive theory of multimedia learning, Condition WP was the highest in instructional efficiency for retention and transfer. However, contrary to our predictions, having students organize the multimedia materials was detrimental to transfer. Two follow-up experiments tested the hypotheses that the negative effects of interactivity were due to students' lack of time control (Experiment 2) and the form of feedback (Experiment 3). The findings showed that interactivity was effective when students were asked to evaluate their answers before receiving corrective feedback from the system.*

☐ How can instructors help students construct a meaningful model of a causal system using multimedia technologies? The goal of this study was to extend past research on multimedia learning by examining the role of dual code and interactivity in promoting scientific understanding. First, concerning dual code, we tested the hypothesis that presenting students with dual representations of a causal system (verbal and nonverbal) promotes deeper learning than presenting students with only one representation. Second, regarding interactivity, we tested the hypothesis that asking students to organize the steps in the causal chain of a scientific system promotes deeper learning than presenting students with a set of preorganized causal chain steps.

Multimedia environments have the potential of promoting meaningful learning by varying both the number of representations provided to students and the degree of student interactivity. In the present study, we varied the number of representations by comparing how students learn from one representation of a scientific system (i.e., pictures or words) versus two representations (i.e., pictures and words). In addition, we varied the interactivity degree by comparing how students learn from a multimedia program where they either need to organize the causal steps of the system to be learned or spend the same amount of time studying preorganized causal steps.

We tested the dual code and interactivity hypotheses by having college students learn about the process of lightning formation with a

ETR&D, Vol. 53, No. 3, 2005, pp. 35–45 ISSN 1042–1629 

multimedia program. To test the effectiveness of dual code, we compared the learning of three groups of students. (a) Group P was presented with a set of frames illustrating with pictures the main steps in the causal chain that lead to the process of lightning. (b) Group W was presented with a set of frames describing in words each one of the corresponding causal chain steps. (c) Group WP was presented with dual codes—a set of picture frames with their corresponding descriptions in words.

To test the effectiveness of interactivity by means of a self-organization technique, we divided the P, W, and WP groups in two. One group of students was presented with a set of frames about the process of lightning formation organized in a temporal sequence that explained the cause and effect chain from beginning to end (not-interactive or NI group). A second group of students learned with an identical set of frames, with the exception that the frames were not presented in the right sequence but had to be organized by the student to explain the cause and effect chain (interactive or I group). We focused on the following measures of learning: (a) *retention,* where we asked students to write down everything they could remember about the lightning process; (b) *transfer,* where we asked students to solve a set of four problem-solving transfer questions; (c) *mental load,* where we asked students to rate how difficult to learn the program was; and (d) the *relative efficiencies* of the different learning conditions, where mental load and performance measures were combined to yield an instructional condition efficiency score.

## Theoretical Framework

According to Mayer and Moreno's (2003) cognitive theory of multimedia learning (CTML), meaningful learning is active learning in which the learner possesses and uses a variety of cognitive processes to make sense out of the presented information. The major cognitive processes that lead to meaningful learning include selecting relevant information, organizing that information into coherent representations, and integrating these representations with existing knowledge. Based on research in cognitive science, CTML makes the following three assumptions: (a) the dual channel assumption, (b) the active processing assumption, and (c) the limited capacity assumption (Mayer & Moreno).

According to CTML, the representation and processing of information in humans is handled cognitively by two separate subsystems: one concerned with verbal materials, and the other concerned with nonverbal materials (Clark & Paivio, 1991). This assumption supports a dual-code hypothesis according to which teaching students about a causal system in both verbal (i.e., a description in printed words) and nonverbal codes (i.e., a set of graphics depicting the system), results in stronger encoding than teaching them with a verbal or nonverbal code alone. For the dual-code hypothesis to apply, the information contained in both codes should not be redundant, and the two codes should be integrated so that students are not forced to split their attention between the two codes (Kalyuga, Chandler, & Sweller, 1999; van Merriënboer & Ayres, 2005).

In addition, this research is concerned with the active processing assumption of CTML, according to which, meaningful learning requires significant conscious processing within the verbal and visual channels. This assumption supports the interactivity hypothesis. Having students organize rather than study the pre-organized materials is cognitively engaging, and thus promotes deeper learning. Similar to past research in the study strategy area, we expected learning to increase when students were asked to generate their own context for meaning by self-organizing the materials (Bruning, Schraw, & Ronning, 1999). For example, research on the generation effect (Rabinowitz & Craik, 1986), elaborative interrogation (Martin & Pressley, 1991), and the self-explanation effect (Chi, Bassok, Lewis, Reimann & Glaser, 1989) has consistently shown that learning improves when students make, rather than take, meaning. Similarly, our interactive treatment was designed to facilitate the cognitive organization process during students' meaning making (Schnotz & Rasch, 2005).

Lastly, because cognitive overload is very likely to occur when learning from a multiple representation interactive environment (Moreno & Durán, 2004), our study is concerned with the limited capacity assumption of CMTL, a central thesis of cognitive load theory (CLT), (Chandler & Sweller, 1991; Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Sweller, van Merriënboer, & Paas, 1998). According to this assumption, because each working memory channel can process only a limited amount of information at any time, a potential problem is that the processing demands required by interactive multimedia may exceed the processing capacity of the cognitive system and prevent learning (Baddeley, 1998; Mayer & Moreno, 2003). Therefore, an additional goal of this research was to examine the cognitive load effects of the instructional methods used in our studies.

### Hypotheses and Predictions

Consistent with the dual-code hypothesis, we predicted that, compared to Groups W and P, Group WP would have higher mean scores in both learning measures, and report lower levels of mental effort. The combination of relatively higher performance and lower cognitive load for Group WP should, in turn, lead to higher instructional efficiency. Consistent with the interactivity hypothesis, we predicted that groups using the interactive strategy (I) would outperform groups not using the interactive strategy (NI) on both learning measures. In addition, because a basic assumption of CLT is that asking students to consciously engage in the processing of new information results in *germane load*—load that is dependent on how much effort the learner invests to comprehend the instructional materials (Paas et al., 2003; Paas, Tuovinen, van Merriënboer, & Darabi, 2005), we expected Group I to report higher levels of mental effort than Group NI. However, predictions about differences in the relative instructional efficiency of I and NI conditions are unclear. The predicted combination of relatively higher performance and higher cognitive load for Group I and relatively lower performance and lower cognitive load for Group NI may lead to equivalent intermediate efficiency conditions. Finally, no interactions between code and interactivity were predicted: According to CTML, dual coding promotes learning from interactive or non-interactive environments, and interactivity promotes learning from words and/or pictures.

## EXPERIMENT 1

### Method

*Participants and design.* The participants were 98 undergraduate students from the Educational Psychology Subject Pool at a southwestern university in the United States (84 females and 14 males), who lacked substantial knowledge in meteorology as assessed by a pretest questionnaire. There were 17 students in the I-P group, 16 in the I-W group, 15 in the I-WP group, 16 in the NI-P group, 15 in the NI-W group, and 19 in the NI-WP group. The mean age of the participants was 27.31 ($SD$ = 8.49). Neither age nor gender differed significantly among the groups.

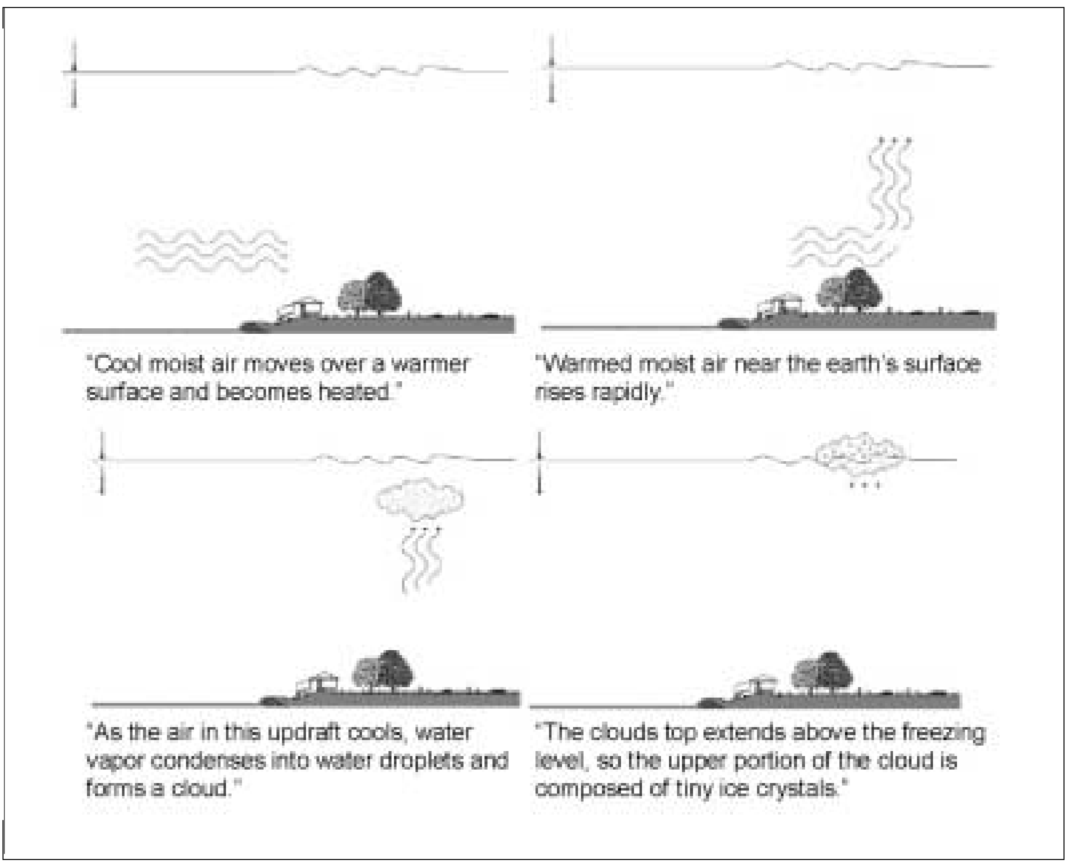
*Materials and apparatus.* For each participant, the paper-and-pencil materials consisted of a subject questionnaire, a difficulty rating, a retention test, and a four-page transfer test, each typed on 8.5 × 11″ sheets of paper. The subject questionnaire solicited demographic and prior knowledge (meteorology) information. Meteorology knowledge was assessed with the six-item knowledge checklist and five-item self-rating used by Moreno and Mayer (2000a).

The difficulty rating sheet contained the following question followed by a 7-point rating scale (1 as *very easy* and 7 as *very hard*) and was intended to assess the learner's perception of learning difficulty: "How difficult was it for you to learn about the process of lightning?" The retention test contained the following instructions at the top of the sheet: "Please write down an explanation of how lightning works." The transfer test consisted of the following four questions, each typed on separate sheets: (a) "What can be done to decrease the intensity of lightning?" (b) "Suppose you see clouds in the sky, but no lightning. Why not?" (c) "What does air temperature have to do with lightning?" (d) "What do electrical charges have to do with lightning?"

The computerized materials were six computer programs for multimedia presentations on how the lightning process works (I-W, I-P, I-WP, NI-W, NI-P, and NI-WP versions), with each consisting of a sequence of 16 frames depicting the causal chain of events leading to the process of lightning formation (i.e., condensation, formation of crystals, polarization, etc.). The 16 frames represented these events either in words (W), pictures (P), or words and pictures (WP). After a short training session about dragging and placing frames in the right order, the three interactive groups (I) were presented with a random sequence of 4 frames at a time and given 3 min to organize them in the right sequential order. Corrective feedback was given for each frame movement by displaying the word *Correct!* on top of the frame when it was successfully dragged to its place, or by returning the frame to the original position when wrongly placed. After the 3 min was over, the 4 frames were shown in the correct order for an additional 1 min. Then, the next set of 4 randomly ordered frames was presented, and the same process was repeated. Groups in the noninteractive condition (NI) were presented with an identical program, except that each set of 4 frames was presented in the right sequential order for 4 min. Figure 1 shows a selected set of frames from the NI-WP version of the program. On the computer screen, the P and W versions looked identical to the WP versions, except that the bottom text or the top picture was deleted, respectively. The multimedia presentations were developed using Flash MX™ (Macromedia, 2002). The apparatus consisted of 6 Pentium III PC computer systems, each with a 15-in monitor.

*Procedure.* One to six participants were tested per session. Each participant was randomly assigned to one of the six treatment groups, and

Figure 1 ☐ Selected frames from the NI-WP version of the program.



seated in front of an individual computer. After completing the subject questionnaire, the experimenter indicated that the computer would show a presentation of how the process of lightning works and that, when the program ended, participants would be questioned to assess how much they had learned. Next, the respective version of the program was presented once to all participants. Once the presentation was finished, students were given the difficulty-rating sheet to complete at their own pace, 5 min to answer the retention sheet, and 3 min to answer each of the four problem-solving sheets of the transfer test.

*Scoring*. Two scorers determined the experience, difficulty, retention, and transfer scores for each participant. Agreement between both scorers was 92% on the retention test, 87% on the transfer tests, and 100% on the experience and difficulty scores. Differences between scorers were solved by agreement. The experience score was computed using the procedure described by Moreno and Mayer (2000a). This questionnaire was designed to exclude those students who reported significant prior experience in the subject domain. Data for students who scored above the median were eliminated ($n$ = 9). The difficulty score was the number that participants had circled in the difficulty-rating sheet. Retention scores were computed by counting the number of major idea units (out of 19 possible) that the participant produced on the retention test. Transfer scores were computed by counting the number of acceptable answers that the participant produced across the four transfer problems. Finally, we calculated the instructional efficiency of each treatment. Using Paas and van Merriënboer's (1993) computational method, we combined participants' mean ratings of difficulty with their mean learning scores to compare the efficiency of each treatment for retention and transfer.

## Results and Discussion

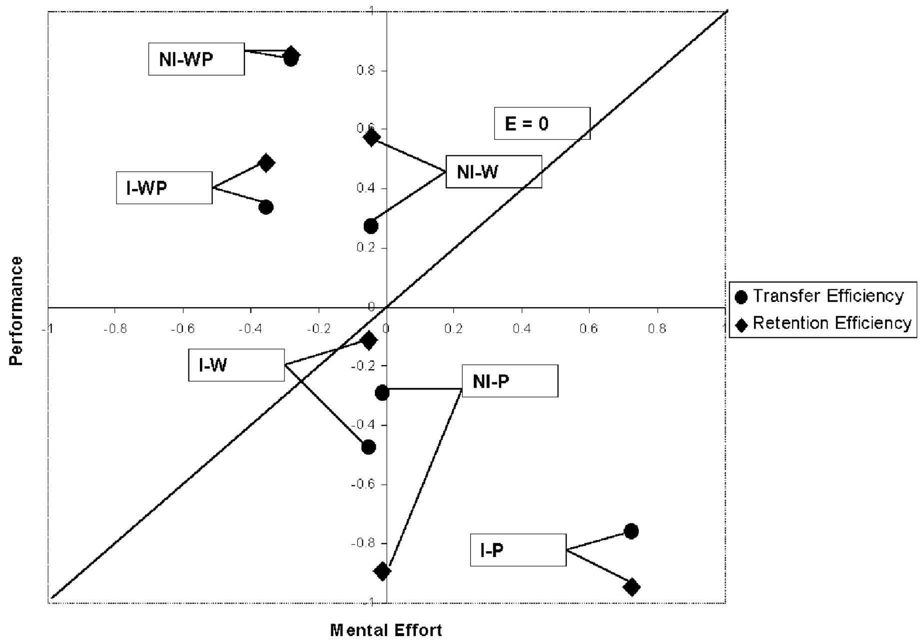
We conducted a multivariate analysis of variance (MANOVA) with code (P, W, or WP) and interactivity (I vs. NI) as between-subjects factors, and retention, transfer, and difficulty as dependent measures, to determine if the treatment groups differed on the dependent measures. Data were screened for extreme or missing values, and statistical assumptions for MANOVAs were considered. Alpha was set at .05 when evaluating tests of statistical significance, and Type I error rates due to multiple two-way analyses of variances (ANOVA) on each dependent variable (conducted as follow-up tests to the MANOVA) were controlled using a Bonferroni adjustment of alpha (0.017). All pairwise comparisons were conducted using Tukey's HSD. Table 1 shows the mean scores

Table 1 ☐ Mean scores (and standard deviations) for six groups on retention, transfer, and difficulty ratings—Experiment 1.

| | | *Retention* | | *Transfer* | | *Difficulty* | |
|---|---|---|---|---|---|---|---|
| *Group* | *N* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| NI-P | 16 | 2.50 | 2.07 | 2.12 | 1.63 | 3.25 | 1.44 |
| NI-W | 15 | 7.80 | 3.32 | 3.40 | 1.72 | 2.73 | 1.49 |
| NI-WP | 19 | 9.21 | 3.69 | 4.42 | 2.19 | 2.84 | 1.50 |
| I-P | 17 | 2.29 | 1.79 | 1.18 | 1.18 | 4.35 | 1.45 |
| I-W | 16 | 5.50 | 2.19 | 1.75 | 1.00 | 3.19 | 1.11 |
| I-WP | 15 | 8.13 | 2.67 | 3.27 | 2.19 | 3.20 | 1.61 |

*Note.* Actual score ranges were 0 to 14 for retention, 0 to 7 for transfer, and 1 to 6 for difficulty.
NI-P = noninteractive–pictures; NI-W = noninteractive–words; NI-WP = noninteractive–words and pictures; I-P = interactive–pictures; I-W = interactive–words; I-WP = interactive–words and pictures.

Figure 2 ☐ Graphic representation to visualize the instructional efficiency of six treatment groups—Experiment 1.




and standard deviations for the six groups on measures of retention, transfer, and difficulty ratings.

The MANOVA revealed significant differences among the different code groups, Wilks's $\Lambda = .43$, $F(4, 89) = 11.54$, $p < .0001$, and between the different interactivity groups Wilks's $\Lambda = .84$, $F(4, 89) = 4.16$, $p = .004$, on the dependent measures. There were no significant interactions between code and interactivity. Figure 2 shows the standardized mental effort and performance scores on retention and transfer for each of the six treatment groups. The resulting distance of the efficiency points from the E = 0 line gives a measure and direction of the instructional efficiency.

*Issue 1: Do students construct better models of causal systems when they are presented with two representation modes rather than one*? Using retention as a dependent measure, there was a large and significant main effect for code, $F(2, 92) = 45.52$, $MSE = 339.80$, $p < .0001$, $\eta^2 = .50$. Mean retention scores showed that Group WP had higher scores than Group W, which had higher scores than Group P (*M*s = 8.67, 6.65, and 2.40; *SD*s = 2.74, 2.73, and 2.76 for the WP, W, and P groups, respectively). There was also a moderate and significant main effect for code on transfer, $F(2, 92) = 13.62$, $MSE = 40.33$, $p < .0001$, $\eta^2 = .23$. Group mean transfer scores showed a similar pattern as was seen for retention mean scores, with Group WP mean transfer score higher than the other groups and Group W higher than Group P on tests of transfer to problem solving (*M*s = 3.84, 2.58, and 1.65; *SD*s = 1.73, 1.72, and 1.72 for the WP, W, and P groups, respectively). In addition, there was a marginal but nonsignificant effect for code on difficulty, $F(2, 92) = 3.45$, $MSE = 7.18$, $p = .04$, $\eta^2 = .07$. Students' difficulty ratings were higher for P conditions than for W and WP conditions (*M*s = 3.80, 2.96, and 3.02; *SD*s = 1.44, 1.44, and 1.45, respectively). Finally, Group WP had the highest instructional efficiency for retention and transfer (*E* = .62 and .51, respectively), followed by Group W (*E* = .28 and .09, respectively). Group P showed relatively higher cognitive load and lower performance on retention and transfer ($E$ = –.90 and –.62, respectively), a case of low-instructional efficiency.

*Issue 2: Do students construct better models of causal systems when they are asked to organize the causal chain themselves rather than when the materials are preorganized*? Using retention as a dependent measure, there was a marginal effect for interactivity which failed the Bonferroni adjustment, $F(1, 92) = 4.65$, $MSE = 34.72$, $p = .03$, $\eta^2 = .05$, indicating that, contrary to our predictions, Group NI scored marginally higher than Group I on this measure (*M*s = 6.50 and 5.31; *SD*s = 2.74 and 2.74, respectively). There was a small and significant main effect for interactivity on transfer in the same direction, $F(1, 92) = 12.86$, $MSE = 38.09$, $p = .001$, $\eta^2 = .12$, (*M*s = 3.31 and 2.06; *SD*s = 1.73 and 1.72, respectively). Additionally, there was a marginal and nonsignificant effect for interactivity on difficulty, $F(1, 92) = 4.77$, $MSE = 9.92$, $p = .03$, $\eta^2 = .05$, indicating that Group I had marginally higher difficulty ratings than Group NI (*M*s = 3.58 and 2.94; *SD*s = 1.44 and 1.45, respectively). Finally, whereas Condition NI showed high instructional efficiency for retention and transfer ($E$ = .26 and .36), Condition I showed low instructional efficiency ($E$ = –.26 and –.37).

In summary, the findings from Experiment 1 show that, consistent with our predictions, students learn best when the instructional materials present two representation codes rather than one, as a result of the relatively higher performance obtained with relatively lower cognitive load. On the other hand, contrary to our predictions, the self-organization technique hurt students' learning from the multimedia program. Interactive conditions had low instructional efficiency for both learning measures as a combination of the relatively lower performance and higher cognitive load.

A possible explanation for the negative interactivity effects may be that the time to organize the materials was arbitrarily imposed by the computer program. Had we given students unconstrained time to organize the materials, the results might have been different. A second explanation resides on the form of feedback. Because Group I could see the word *Correct!* immediately after placing each frame into its correct position, we hypothesized that the feedback design may have encouraged students to merely fit the frames in place by trial-and-error rather than promote the mindful processing of the entire chain of events. The goal of the next two experiments was to test these two hypotheses.

## EXPERIMENT 2

The purpose of Experiment 2 was to determine if the negative interactivity effects were due to participants' lack of control of the time deemed necessary to organize the multimedia materials. Most researchers in the area of computer-based instruction agree that effective interactivity requires the learner to be in control of the pace of the presentation (Cairncross & Mannion, 2001; Stanton, Porter, & Stroud, 2001), and past research has shown that, on measures of transfer, learners who are allowed to exercise control over the pace of a multimedia message outperform those who are not (Mayer & Chandler, 2001). In Experiment 2, we compared the cognitive load and learning effects of an interactive, system-timed condition (I-ST), an interactive, user-timed condition (I-UT), and a not-interactive, system-timed condition (NI-ST). We predicted Condition I-UT to be the most efficient as a result of relatively higher performance and lower mental effort and, similar to Experiment 1, Condition I-ST to be the least effective.

### Method

*Participants and design.* The participants were 53 undergraduate students from the Educational Psychology Subject Pool at a southwestern university in the United States (45 females and 8 males) who lacked substantial knowledge in meteorology. There were 17 students in the I-UT group and 18 in the I-ST and NI-ST groups. The mean age of the participants was 26.43 (*SD* = 8.72). Neither age nor gender differed significantly among the groups.

*Materials and apparatus.* The paper and pencil materials were identical to those used in Experiment 1. The I-ST version of the program was identical to the I-WP version used in Experiment 1. The I-UT and I-ST versions were identical except that Group I-UT had no time limitations on interacting with the frames. The program moved to the next screen only after each set of events was successfully organized.

*Procedure.* The procedure was identical to the one used in Experiment 1. Data for students who scored above the median on the experience questionnaire were eliminated ($n$ = 4).

### Results and Discussion

The top three rows in Table 2 show the mean scores and standard deviations for the three groups on measures of retention, transfer, and difficulty ratings. A MANOVA, with treatment group as between-subjects factor, and retention, transfer, and difficulty as dependent measures, failed to reveal significant differences among the different groups on the dependent measures, Wilks's $\Lambda = .43$, $F(6, 96) = 0.65$, $p = .69$. Consequently, the treatments yielded comparable instructional efficiency values for retention and transfer. In summary, we did not find support for the hypothesis that the negative interactivity effects found in Experiment 1 were due to participants' lack of control of the time necessary to organize the multimedia materials. In fact, contrary to our expectations, Group I-UT spent less rather than more time on the multimedia program than the time given to the other groups by the system, $F(2, 50) = 8.90$, $MSE = 211{,}535.28$, $p < .001$.

## EXPERIMENT 3

The purpose of Experiment 3 was to test the hypothesis that the type of feedback given to Group I in Experiment 1 hurt learning by promoting the superficial processing of the multimedia materials. Different feedback forms may prompt the student to process the materials mindfully or not mindfully (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). In a meta-analysis of the effects of feedback in computer-based instruction, Azevedo and Bernard (1995) concluded that feedback messages, to be effective, should stimulate the cognitive processes necessary to gain deep understanding. Because our interaction version included frame-by-frame immediate corrective feedback, we hypothesized that students may have merely resorted to select by trial-and-error the correct frame positions rather than engage in deep processing (Morrison, Ross, Gopalakrishnan, & Casey, 1995). Had we asked students to assess the correctness of the sequence as a whole before submitting their answer, we might have found

Table 2 ☐ Mean scores (and standard deviations) for four groups on retention, transfer, and difficulty ratings—Experiments 2 and 3.

| | | *Retention* | | *Transfer* | | *Difficulty* | |
|---|---|---|---|---|---|---|---|
| *Group* | *N* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| *Experiment 2* | | | | | | | |
| I-ST | 18 | 8.06 | 3.17 | 3.28 | 1.64 | 3.00 | 1.37 |
| I-UT | 17 | 8.59 | 3.17 | 3.53 | 1.54 | 2.76 | 1.25 |
| NI-ST | 18 | 7.61 | 3.17 | 4.11 | 1.64 | 2.72 | 1.27 |
| *Experiment 3* | | | | | | | |
| I-SF | 15 | 7.80 | 2.80 | 3.13 | 1.50 | 3.37 | 1.26 |
| I-UF | 16 | 8.75 | 2.77 | 4.19 | 1.51 | 3.53 | 1.50 |

*Note.* Actual score ranges were 0 to 15 for retention, 0 to 7 for transfer, and 1 to 6 for difficulty.
I-ST = interactive group–system-timed condition; I-UT = interactive group–user-timed condition;
NI-ST = not-interactive group–system-timed condition; I-SF = interactive group–system-feedback condition;
I-UF = interactive group–user-feedback condition.

positive interactivity effects. In Experiment 3, we compared the cognitive load and learning outcomes of an interactive, system-feedback (I-SF) condition that presented students with frame by frame corrective feedback, and an interactive, user-feedback (I-UF) condition that required student evaluation of the correctness of the organized set of frames before submitting the sequence for system feedback.

### Method

*Participants and design.* The participants were 31 undergraduate students from the Educational Psychology Subject Pool at a southwestern university in the United States (27 females and 4 males) who lacked substantial knowledge in meteorology. There were 15 students in the I-SF group, and 16 in the I-UF group. The mean age of the participants was 22.68 ($SD = 5.33$). Neither age nor gender differed significantly among the groups.

*Materials and apparatus.* The paper and pencil materials were identical to those used in Experiment 1. The I-SF version of the program was identical to the I-UT version used in Experiment 2. The I-SF and I-UF versions were identical except that the latter version instructed students to arrange the four frames in order and to press the SUBMIT button (placed at the bottom of the screen) when they considered that the four frames represented the causal chain sequence in the right order.

*Procedure.* The procedure was identical to the one used in Experiment 2. Data for students who scored above the median on the experience questionnaire were eliminated ($n = 2$).

### Results and Discussion

The bottom two rows of Table 2 show the mean scores and standard deviations for the two groups on measures of retention, transfer, and difficulty ratings. A MANOVA, with treatment group as between-subjects factor, and retention, transfer, and difficulty as dependent measures, revealed significant differences among the different groups on the dependent measures, Wilks's $\Lambda = .76$, $F(3, 27) = 2.88$, $p = .05$. ANOVAs on each dependent variable showed that, although groups did not differ in measures of retention or difficulty, consistent with our predictions, there were large and significant group differences on mean transfer scores $F(1, 29) = 6.54$, $MSE = 8.60$, $p < .0167$, $\eta^2 = .71$. The mean transfer score of Group I-UF was higher than that of Group I-SF, leading to higher instructional efficiency ($E = .25$ and $-.26$ for I-UF and I-SF, respectively). In summary, the hypothesis that effective interactivity requires feedback designs that promote the intentional and purposeful processing of the information was supported.

## GENERAL DISCUSSION

The goal of this study was to extend past research on multimedia learning by examining the role of dual coding and interactivity in promoting scientific understanding. The findings of Experiment 1 strongly supported the dual-code hypothesis, which predicts that students learn better when provided with visual and verbal knowledge representations rather than visual or verbal representations alone. This was demonstrated by finding a dual-code effect on learning, cognitive load, and instructional efficiency. That is, learning about a causal system with non-redundant, integrated words and pictures is significantly more efficient than learning with text or pictures alone.

On the other hand, although the interactivity hypothesis derived from CTML predicts that students learn better when given the opportunity to make rather than take meaning, our self-organization technique proved to hurt learning. The results from Experiments 2 and 3 suggest that, even when students are given a cognitively engaging task—the interactive treatment—and control over the pace of their interaction—the I-UT treatment—deep learning is not promoted unless careful consideration is given to the effects of different feedback strategies. This is particularly important for the case of novice learners such as the ones used in the present study, who may lack the necessary schemas to guide them in the process of meaning making (Clarke, Ayres, & Sweller, 2005; Kalyuga & Sweller, 2005). Because interactivity in e-learning is a two-way communication between a learner and the computer system, it is important to focus on both components as potential sources to promote students' cognitive engagement. In our study, asking students to evaluate a set of frames before submitting an answer was all that was needed to promote deeper learning from a self-organization technique.

On the practical side, the present study contributes to the growing research base on the design of e-learning applications (Bruning, Horn, & PytlikZillig, 2003) by pointing out two effective methods that instructional designers may use in interactive learning environments. First, designers of e-learning environments should combine and coordinate verbal explanations with visual presentations of science topics to reduce the cognitive load of students and increase learning. Second, they should take advantage of the transformation capabilities of computers to aid learners in the elaboration of their mental models (Kozma, 1991). Apparently, the beneficial effects of interactivity in a multimedia science lesson are less dependent on the behavioral interaction required during a computer lesson than on the mental interaction needed to actively involve the learner in the process of understanding (Moreno & Mayer, 2000b). The present study reinforces the importance for instructional designers not just to present feedback, but to include techniques that invite learners to evaluate their own actions before submitting their answers for feedback.

Finally, our research is limited because it dealt with only one kind of multimedia lesson (i.e., a scientific multimedia explanation), one kind of interactivity (i.e., having students organize a causal chain of events), two feedback forms (i.e., frame-by-frame vs. sequence-corrective feedback), and one kind of learner (i.e., college students who were unfamiliar with meteorology). While our findings begin to provide a better understanding of the cognitive processes required for effective e-learning, further research is needed to provide additional information concerning the roles of interactivity and feedback. □

Roxana Moreno [moreno@unm.edu] and Alfred Valdez [f.valdez1@comcast.net] are with the Educational Psychology Program at the University of New Mexico.

Correspondence concerning this article should be addressed to either author c/o Educational Psychology Program, Simpson Hall 123, University of New Mexico, Albuquerque, NM 87131.

This research was supported by the CAREER 0238385 grant from the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Casey Frechette and Paul Gibson created the multimedia materials used in these studies.

## REFERENCES

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research, 13*(2), 111–127.

Baddeley, A. (1998). *Human memory*. Boston, MA: Allyn & Bacon.

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213–238.

Bruning, R., Horn, C. A., & PytlikZillig, L. M. (2003). *Web-based learning: What do we know? Where do we go?* Greenwich, CT: Information Age Publishing.

Bruning, R. H., Schraw, G. J., & Ronning, R. (1999). *Cognitive psychology and instruction*. Upper Saddle River, NJ: Prentice Hall.

Cairncross, S., & Mannion, M. (2001). Interactive multimedia and learning: Realizing the benefits. *Innovation in Education and Teaching International*, *38*(2), 156–164.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145–182.

Clark, J.M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*, 149–210.

Clarke, T., Ayres, P., & Sweller, J. (2005). The impact of sequencing and prior knowledge on learning mathematics through spreadsheet applications. [This special issue]. *Educational Technology Research and Development, 53*(3), 15–24.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351–371.

Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. [This special issue]. *Educational Technology Research and Development, 53*(3), 83–93.

Kozma, R. B. (1991). Learning with media. *Review of Educational Research, 61*, 179–211.

Macromedia. (2002). *Flash MX™* [Computer program]. San Francisco, CA: Author.

Martin, V. L., & Pressley, M. (1991). Elaborative-interrogation effects depend on the nature of the question. *Journal of Educational Psychology, 83*, 113–119.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*, 390–397.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*, 43–52.

Moreno, R., & Durán, R. (2004). Do multiple representations need explanations? The role of verbal guidance and individual differences in multimedia mathematics learning. *Journal of Educational Psychology, 96,* 492–503.

Moreno, R., & Mayer, R. E. (2000a). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology, 92*, 117–125.

Moreno, R., & Mayer, R. E. (2000b). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology, 92, 724–733.*

Morrison, G. R., Ross, S. M., Gopalakrishnan, M., & Casey, J. (1995). The effects of feedback and incentives on achievement in computer-based instruction. *Contemporary Educational Psychology, 20*, 32–50.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive theory. *Educational Psychologist, 38,* 63–71.

Paas, F., Tuovinen, J. E., van Merriënboer, J. J. G., & Darabi, A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learners' involvement in instructional conditions. [This special issue]. *Educational Technology Research and Development, 53*(3),25–33.

Paas, F., & van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors, 35*, 737–743.

Rabinowitz, J. C., & Craik, F. I. M. (1986). Specific enhancement effects associated with word generation. *Journal of Memory and Language, 25*, 226–237.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. [This special issue]. *Educational Technology Research and Development, 53*(3), 47–58.

Stanton, N. A., Porter, L.J., & Stroud, R. (2001). Bored with point and click? Theoretical perspectives on designing learning environments. *Innovation in Education and Teaching International*, *38*(2), 175–182.

Sweller, J., van Merriënboer, J.J.G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

van Merriënboer, J. J. G., & Ayres, P. (2005). Research on cognitive load theory and its design implications for e-learning. [This special issue]. *Educational Technology Research and Development, 53*(3), 5–13.