EFFECTS OF MULTIMODAL INFORMATION ON LEARNING PERFORMANCE AND JUDGMENT OF LEARNING

GONGXIANG CHEN
XIAOLAN FU
Institute of Psychology, Chinese Academy of Sciences

ABSTRACT

Two experiments were conducted to investigate the effects of multimodal information on learning performance and judgment of learning (JOL). Experiment 1 examined the effects of representation type (word-only versus word-plus-picture) and presentation channel (visual-only versus visual-plus-auditory) on recall and immediate-JOL in fixed-rate learning conditions. Experiment 2 examined the effects of representation type (word-only versus word-plus-picture) and presentation media (computer versus paper) on recall and delayed-JOL in self-paced learning conditions. The results showed that recall performance was better in word-only conditions than in word-plus-picture conditions in Experiment 1, and better in computer conditions than in paper conditions in Experiment 2. Multimodal information had no influence on magnitude of people's judgment. Participants were overconfident in all conditions, but more overconfident in computer conditions than in paper conditions.

INTRODUCTION

There has been a great deal of interest in learning performance and judgment of learning (JOL) (e.g., Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Dunlosky &

349

Matvey, 2001; Dunlosky & Nelson, 1992, 1994; Koriat, 1997). JOL is a prediction of the likelihood of eventual learning performance for recently studied items (Dunlosky & Nelson, 1992). This article is concerned with the effects of multimodal information on learning performance and JOL.

Effects of Multimodal Information on Learning Performance

Multimodal information refers to the information presented in multiple modes such as visual and auditory modes. In general, three different theories could be regarded as the base of analyzing the effects of multimodal information on learning performance. One is Paivio's dual-coding theory (Claik & Paivio, 1991; Paivio, 1991), which stated that pictures were more likely to be coded both verbally and pictorially than words alone and thus led to better learning than words (e.g., Gerlič & Jaušovec, 2001; Larkin & Simon, 1987; Stenning & Oberlander, 1995). However, some researchers (e.g., Ainsworth, Bibby, & Wood, 1998; Reimann, 1999; Rogers, 1999) argued that graphical representations could take away learning opportunities and impair learning performance.

Dual-coding theory also supported the additive effect of multi-channel presentations. Paivio (1990) assumed that the combination of redundant channels of information would make concepts dual coded and thus result in increased learning and retention. Although many researchers (e.g., Ednerson & Tumey, 1984; Hsia, 1969; Mayer & Anderson, 1991, 1992; Mousavi, Low, & Sweller, 1995) found that learning was improved when the information was referentially processed through two channels, some researchers (e.g., Edwardson, Groms, & Pringle, 1976; Gunter, 1987) proposed that audio/video redundancy was harmful to memory.

The second theory is the cognitive load theory (Chandler & Sweller, 1991, 1992; Sweller & Chandler, 1994). Some researchers (e.g., Hegarty & Steinhoff, 1997; Zhang, 1997) found that graphical displays could reduce cognitive load required to comprehend the information presented because information was depicted in the environment. However, learning could be impaired if the learning material caused a cognitive overload. Since the capacity of working memory was very limited, presenting different sources of information in the same modality (e.g., only visually) may easily result in a split-attention effect, which led to poor learning performance (Chandler & Sweller, 1992).

Moreover, many researchers explored the differences between computer presentation and paper presentation, but arrived at two opposite conclusions. Some researchers (e.g., Askwall, 1985; Gerlič & Jaušovec, 2001) found the differences between computer presentation and paper presentation, but the others (e.g., Mason, Patry, & Bernstein, 2001; Powers, 2001; Rice, 1994) did not find any difference between them.

Baddeley (2000) proposed a new version of the multi-component working memory model. This theory suggested a central executive and two supporting systems, a "visual-spatial sketchpad" for dealing with visual images and a "phonological loop" for processing verbal information, except a new component served as episodic buffer. The episodic buffer was assumed to be a limited-capacity temporary storage that was capable of integrating information from a variety of sources held in multimodal codes. It was capable of binding information from the subsidiary systems (the phonological loop and the visuospatial sketchpad) and from long-term memory. Therefore, multiple sources of information are integrated in the episodic buffer.

Magnitude and Accuracy of JOL

JOL can be made immediately or later after items have been studied. When one student makes a JOL for an item immediately after studying it, the JOL is called an immediate JOL. In contrast, while one student studies an item and then waits for a short time (e.g., 30 seconds) before making a JOL for the item, the JOL is called a delayed JOL. In studying new materials, learners need to monitor what they have learned and decide whether to go over the materials again based on their immediate and/or delayed JOL.

Koriat (1997) advanced the cue-utilization model to explain how people make JOL. His model is based on the distinction between three general classes of cues for JOL: intrinsic, extrinsic, and mnemonic cues. Intrinsic cues refer to inherent characteristics of the study items that disclosed their a priori difficulty. Extrinsic cues pertain to the conditions of learning or to the encoding operations applied by the learner. Koriat (1997) considered that extrinsic factors would be discounted in judgment of future recall, and, in fact, sometimes failed to have any effect on JOL at all. As extrinsic cues, information presentation characteristics, such as representation type (e.g., word, picture), presentation channel (e.g., visual, auditory), and presentation media (e.g., computer, paper), might have no influence on the magnitude of people's judgments. In other words, JOL may not monitor any effects of multimodal information on learning performance.

Knowledge of our own memory system (called metamemory) stems from the self-monitoring of one's own memory (Leonesio & Nelson, 1990). One common metamemory measure is JOL (Kelemen & Weaver III, 1997). The accuracy of JOL is assessed by comparing the magnitude of JOL with future recall performance (Dunlosky & Nelson, 1992). Relative accuracy is how well a person predicts the likelihood of correct recall for one item relative to another item. In other words, relative accuracy measures whether an item given a higher JOL has a greater probability of correct recall than did an item given a lower JOL. By contrast, absolute accuracy is how well a person predicts the likelihood of correct recall for a particular set of items. In other words, absolute accuracy measures the actual percentage of correct recall for items given a particular JOL, and most

important, how close the predicted percentage is to the actual percentage of correct recall. Absolute accuracy can tell us whether a participant is overconfident or underconfident. The most common finding is that people are usually over-confident in the accuracy of their performance (Lichtenstein, Fischhoff, & Phillips, 1977). The previous research (Dunlosky & Nelson, 1992; Nelson & Dunlosky, 1991) found that both relative accuracy and absolute accuracy of delayed-JOL were higher than that of immediate-JOL.

Purposes and Hypotheses of Research

In sum, although multimodal information has been widely investigated for several years, there have been no consistent conclusions about the effects of multimodal information on learning performance. Furthermore, little was known about the effects of multimodal information on JOL. In the present study, two experiments were conducted to investigate the effects of multimodal information on recall performance and the magnitude and accuracy of JOL in different learning conditions. Experiment 1 aimed to evaluate the effects of representation type (word-only versus word-plus-picture) and presentation channel (visual-only pre-sentation versus visual-plus-auditory presentation) on immediate-JOL and its accuracy in fixed-rate learning conditions. Experiment 2 aimed to examine the effects of representation type (word only versus word-plus-picture) and presenta-tion media (computer versus paper) on delayed-JOL and its accuracy in self-paced learning conditions.

Based on the previous theories and research, we predicted that recall per-formance would be improved when the information was referentially processed through two channels rather than one channel; however, multiple sources of information presented simultaneously would impair recall performance. We also predicted that recall performance would be better in computer conditions than in paper conditions, and better in self-paced learning than in fixed-rate learning. Finally, based on Koriat's model, we predicted that information presentation characteristics would have no influence on the magnitude of people's judgments; however, little is known about whether multimodal information would influence the accuracy of JOL.

EXPERIMENT 1

Method

Participants

Seventy-five undergraduate students (37 women and 38 men) from two univer-sities in Beijing, China participated in the study. They ranged in age from 18 to 23 years (M = 20.31 years). All reported normal or corrected-to-normal vision, without deficiency in hearing.

Design

A 2 × 2 (representation type × presentation channel) between-participants design was used. There were four experimental conditions: (a) word-only in visual mode (W), (b) word-plus-picture in visual mode (WP), (c) word-only in visual-plus-auditory mode (WA), and (d) word-plus-picture in visual-plus-auditory mode (WPA).

Materials

The schematic pictures were selected from "A standardized set of 260 pictures" (Snodgrass & Vanderwart, 1980). Shu, Cheng, and Zhang (1989) measured the pictures and found 235 pictures fitting for Chinese participants. We selected 108 pictures with higher name agreement (0–1.93), higher familiarity (2.03–5), and the double-character names of these pictures. The pictures' corresponding double-character names composed 54 cue-target pairs (the first word is cue, and the second word is target). The frequencies of these words were between 0.000176 and 0.000936 (Institute of Language Teaching and Research, 1986). Thirty participants were instructed to judge the familiarity of 108 words and other 30 participants were instructed to judge the strength of the preexisting cue-to-target. Based on the results of the familiarity questionnaire and the semantic association questionnaire, 40 unrelated word pairs with their pictures were selected. The association strength between the cues and targets varied from 0 to 2.26 ($M = 1.82, SD = 0.24$). The familiarity of the 80 words varied from 3.60 to 5 ($M = 4.51, SD = 0.40$). The name agreement of the pictures varied from 0 to 1.87 ($M = 0.85, SD = 0.50$). The familiarity of 80 schematic pictures varied from 1.87 to 5 ($M = 3.90, SD = 0.75$).

Forty pairs of words were digitized into a male voice, about 1000 ms each word with a normal speech pitch, to be used in both WA and WPA conditions.

Procedure

The learning task was programmed in E-prime and run on computers. Participants were randomly assigned to one of the four experimental conditions. They were instructed that they would be studying 40 cue-target pairs for an upcoming test of paired-associate recall, and should then make a judgment about the possibility to recall the target as accurately and quickly as they could. Participants completed the tasks in the order described below.

Study— Participants were presented with a list of 40 cue-target pairs to study for a coming paired-associate recall test. Each pair (cue-target pairs in W and WA conditions; cue-target pairs with corresponding pictures pair in WP and WPA conditions) was displayed individually for 6s to study. Pairs were presented once binaurally through headphones plugged into the computer with the visual presentations in WA and WPA conditions. After the offset of the presentation of a pair, a JOL was made.

Judgments of learning—The cue word of a pair was presented in all conditions. The participant was asked to judge the likelihood of correctly retrieving the target of the pair on a subsequent test of memory along with the query, "How confident are you that in about 10 minutes you will be able to remember the second word of the pair above when prompted with: cue- ( )." Participants were instructed to make the judgments on a scale from 0 (definitely not recall) to 100 (definitely recall). Any whole number from 0 to 100 could be used.

Paired-associate recall—Following the study-JOL phase, the cue was presented in an order of studying, and the participant was instructed to respond with the word that was originally paired with that cue. There was no time limit for participants to respond to each cue.

Measures

For each participant, the proportion of correct recall was computed in each condition. Both relative and absolute JOL accuracy were evaluated in the present study.

Relative accuracy of immediate-JOL in predicting eventual recall was operationalized via a Goodman-Kruskal gamma correlation between JOL and recall for each participant. Gamma had a possible range from -1.0 to $+1.0$, with 0 representing nil predictive accuracy and with $+1.0$ representing perfect predictive accuracy (Nelson, 1984).

Absolute accuracy of immediate-JOL was assessed with bias scores. The bias index reflects participants' overall overconfidence or underconfidence in a particular condition (Yates, 1990). Bias scores were derived by obtaining the signed difference between mean JOL magnitude and mean recall performance in each condition for each participant. A score greater than 0 indicated overconfidence, and a score less than 0 indicated underconfidence.

Results and Discussion

Recall performance

Percentage of correct recall was computed for each participant. Means across participants' values are reported in Table 1. A 2×2 ANOVA revealed a main effect of representation type on recall performance, $F(1, 74) = 4.58, p < .05$. The main effect of presentation channel was not significant, $F(1, 74) = 0, p > .05$, and the interaction between representation type and presentation channel was not significant either, $F(1, 74) = 0.92, p > .05$.

Magnitude of Immediate-JOL

To evaluate the effects of representation type and presentation channel on the magnitude of immediate-JOL, we computed the median across each participant's

Table 1. Mean Proportion of Correct Recall and Magnitude of
Immediate-JOL in Experiment 1

Presentation channel	Representation type									
	Word only					Word-plus-picture				
		Recall		JOL			Recall		JOL	
	N	M	SD	M	SD	N	M	SD	M	SD
Visual-only	18	42.67	24.54	53.97	22.73	20	36.53	26.63	57.75	21.73
Visual-plus-auditory	18	48.00	24.12	66.90	20.65	18	30.50	17.51	54.92	20.12

JOL. A 2 × 2 ANOVA revealed that representation type had no influence on immediate-JOL, $F(1, 74) = 0.67$, $p > .05$; the main effect of presentation channel was not significant, $F(1, 74) = 0.85$, $p > .05$; and their interaction was not significant either, $F(1, 74) = 2.90$, $p > .05$.

Accuracy of Immediate-JOL

JOL-recall gamma correlations calculated over the entire list of 40 items averaged 0.34, significantly different from 0, $t(74) = 9.00$, $p < .001$. It suggested that the participants did not make a judgment based on a guess only. The bias score (19.29) was greater than 0, $t(74) = 6.77$, $p < .001$, however, the main effect of representation type was not significant, $F(1, 74) = 1.48$, $p > .05$; the main effect of presentation channel was not significant either, $F(1, 74) = 0.44$, $p > .05$. This finding was consistent with Lichtenstein et al. (1977) and indicated that participants were reliably overconfident in each condition.

The main effect of representation type on recall performance was significant, which meant that pictures impaired learning performance. However, JOL couldn't monitor this effect.

EXPERIMENT 2

Method

Participants

Seventy-two undergraduate students (36 women and 36 men) from two universities in Beijing, China participated in Experiment 2. They ranged in age from 18 to 23 years ($M = 19.23$ years). All reported normal or corrected-to-normal vision. None participated in Experiment 1.

Design

A 2 × 2 (representation type × presentation media) between-participants design was used. There were four experimental conditions: (a) word-only presented via computer (W), (b) word-plus-picture presented via computer (WPC), (c) word-only presented via paper (WP), and (d) word-plus-picture presented via paper (WPP).

Materials

The materials were the same 40 pairs of cue-target words and corresponding pictures as used in Experiment 1.

Procedure

The basic procedure in the conditions of computer presentation was identical to Experiment 1 except that: (a) participants pressed the spacebar to regulate studying pace, and (b) they wrote their recalled target on the paper.

Participants studied the cue-target pairs printed on a card one by one at their own pace in the conditions of paper presentation. After studying all 40 cue-target pairs once, they wrote their JOLs, and finally the target words on paper.

Measures

The methods to compute the recall and the magnitude and accuracy of delayed-JOL were identical to Experiment 1.

Results and Discussion

Recall Performance

Percentage of correct recall was computed for each participant. Means across participants' values are reported in Table 2. A 2 × 2 ANOVA revealed a marginal significant effect of presentation media on recall performance, $F(1, 67) = 3.74$, $p = .058$. The main effect of representation type was not significant, $F(1, 67) = 1.15$, $p > .05$; and the interaction between representation type and presentation media was not significant either, $F(1, 71) = 0, p > .05$.

Magnitude of Delayed-JOL

To evaluate the effects of representation type and presentation media on the magnitude of delayed-JOL, we computed the median across each participant's JOL. A 2 × 2 ANOVA revealed that representation type had no influence on delayed-JOL, $F(1, 71) = 0.79, p > .05$; the main effect of presentation media was not significant, $F(1, 71) = 0.22, p > .05$; and their interaction was not significant either, $F(1, 71) = 0.09, p > .05$.

Table 2. Mean Proportion of Correct Recall and Magnitude of
Delayed-JOL in Experiment 2

Presentation media	Representation type									
	Word only					Word-plus-picture				
		Recall		JOL			Recall		JOL	
	N	M	SD	M	SD	N	M	SD	M	SD
Computer	18	73.64	18.14	67.94	26.28	18	76.14	20.02	72.49	20.14
Paper	18	63.19	22.78	59.31	26.11	18	72.50	20.84	71.25	19.69

Accuracy of Delayed-JOL

JOL-recall gamma correlations calculated over the entire list of 40 items averaged 0.78, significantly different from 0, $t(71) = 18.51, p < .001$, which indicated that the participants did not make a judgment based on a guess only. The bias score (6.90) was greater than 0, $t(71) = 4.53, p < .001$. Participants were reliably overconfident in each condition. The main effect of presentation media was marginally significant, $F(1, 68) = 3.26, p = 0.07$. This conclusion meant that participants overestimated their recall via computer presentation more than via paper presentation.

Recall performance was better in computer conditions than in paper conditions, and better in self-paced learning conditions than in fixed-rate learning conditions. However, JOL couldn't monitor this effect, which was consistent with Experiment 1.

GENERAL DISCUSSION

Recall Performance

Representation type had significant influence on recall performance in Experiment 1. The recall performance was higher in word-only conditions than in word-plus-picture conditions. The result confirmed the split-attention effect. The participants might not process the unrelated pairs deeply enough when presented fixed-rate in Experiment 1, and could not build the referent connection between the words and the pictures. But the participants might process the unrelated pairs at their own pace in Experiment 2; therefore, there was no significant difference between the word-only and word-plus-picture conditions on recall performance.

Accordingly, recall performance in WA condition was the highest in the four conditions of Experiment 1. That is to say, the performance increased when words were presented in visual combined with auditory mode. It is similar to the prediction of the dual-coding theory of Paivio (1990) that information processed verbally and visually should increase the memory performance. However, it is critical to note that the performance in WAP condition was the lowest in the four conditions of Experiment 1, which means that recall performance decreased by using pictures. According to Baddeley's working memory model, students learn better when words and pictures are presented in separate modality than presented in the same modality. When pictures and words are both presented visually, working memory would be overloaded because the amount of information to be coordinated or integrated at a time exceeds the capacity of the central executive. Participants find it difficult to connect the two unrelated words because the irrelevant visual features hinder their association.

Computer-based information presentation appears to offer general learning advantages over the traditional paper presentation. The recall performance was higher in computer presentation than in paper presentation in Experiment 2. Salomon, Perkins, and Globerson (1991) examined how a computer aided cognitive processing, and suggested that a computer could support intellectual performance and enrich individuals' minds. The interface of a computer is active, flexible, and easy to maintain a high level of interest so it reduces the user's cognitive load. Generally, a computer is an effective instructional facility that can reduce time needed to complete tasks and improve the level of learning.

Magnitude of JOL

In the two experiments of this study, the results of people's judgments confirm our prediction based on Koriat's model. Representation type, presentation channel, and presentation media all had no influence on the magnitude of JOL. Considering the significant main effects of representation type in Experiment 1 and presentation media in Experiment 2 on recall performance, the influence of representation type and presentation media were discounted in predicting recall performance by immediate-JOL in Experiment 1 and delayed-JOL in Experiment 2, respectively.

The interpretation of this result pattern, however, might take two different lines. One account, aligned with the explanation of Koriat's cue-utilization model, assumed that JOL was not sensitive to the effect of extrinsic cues. In each condition of this study, participants made JOL according to the memory effects of different items but not different multimodal presentations. Furthermore, JOL was participants' subjective feeling so it was possible for participants to use the same whole number indicating different memory efficacy in different conditions.

The other account could be derived from Baddeley's working memory model. Only when the on-line information retrieved from different sources accumulated in episodic buffer could participants apply to make judgments. Considering the limited capability of working memory, participants could use limited information although the information presented was different in different experimental conditions so that their JOLs had no significant difference.

Absolute and Relative Accuracy of JOL

The results showed that the relative accuracy was higher in Experiment 2 (0.78) than in Experiment 1 (0.34). It was consistent with the findings of Nelson and Dunlosky (1991). They also found that the relative accuracy of delayed-JOL was higher than that of immediate-JOL. This delayed-JOL effect has important implications in education. Namely, to best predict eventual recall, a person should make JOL shortly after study instead of immediately after study (Dunlosky & Nelson, 1992). Furthermore, the participants were relatively well calibrated in Experiment 2, and predicted their memory effect more accurately than did the participants in Experiment 1. It indicated that the accuracy of judgment could be increased in self-paced learning conditions.

This pattern of findings could also be well explained in Baddeley's working memory model by assuming that when people access the likelihood of eventual recall for recently studied information, they need to simultaneously monitor both working memory and long-term memory. All the information should be integrated in episodic buffer, and people can make judgments based on such information. People make immediate-JOL mainly based on the information on-line they can use. However, the recall performance needed the information mainly from long-term memory. This is the reason that the accuracy of immediate-JOL was reduced. By contrast, the information people needed to make delayed-JOL has better interrogation of long-term memory via retrieval of the information contained therein.

Participants were all overconfident and this effect was robust in two experiments. People tend to overestimate the precision of their knowledge (Alpert & Raiffa, 1982). Researchers also found that people overestimate their ability to do well on tasks and these overestimates increase with the personal importance of the task (Frank, 1935). People are also unrealistically optimistic about future events. Taylor and Brown (1988) argued that exaggerated beliefs in one's abilities and unrealistic optimism might lead to "higher motivation, greater persistence, more effective performance, and ultimately, greater success." Being identical to above opinions, we found the recall performance was better in computer conditions than in paper conditions. We can conclude that multimodal information presentation makes people feel that it is easy to learn and they can maintain long attention, which will benefit the learning process and increase the learning performance.

Summary

A major goal of the present research was to empirically evaluate the influence of multimodal information on memory performance, JOL, and its accuracy. We found that, most notably, a split-attention effect occurred when words and pictures were presented in the same modality simultaneously in Experiment 1. Recall performance was better in computer-based learning than in paper-based learning, which indicated that the computer is an effective tool to use for learning. In the present study, multimodal information had no influence on magnitude of people's judgments. The participants were all overconfident in all conditions, but more overconfident in computer conditions than in paper conditions.

REFERENCES

Ainsworth, S., Bibby, P. A., & Wood, D. (1998). Analyzing the costs and benefits of multi-representational learning environments. In M. van Someren, E. Boshuizen, T. de Jong, & P. Reiman (Eds.), *Multiple representations for problem solving and learning* (pp. 103-208). London: Elsevier.

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and Biases.* Cambridge University Press, Cambridge and New York.

Askwall, S. (1985). Computer supported reading versus reading text on paper: A comparison of two reading situations. *International Journal of Man-Machine Studies, 22,* 425-439.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Begg, I., Duft, S., Lalonde, R., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language, 28,* 610-632.

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

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.

Claik, J. M., & Paivo, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3,* 149-210.

Dunlosky, J., & Matvey, G. (2001). Empirical analysis of the intrinsic-extrinsic distinction of judgments of learning (JOLs): Effects of relatedness and serial position on JOLs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(5), 1180-1191.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOLs) and the delayed-JOL effect. *Memory & Cognition, 20,* 374-380.

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language, 33*(4), 545-565.

Ednerson, M., & Tumey, D. (1984). An experiment with speech on an electronic text service. *Videodisc and Optical Disc, 4*(4), 306-317.

Edwardson, M., Groms, D., & Pringle, P. (1976). Visualization and TV news information gain. *Journal of Broadcasting & Electronic Media, 20,* 373-380.

Frank, J. D. (1935). Some psychological determinants of the level of aspiration. *American Journal of Psychology, 47,* 285-293.

Gerlič, I., & Jaušovec, N. (2001). Differences in EEG power and coherence measures related to the type of presentation: text versus multimedia. *Journal of Educational Computing Research, 25*(2), 177-195.

Gunter, B. (1987). *Poor reception: Misunderstanding and forgetting broadcast news.* Hillsdale, NJ: Lawrence Erlbaum.

Hegarty, M., & Steinhoff, K. (1997). Individual differences in use of diagrams as external memory in mechanical reasoning, *Learning and Individual Differences, 9,* 19-42.

Hsia, H. J. (1969). Intelligence in auditory, visual, and audiovisual information processing. *AV Communication Review, 17,* 272-282.

Institute of Language Teaching and Research (1986). *A frequency dictionary of Modern Chinese.* Beijing: Beijing Language Institute Press.

Kelemen, W. L., & Weaver III, C. A. (1997). Enhanced metamemory at delays: Why do judgments of learning improve over time? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(6), 1394-1409.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology, General, 26,* 349-370.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65-100.

Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 423-434.

Lichtenstein, S., Fischhoff, B., & Phillips, L. S. (1977). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Mason, B. J., Patry, M., & Bernstein, D. J. (2001). An examination of the equivalence between non-adaptive computer-based and traditional testing. *Journal of Educational Computing Research, 24*(1), 29-39.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83,* 484-490.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84,* 444-452.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 2*(87), 319-334.

Nelson, T. O. (1984). A comparison of current measure of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95,* 109-133.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science, 2,* 267-270.

Paivio, A. (1990). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255-287.

Powers, D. E. (2001). Test anxiety and test performance: Comparing paper-based and computer-adaptive versions of the Graduate Record Examination (GRE(c)) General Test. *Journal of Educational Computing Research, 24*(3), 249-273.

Reimann, P. (1999). The role of external representations in distributed problem solving. *Learning and Instruction, 9*, 411-418.

Rice, G. E. (1994). Examining constructs in reading comprehension using two presentation modes: Paper vs. computer. *Journal of Educational Computing Research, 11*(2), 153-178.

Rogers, Y. (1999). What is different about interactive graphical representations? *Learning and Instruction, 9*, 419-425.

Salomon, G., Perkins, D. N., & Globerson, T. (1991). Patterns in cognition: Extending human intelligence with intelligent technologies. *Educational Research, 20*, 2-9.

Shu, H., Cheng, Y. S., & Zhang, H. C. (1989). Name agreement, familiarity, image agreement and visual complexity for 235 pictures. *Acta Psychologica Sinica, 4*, 389-396.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for naming agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174-215.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science, 19*, 97-140.

Sweller, J., & Chandler, P. (1994). Why is some material difficult to learn? *Cognition and Instruction, 12*, 185-233.

Taylor, S., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.

Yates, J. F. (1990). *Judgment and decision making.* Englewood Cliffs, NJ: Prentice Hall.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science, 21*, 179-217.

Direct reprint requests to:

Dr. Xiaolan Fu
Institute of Psychology
Chinese Academy of Science
Beijing 100101
China
e-mail: fuxl@psych.ac.cn