The Role of Working Memory and External Representation in Individual Decision Making

Jozsef A. Toth
Learning Research and Development Center and
School of Information Science
University of Pittsburgh
Pittsburgh, PA 15260, USA
Tel: 1-412-624-2189
E-mail: jtoth+@pitt.edu

C. Michael Lewis
School of Information Science
University of Pittsburgh
Pittsburgh, PA 15260, USA
Tel: 1-412-624-9426
E-mail: ml@sis.pitt.edu

Abstract

Exploratory research into the role of visual and verbal working memory in diagrammtic reasoning is presented. Eighty subjects participated in an experiment where thirty-four different gain/loss problems were represented in sentential and graphical form. In order to test the role of each component of Baddeley's model of working memory, subjects performed verbal, visual, and mental suppression tasks while reasoning about sentential and graphical gain/loss problems, yielding six different conditions. In two additional control conditions, no suppression tasks were performed. Interference and preference reversals occurred in all six conditions involving suppression tasks, although no interference was predicted in conditions involving the graphical representation and verbal suppression task, or the sentential representation and visual suppression task. In the control conditions subjects made responses consistent with prospect theory with little interference. The data suggests that certain gain/loss problems require both visual and verbal resources and that the taxing of these resources results in strategies that may favor a minimal inference strategy.

INTRODUCTION

Much has been written about short-term memory (hereafter *working memory*) in planning, problem solving, and decision making. In the realm of visual and diagrammatic reasoning, however, very little empirical evidence exists describing the role of this phase of human information processing. The goal of this work is to understand the relationship between verbal and visual isomorphic representations and the different components of working memory that may manage relevant information as a decision maker reasons about a given task.

Baddeley's influential model of working memory consists of three components: the articulatory rehearsal, or phonological loop, the visuo-spatial sketchpad, and the executive control system (Baddeley 1986; Baddeley 1992). It is also assumed that working memory is defined as an activated subset of long-term memory.

Rather than an intermediate repository for encoded data, working memory is part of a closed-system which is instead governed by the central executive. By performing dual task experiments whereby a load is placed on one of the three components, it has been demonstrated that, at some level of abstraction, these visual (sketchpad), verbal (articulatory loop), and controller-like (central executive) components indeed exist.

The work described herein will test this model of memory against behavioral data from a well known theory in the decision making literature known as *prospect theory.* According to the basic tenets of this theory, decision-making individuals tend to be risk-averse in situations involving gain and risk-seeking in situations involving loss, when deciding on alternatives such as the following:

(gain) 4,000 with probability .8; 3,000 with probability 1.0
(loss) -3,000 with probability .9; -7,000 with probability .45

Thus, in a situation involving gain, the decision maker will prefer the more certain outcome (3,000 with probability 1.0) over the riskier or less certain outcome (4,000 with probability 0.8). Likewise, in a situation involving loss, the decision maker will prefer the riskier loss (-7,000 with probability .45) over the more certain loss (-3,000 with probability .9). Although the certain loss of -3,000 has a higher expected utility (-2,700), decision makers typically go for the less certain option with the lower expected utility (-3,150). The received interpretation of the theory suggests that the decision maker's *perception* of the problem alternatives—rather than adherence to normative principles such as Expected Utility Theory—contributes to the counterintuitive results in this classic set of experiments. Decision makers are thought to view the problems in terms of gains and losses. These perceptions are framed according to a neutral reference point, corresponding to the current asset position, which is assumed to be a null wealth (i.e., $0). One principle that appears to contribute to this phenomenon is the *overweighting of certainty.* In their words, "it appears that certainty increases the aversiveness of losses [by eschewing the certain loss and favoring the less probable loss] as well as the desirability of gains [by favoring the certain gain and eschewing the less probable gain]" (Kahneman and Tversky 1979, pg. 269). More recent experiments by Tversky and Kahneman (1992) have revealed that the effect is re-

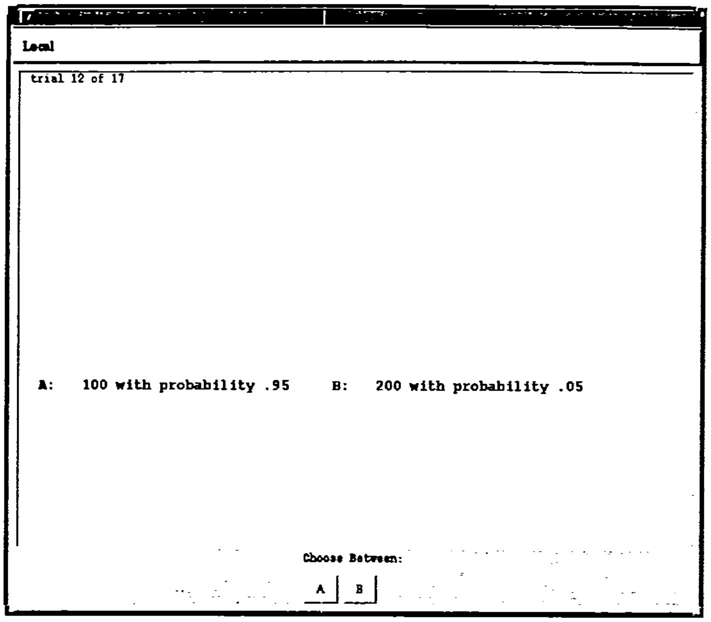

Figure 1: Layout of workstation screen for the experiment illustrating sentential representation of gain-loss problem.

versed in cases involving low probabilities; risk seeking for gains, and risk averse for losses.

Three aspects of this problem that have received little or no attention, however, pertain to (a) the external representation of the gain–loss problem, (b) the relationship between the external representation and various cognitive and perceptual resources, most notably working memory, and (c) how the external representation bears on the corresponding internal representation, whether it be visual or verbal. To date, gain–loss problems have historically been represented in a quasi-tabular sentential form.

Regarding prospect theory and the gain–loss problem, at issue is which components of working memory, visual, verbal, or executive, are involved in the process of (1) viewing the external representation of the problem, (2) internal activation of the external problem constituents in either visual memory (sketchpad), verbal memory (articulatory loop), or both, (3) drawing inferences based on the information present in visual or verbal working memory, then making a decision (central executive), and (4) providing a response. These four stages of processing closely follow the experimental paradigm of Sternberg (1969). In this case, however, Baddeley assumes that a systemic interaction exists among these four components whereas in the latter case, processing has been assumed to transpire serially, from (1) to (4). It has been determined in recent work in syllogistic reasoning that taxing the central executive component of working memory has effects on the decision-making process but taxing the sketchpad and articulatory loop does not (Gilhooly et al. 1993). To be sure, experimental task design is paramount.

In the case of the gain–loss problem it is expected that marked changes in decision-making behavior will result if any of the three memory components are similarly taxed. Since the standard representation of the problem is primarily sentential, it is assumed that verbal working memory will be most utilized during the solving of the problem. It can, however, also be assumed that certain individuals will also consider such problems in a visual sense as well. Since magnitudes of wealth are considered, a decision maker may also associate a gain or a loss with a personalized mental depiction of the problem constituents (see also Larkin and Simon 1985; Koedinger and Anderson 1990).

Also to be considered is the external representation of the gain–loss problem. The original representation, as mentioned, is in a quasi-tabular sentential form. Re-representing the problem into a graphical form should have certain effects on the activation of visual versus verbal working memory. A graphical representation of the same problem should result in a higher activation of visual memory (sketchpad), but again, aspects of the problem may also activate verbal memory (articulatory loop) to a lesser extent, depending on the individual's predisposition to considering visual elements in a verbal fashion.

HYPOTHESES

Sentential External Represenation: A problem externally represented in sentential form should result in a primary activation of verbal working memory. Since the quantities ordinarily reasoned about, probability and payoff, are postulated to be active in verbal working memory, and to a much lesser extent in visual working memory, a taxing of verbal resources via an articulatory suppression task should inhibit the reasoning process. This should result in interfer-

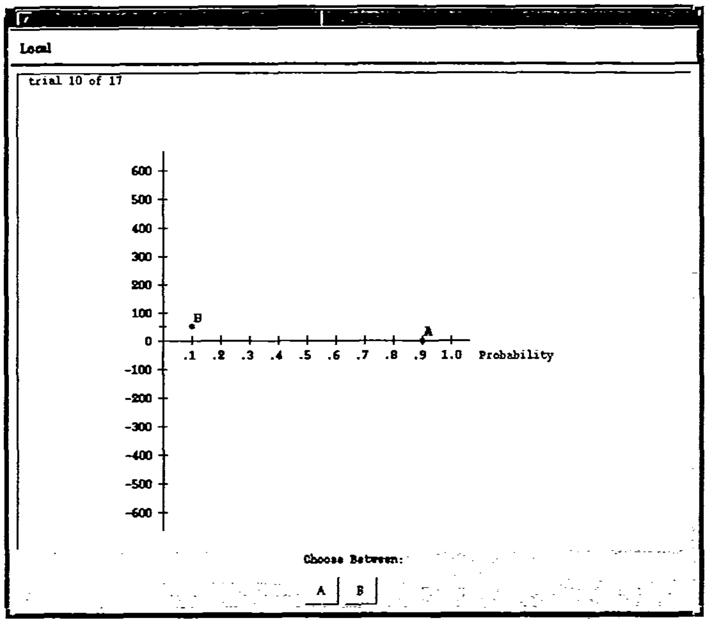

Figure 2: Layout of workstation screen for the experiment illustrating graphical representation of gain-loss problem. Payoff and probability of choices A and B are denoted spatially in a two-dimensional x, y coordinate system.

ence and decisions that are inconsistent with prospect theory. The same should hold true for the loading of the central executive with a secondary task since this phase of processing (i.e., Sternberg's phase (3)) is presumed to be involved. If an attempt is made to load visual working memory with an articulatory suppression task, however, decisions should remain more consistent with prospect theory since it is presumed that with a sentential external representation, the main interaction lies between the verbal and central executive components of the working memory system.

Diagrammatic External Representation: A problem externally represented in graphical form should result in a primary activation of visuospatial working memory. Since the quantities reasoned about, probability and payoff, are postulated to be active in visual working memory, and to a lesser extent in verbal working memory, a taxing of visuospatial resources should inhibit the reasoning process and result in decisions that are inconsistent with prospect theory. The same should hold true for the loading of the central executive with a secondary task. If an attempt is made to load verbal working memory is loaded with a visuospatial suppression task, decisions should remain consistent with prospect theory since it is presumed that with a visual external representation, the main interaction lies between the visual and central executive components of the working memory system.

METHOD

Design

There were two independent variables in the 2x4 factorial design, (A) external representation of the gain–loss problem, and (B) working memory load. External representation of the gain-loss problem comprised two levels, (1) the default sentential representation illustrated in Figure 1, and (2) a diagrammatic representation in which probability and payoff were presented in a graphical form (Figure 2). Working memory load comprised four levels: (1) Control, no load on the working memory system, (2) load on the articulatory loop with a secondary articulatory suppression task, (3) load on the visuospatial sketchpad with a spatial suppression task, and (4) load on the central executive with a secondary suppression task. Thus, there were eight total experimental conditions. For the purposes of clarity, the eight conditions will be indicated as follows: *t* sentential text representation with no suppression task, *tartic* text with articulatory suppression task, *tsketch* text with spatial suppression task, *tcont* text with executive controller suppression task, *g* graphical representation with no suppression task, *gartic* graphics with articulatory suppression task, *gsketch* graphics with spatial suppression task, and *gcont* graphics with executive controller suppression task.

Thirty-four gain-loss problems were randomly selected from a pool of seventy gain-loss problems present in the literature (Kahneman and Tversky 1979; Tversky and Kahneman 1992, pg. 307). Each of the thirty-four problems was represented once in a text-based condition and again in a graphics-based condition. The subject pool was divided into two groups. The first group performed conditions *t*, *gartic*, *tsketch*, and *gcont*. The second group performed conditions *g*, *tartic*, *gsketch*, and *tcont*. The four conditions in the experiment were presented as randomly ordered blocks and the seventeen trials within each block were also random-

Pair: ID (A;B)	*Expected*	*t v. g*	*t v. tartic*	*t v. tsketch*	*t v. tcont*	*g v. gartic*	*g v. gsketch*	*g v. gcont*
0 (.1,0; .9,50)	B				.066	.090		
3 (.5,0; .5,-50)	A		.090					
4 (.9,0; .1,50)	B						.073	**.001**
11 (.5,0; .5,-100)	A			.090				
13 (.75,0; .25,-100)	A	.036				.002	.014	
15 (.95,0; .05,-100)	A					.090		
16 (.01,0; .99,200)	B			.090	.090			
19 (.1,0; .9,-200)	A			.036	.090		.090	.036
21 (.5,0; .5,-200)	A				.090	.014	.090	.036
23 (.9,0; .1,-200)	A	.036	.036	.090	.090			
24 (.99,0; .01,200)	B	**.108**				**.051**	**.051**	
25 (.99,0; .01,-200)	A	.014	.090	.036	.090			
26 (.01,0; .99,400)	B		.090		.006			.066
27 (.01,0; .99,-400)	A					.090		
28 (.99,0; .01,400)	B					**.108**	**.055**	**.108**
31 (.1,-50; .9,-100)	A						.073	
35 (.9,-50; .1,-100)	A					.016		
41 (.5,-50; .5-150)	A						.090	
43 (.75,-50; .25,-150)	A		.088	.088				.073
47 (.05,-100; .95,-200)	A						.030	
50 (.5,100; .5,200)	B				.030			
56 (.8,4000; 1.0,3000)	B			**.055**				
66 (.5,1000; 1.0,500)	B			**.055**	**.108**			**.056**

Table 1: Significant results from experiment based on response frequencies (*N=20 observations per condition, df=1*). *p* values in table entries depicted in plain font indicate interference, those in boldface indicate a preference reversal.

ized.

Subjects

Eighty subjects participated in the experiment, comprising graduate, undergraduate, and staff at the University of Pittsburgh. The subjects were assigned to each experimental session based solely on the criteria of availability. Subjects were paid five dollars or received extra credit for a course, which had been prearranged with the course instructor. In some cases, the subjects volunteered and did not receive any compensation.

Materials

The subjects were seated at a DEC 5000 Workstation with a 13 inch chromatic monitor. The software for the experiment was implemented in Harlequin Common Lisp using the Harlequin Common Application Programming Interface in the X Windows display environment. In both the sentential (Figure 1) and graphical (Figure 2) conditions, the gain loss constituents were depicted in black, alphanumeric characters were displayed in 10, 12 and 14 Point Courier Roman Bold font with a white background. In the graphical condition, probability was represented on the x axis and payoff was represented on the y axis (Figure 2). The two choices A and B along with the corresponding x, y dot were displayed in red and the rest of the display constituents (axes, tick marks, etc.) were in black on a white background. The design was intended to maintain as many isomorphisms between the sentential and graphic representations as possible (Zhang and Norman 1994; Lewis and Toth 1992; Kotovsky et al. 1985). Subdivisions of payoff and probability were abstracted to minor hash-marks rather than an explicit value (e.g., see payoff for choice B in Figure 1). However, in the case of the probabilities .01 and .99, they were made explicit in the 10 point font on the probability scale.

Procedure

In the articulatory suppression task, subjects uttered out loud the same sequence of digits (*"1, 2, 3, 4, 5"*) while presented with a sentential (i.e., *tartic*) or graphical (i.e., *gartic*) gain–loss problem in the computer display. The subject was instructed to synchronize with an activated metronome at the rate of one utterance per second. When the subjects were ready to choose one of the alternatives, they clicked either the "A" or "B" button in the display using the mouse. The problem disappeared from the display, and a new problem was presented. For the visuospatial suppression task, subjects typed the sequence of four characters on the computer keypad portion of a keyboard in a clockwise circular fashion (4, 8, 6, 2) with their non-dominant hand at the rate of one per second; again to the beat of the metronome. While still typing the sequence, the subject was presented with a gain–loss problem either in sentential (*tsketch*) or graphical form (*gsketch*). The subject then chose either "A" or "B", the problem disappeared from the display, and a new problem was presented. For the executive controller suppression task, subjects were instructed to utter a random sequences of digits from the set of numbers 1–5 (*e.g., "2, 1, 4, 5, 3 ..."*) in

Representation/Load	No Load	Articulatory	Visuopatial	Controller
Text	4.792	4.673	5.246	6.076
Graph	3.636	3.845	4.290	4.507

Table 2: Average response times (sec.) for eight experimental conditions (*N=680 per condition, F(7,79)=64.46, p=.0001*).

which each new sequence was unique, while presented with a gain–loss problem either in sentential (*tcont*) or graphical (*gcont*) form. Selection and problem display proceeded in the same manner as the other two tasks just described. In the two control conditions, subjects simply selected "A" or "B" from the sentential (*t*) or graphical (*g*) representation without performing a secondary task. The metronome remained active through the entire experiment as a task invariant. A block of practice trials was presented before the experiment began which allowed the subject to learn each suppression task. All responses were written out to disk for subsequent analysis.

RESULTS

Tables 1–3 and Figure 3 summarize the data collected from the experimental trials. *N=20* samples were recorded from the eighty subjects for each of the eight conditions. Table 1 displays the results of the Likelihood Ratio χ^2 analysis using the *SAS* FREQ procedure. Each cell reflects the *p* value which is the result of the pairwise comparison of two conditions. Ten pairs of conditions were analyzed.

Only those gain/loss problems in which the observed responses varied significantly from the expected responses appear in the table along with the ID number from the original pool of seventy. The expected responses were derived from condition *t*. The observed responses were organized into two different categories, interference and reversal. *Interference* responses are shown in the Table as *p* values in plain typeface. These responses varied significantly from the expected responses, but more than fifty percent of the twenty possible observed responses were still congruent with the expected response. These kinds of responses suggest error in which subjects were not able to provide the expected response owing to interference from the secondary task. The second kind of response, called a *reversal* response, is indicated with a *p* value in boldface type. With a reversal, more than fifty percent of the observed responses were incongruent with the expected response. For example, in the *g v. gcont* pair of gain/loss problem 4, where the expected response was B, in *g* 3 subjects chose A (.9,0), and 17 subjects chose B (.1,50), but in *gcont,* 13 subjects chose A, and 7 chose B. This data is highly suggestive of a reversal owing to the controller suppression task (*p=.001*). In the *g v. gartic* pair of gain/loss problem 35 the responses were evenly divided for A and B. Regarding the *t v. g* pair, three interferences occurred, with only one reversal (problem 24) where the expected response was considered choice B, the higher expected utility. In summary, interferences and reversals occurred in many of the problems, but did not exactly conform to the postulates enumerated earlier.

Response times for each of the eight conditions are summarized in Table 2. The response times for the eight conditions varied significantly (*F(7,79)=64.46, p=.0001*). Within all eight conditions involving *t* and *g*, response times varied when loading the sketchpad and the controller, but not the articulatory loop. Between all the conditions involving *t* and *g*, response times were consistently and significantly longer for the four *t* conditions compared to the four *g* conditions. This suggests that the graphics-based reasoning was more effortless than the text-based reasoning. For instance, the difference between the means of *tcont* and *gcont* was 1.569 seconds (*p=.0001*). Many subjects reported that the random number generation task was the most difficult of the three suppression tasks, even though there was no significant difference between *t*, presumably the easiest of the four text-based conditions and the controller-loaded *gcont* condition, presumably the most difficult of the four graphics-based conditions.

DISCUSSION

The data described in the last section clearly indicates that the tasks designed for this experiment caused interference where none was expected and lacked interference in cases where it was predicted. For example, surprisingly little interference occurred in the *tartic* condition, but even proportionately *more* interference occurred in *gartic,* where hardly none was predicted. Likewise, the visuospatial suppression task caused interference in *tsketch* where little was thought to occur but interference did occur in *gsketch.* As predicted expected interference occurred in *tcont* and *gcont.*

Regarding the articulatory suppression tasks, two explanations readily come to mind. In *tartic,* it may simply be the case that the repetition of the number sequence was not sufficient to cause the expected interference. In earlier experiments by Baddeley, interference increased only as the verbal task became more difficult. Regarding *gartic,* there are components of the graphical representation that are clearly verbal in nature, and not as automatic as the verbal information presented in *tartic.* Thus, it might be the case that in the course of extracting available information from the graph, probability and payoff, this extra processing is not as automatic as it is in the *tartic* case. It can be easily assumed that non-automatic information becomes ensnared in verbal working memory whereas automatic information does not (Shiffrin and Schneider 1977). More importantly, however, is the idea that *both* visual and verbal working memory appears to be active during the visual reasoning process.

Puzzling however, is the interference, even reversals, generated in the *tsketch* condition (problems 56 and 66, *p=.055*). There appears to be a spatial component to the reasoning about the problem when presented in its sentential form; perhaps more than previously expected. Given the

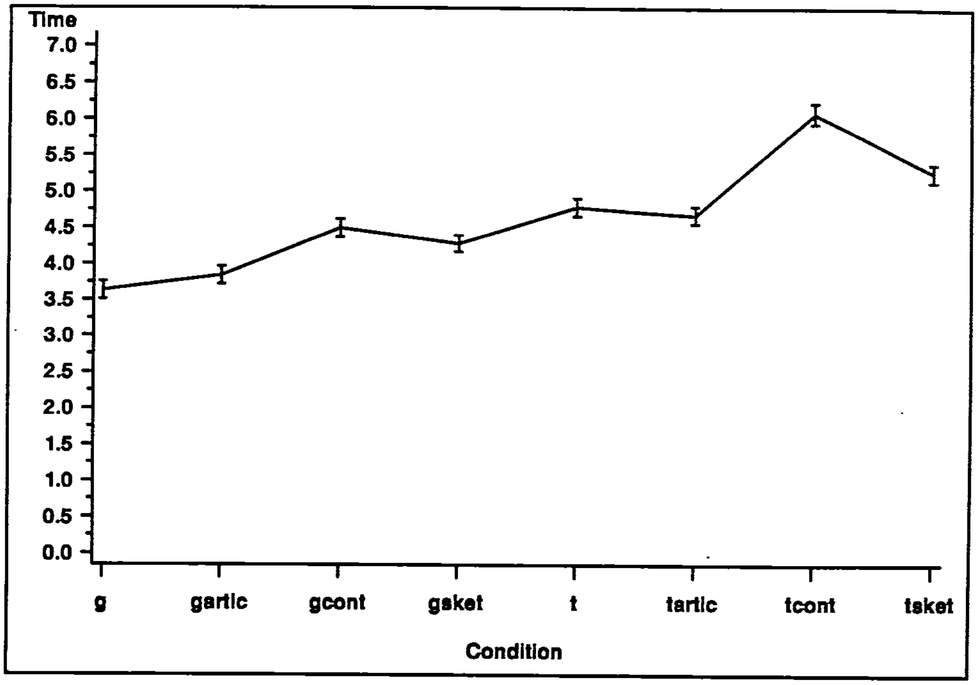

Figure 3: Response times (sec.)

template:

A: w with probability x; B: y with probability z

the data suggests that the decision makers utilized the spatial relationships among the choice constituents to arrive at a decision. The secondary typing task and the random number generation task, are activities that a typical subject performs on a less regular basis than talking and acting at the same time. In fact, many subjects exhibited great difficulties with these two tasks as opposed to the latter task. Thus, the interference appears to be genuine. In contrast, a few subjects were able to chew gum *and* shake one leg with a nervous tick *while* performing the secondary task *and* deciding on the choice alternatives, with relative ease—i.e., *four* tasks at once.

More puzzling is the fact that interference occurred in some problems, but not in other, similar problems, and in some conditions, within a particular problem, but not in others. The thinking at the time of this writing is that there were spatial and informational features unique to some problems that were highly dependent on the full availability of resources whereas in other cases, such resources were not an issue.

Readers can convince themselves of the power of interference by repeating random numbers while trying to choose between *A* and *B* in Figures 1 and 2. Particularly in Figure 2 (gain/loss ID 4), in *gcont* where the mind is partially consumed with having to think of the next random number to say (from the set 1,2,3,4,5), different percpetual strategies appear to pop-up that would not ordinarily be pursued. Thus, in Figure 2 even though choice *B* makes the most irrefutable sense (choice *A* lacks any payoff), the heuristic that seems to arise is the selection of the choice that is farther to the right, a payoff that is not associated with a minor hashmark (as it is in *B*, which requires additional visual inference), and that is anchored on the X axis. This is substantiated by the reversal in *gcont* ($p=.001$) and the interference in *gsketch* ($p=.073$). It is presumed that such unique strategies account for much of the interference and reversals present in the data and is referred to as the "minimal inference" strategy.

Subjects reported developing other visual strategies for some of the problems. For instance, a payoff further to the right and to the top of the display was better than any other payoff. Indeed, interference with such heuristics does not appear in the data. Space does not permit an individual analysis of each gain-loss problem, but it is hoped that some of the thoughts provided in this section convince the reader that general theories about verbal and visual representations and their memory requirements are not yet within our reach. Only further research and the tuning of task and task constituents will further reveal the relationships among the entities described in this report. For instance, the addition of a third probability/payoff choice, *C*, would futher tax working memory.

CONCLUSIONS

In summary, it appears that in this study the visual, verbal, and reasoning components of memory were active when reasoning about sentential and diagrammatic representations. In tasks that are more visual in nature, such as chess, the components of working memory are more clearly defined such that no interference occurs while performing an

condition	tartic	tsketch	tcont	g	gartic	gsketch	gcont
t	.119	.454	.1284	1.156	.947	.502	.285
	n.s.	*.005*	*.0001*	*.0001*	*.0001*	*.04*	*n.s.*
tartic		.573	1.403	1.037	.827	.382	.166
		.005	*.0001*	*.0001*	*.0001*	*.0001*	*n.s.*
tsketch			.830	1.610	1.401	.956	.739
			.0001	*.0001*	*.0001*	*.0001*	*.0001*
tcont				.2440	2.231	1.786	1.569
				.0001	*.0001*	*.0001*	*.0001*
g					.210	.655	.871
					n.s.	*.0001*	*.0001*
gartic						.445	.662
						.005	*.0001*
gsketch							.217
							n.s.

Table 3: Pairwise significance tests of mean difference absolute values using REGWQ method (*N=680 observations per pair, df=5353.*)

articulatory suppression task, but prohibitive interference occurs when performing a visual suppression task (Baddeley 1992). In the domain described in this work however, it appears that the gain/loss problem, whether represented sententially or graphically, utilizes both visual and verbal components in working memory. More research in this domain will help to further elucidate many of the matters brought to light in this work.

ACKNOWLEDGEMENTS

Many thanks to Walter Schneider, Charles Perfetti, Hermi Tabachneck, Herbert Simon, Alan Lesgold, Terri Lenox, Dan Suthers, and Jeff Shrager. This work was conducted while funded by the Advanced Research Projects Agency Computer Aided Education and Training Initiative, under the title "Collaboration, Apprenticeship, and Critical Discussion: Groupware for Learning," Contract N66001-95-C-8621.

References

Baddeley, A. 1992. Is working memory working? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 44A(1):1–31.

Baddeley, A. 1986. *Working Memory.* New York: Oxford University Press.

Gilhooly, K. J., Logie, R. H., and Wynn, V. 1993. Working memory and strategies in syllogistic reasoning tasks. *Memory and Cognition* 21(1):115-124.

Kahneman, D. and Tversky, A. 1979. Prospect theory: An analysis of decisions under risk. *Econometrica* 47:262–271.

Koedinger, K. R. and Anderson, J. R. 1990. Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science* 14:511–550.

Kotovsky, K., Hayes, J. R., and Simon, H. A. 1985. Why are some problems hard? Evidence from tower of hanoi. *Cognitive Science* 17:248–29.

Larkin, J. and Simon, H. A. 1987. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science* 11:65–99.

Lewis, C. M. and Toth, J. A. 1992. Situated cognition in diagrammatic reasoning. *Reasoning with diagrammatic representations: Papers from the 1992 AAAI spring symposium. Technical Report SS-92-02* 47–52. Menlo Park, CA: AAAI Press.

Shiffrin, R and Schneider, W. A. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84(2):127–190.

Sternberg, S. 1969. The discovery of processing stages: Extensions of Donder's method. *Attention and Human Performance II. Acta Psychologica* 30:276–315.

Tversky, A. and Kahneman, D. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5:297–323.