LEARNING FROM MULTIMEDIA:

THE LOCUS OF MODALITY EFFECTS

A Thesis
Presented to
The Academic Faculty

by

Jesse S. Zolna

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Psychology

Georgia Institute of Technology
December of 2005

LEARNING FROM MULTIMEDIA:

THE LOCUS OF MODALITY EFFECTS

Approved by:

Dr. Richard Catrambone, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Wendy Rogers
School of Psychology
Georgia Institute of Technology

Dr. Bruce Walker
School of Psychology
Georgia Institute of Technology

Date Approved: June 20, 2005

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Research in educational psychology has focused on facilitating learning by using two presentation modalities (auditory and visual) to convey information. Learning is theorized to improve through an increase in perceptual information flow. I hypothesized that presenting information in two modalities might also provide additional benefits that occur after information is perceived, and while it is being processed for learning. The present study explored whether perceptual effects and cognitive effects of multimedia presentation can be separated by presenting auditory and visual information sequentially or simultaneously. During simultaneous presentation, the typical multimedia effect (that is, facilitating learning by presenting information in two modalities) did not occur, suggesting that the multimedia effect might depend upon more than perceptual effects. Moreover, the manipulation showed significant effects of presentation type during sequential presentation, suggesting that effects previously thought to be a result of reducing perceptual overlap might actually occur after perception. Based on the results of this study, I recommend that information designers reconsider the sources assumed to influence the multimedia learning effect. This would have implications for determining the optimal presentation of information.

Chapter I

LEARNING FROM MULTIMEDIA: THE LOCUS OF MODALITY EFFECTS

Educational psychologists have been trying to understand how *multimedia instructional materials*, that is, materials that present to-be-learned information in more than one modality, can improve learning (Mayer, 2001; Mousavi, Low, & Sweller, 1995; Najjar 1998). Multimedia instruction aims to assist learners by leveraging technology to present information in a more compelling, complete, and engaging manner. Benefits to multimedia instructional design are accomplished by mixing media ingredients (e.g., text, narration, diagrams, animations and sounds) that represent a variety of modal forms (e.g., auditory and visual), to assist in the learning process. Multimedia learning has an efficiency advantage over traditional instructional materials that is the result of the potential for the instructional material to provide the same amount of information in a shorter amount of time by, for example taking advantage of a person's ability to simultaneously listen to a narration and look at a diagram. However, this use of technology might overload cognitive resources that limit human information processing capacity and hinder the learner's ability to perceive new information, operate on it, and/or commit it to long-term memory. When the learner's goal is to fully understand all of the information presented, the person must be provided an opportunity to attend to as much relevant information as possible and create meaningful associations among that information.

Though there are advantages to multimedia instructional material, great care has to be taken with its use in the development of learning materials because multimedia can also have cognitive costs (Ainsworth, 1999). For instance, the addition of music and sound effects can interfere with knowledge construction, reducing retention and transfer from a computer based multimedia lesson (Mayer, 2001). So, even though assimilation of complex material can be enhanced by substituting a narration for text to accompany visual instructional materials (Sweller, 1999), the type of information presented via the auditory channel must be carefully considered. Less research has been done regarding types of visual information, but it seems likely that there is some similar interaction. The present investigation will compare learning when descriptive material is presented via a visual medium (text) or an auditory medium (narration). This descriptive material will accompany visual diagrammatic material.

The need for the kind of research presented here is strong. The current knowledge applied to guidelines for developing multimedia materials is woefully incomplete, and dangerously focused on perceptual effects. Researchers have noted that guidelines for choosing how to present information within multimodal user interfaces (e.g., multimedia instructional material) also needs to include principles based on human information processing capabilities (Reeves, Lai, Larson & Oviatt, 2004). Effective learning can be fostered by instructional design that has been developed with a sufficient understanding of how the human cognitive system interacts with instructional material (Schnotz, 2002). Moreover, cognitive science research has implications for the structure and relationship among elements that comprise multimedia information (Drommi, Ulferts & Shoemaker, 2001).

The goal of this study was to examine these topics and to advance the knowledge applicable to finding the mix of media ingredients that will best cater to the strengths and limitations of human information processing. To do so, we must examine the human information processing system with regard to learning, and gather insight on how this relates to the design of multimedia instructional material.

1.1 Learning

Cognitive learning theorists view learning as the acquisition or reorganization of the cognitive structures humans use to process and store information (Good & Brophy, 1990). A cognitive theory of learning is based on a three-stage information processing model. Learning begins with a transient sensory register that receives vast amounts of input from the senses. Sensory input that is important or interesting is selected and transferred to the limited capacity short-term memory store where cognitive operations are required for analysis and retention. For later use (after minutes or years), information is stored in long-term memory (LTM) which has unlimited capacity and can store information without continued maintenance. Learning is the analysis and incorporation of sensory information into the long term memory store. The success of learning is often measured by recall (retrieving information from long term memory).

Robust and efficient learning is thought to result from associating newly presented information with information that already exists in long term memory. Usually, this is the result of an existing chunk of information being extended or altered to accommodate new information. This existing internal knowledge structure is often called a schema (Sweller, 1999). Some materials are "forced" into LTM by rote memorization and over-learning. Deeper levels of processing such as generating linkages between old and new information are much better for successful retention and transfer of material (Anderson, 1995). Moreover, understanding is a result of a learner's active, engaged involvement in the material and process of learning (Jonassen, 1999).

If learning is based on rote effort and elaborateness of processing, poor or inefficient learning would occur when the learner does not sense new information for encoding or does not actively process the information for understanding. Therefore, a lack of cognitive resources for encoding and actively processing new information for committal to long-term memory impedes learning. Accordingly, it seems safe to say that we should not expect learners to understand instructional material that they are not able to process cognitively. Understanding human information processing might be useful in helping instructors understand how to develop instructional materials that are efficient, compelling and complete, but do not exceed learners' processing capacity. This study attempts to expand extant knowledge regarding modality effects by partitioning the observed effects of modality (both cognitive advantages and disadvantages) into effects specific to stages of initial encoding and active processing. Better understanding of effects during each stage separately will lead to better information design.

1.2 Multimedia Learning

Modality effects have been observed for instructional materials that use multimodal perceptual events to present information to be learned. Traditional instructional materials make use of visual objects such as text, diagrams and illustrations. Multimedia instructional material are distinct from traditional instructional materials in that they may add auditory information or animated visual information. In both types of instructional materials, complementary verbal and pictorial information is presented to fully explicate the subject matter. Such a presentation can take advantage of the strengths and avoid the weaknesses of both visual-diagrammatic and auditory-descriptive material. Diagrammatic instructional materials make relationships between elements explicit; the benefit of descriptive information is that it can express and emphasize precise details (Larkin & Simon, 1987). Empirical evidence supports the contention that student learning is affected positively by presenting descriptions and diagrams together (Mayer & Sims, 1994).

Multimedia instructional materials (combining visual and auditory presentations) attempt to exploit the human's ability to sense information in both forms simultaneously, and might increase learning efficiency because it facilitates the processing verbal and pictorial information simultaneously. Verbal information and pictorial information kept simultaneously in working memory allows the learner to make cross-connections between two different (and complimentary) sets of information (Schnotz, 2002). This facilitates elaborative processing, which has been shown to possibly increase the learner's ability to later retrieve the information and to use it more flexibly for more complicated cognitive operations such as problem solving.

Multimedia instructional materials might increase learning efficiency in two ways: presenting information in a form that is complete and easy to process, and by facilitating elaborative processing. To examine how multimedia materials affect these factors of learning, we must examine how multimodal presentation affects the way humans convert sensory information to internal representations. We must also consider how multimodal presentation affects the way these representations are combined. In particular, this investigation focuses on the interaction between the external form (presentation medium) and internal representation (which may be related to information type) of descriptive learning materials. Comparing the benefits of descriptive and diagrammatic information is not the focus of this study (as it was for Larkin and Simon (1987)).

1.3 Multimodal effects on converting sensory information to internal representations

When observing a simultaneous presentation of visual and auditory material, learners must encode and transfer information from both the visual and auditory sensory store to short term memory. Learners guide their attention to visual objects and extract information with their visual sensory-perceptual system (i.e., the eyes, optic nerve, and visual cognitive structures). Visual perceptual encoding takes place rapidly and can extract and briefly store nearly unlimited amounts of information (Potter, 1976). However, this information must be transferred to a more durable form of memory (Jolicoeur & Dell'Acqua, 1999), and every detail cannot be transferred (Cohen, in press). Likewise, auditory information is perceived by the auditory sensory-perceptual system (i.e., ears, cochlea, and auditory cognitive structures). Physical properties of auditory information are also processed quite easily, but processing capacity limitations require the selection of a subset of the information for semantic processing before it is transferred to a semi-permanent form of storage and rehearsed as it awaits cognitive manipulations (Cherry, 1953; Cowan, 1988).

According to the cognitive theory of learning, the sensory memory store receives and holds physical aspects of information temporarily, but more durable memory traces that contain semantic information are held in the short term memory. This memory store holds some representation of both auditory and visual information to be used as a basis for further processing (Busey & Loftus, 1994). New information is effortfuly stored (Jollicuer, 1999) and undergoes operations that associate pieces of the information (Kroll, 1975) and connects them to schemas in long term memory (Sweller, 1999). Cognitive operations relating information are integral to active processing and result in the formation of strong memories. Therefore, from a cognitive learning view, these operations are important to learning. Moreover, information loss is likely to be associated with short term memory because the sensory-perceptual store has been shown to be virtually unlimited, and the long term store has also not been found to have a limit (Anderson, 1985).

Multimedia instructional materials require learners to process information from both auditory and visual sources. With two streams of information, learners can become overloaded. Learners might need to perform additional cognitive operations and/or experience unpredicted interference among information bits when transferring multimodal information to short term memory and while holding this information in short term memory. Due to processing capacity limits for the short term memory system, these additional cognitive processes are likely to impede learners' ability to actively process the new information for committal to long term memory. Therefore, care must be taken to design multimodal materials that do not create inefficiencies. The structure and functioning of short term memory must be examined to understand what would cause these inefficiencies.

1.4 Short term memory, multiple components, and multimedia information design

The semi-permanent short term memory store is the second stage of the common three stage cognitive processing model (Anderson, 1985). It holds information gleaned from the sensory store and is responsible for cognitive operations that will combine new information with old information in the long term memory store for later use. Information processing within this semi-permanent store is limited in capacity, but is necessary for remembering and understanding new information. When a lot of information is included in instructional materials, processing limits necessitate selection of semantic information to be drawn from the physical stimulus. The need for selection limits the amount of semantic information that can be sufficiently encoded and operated on for transfer to the next stage of the processing model. Therefore, selection is synonymous with information loss. Improving the design of instructional materials in a way that would reduce the need for selection would likely improve learning.

A large body of research regarding this stage of human information processing has evolved from the replacement of the concept of a unitary short-term memory system with the concept of a multi-component working memory (Baddeley & Hitch, 1974). According to this theory, the function of working memory is similar to a mental workbench where the cognitive system temporarily stores information and performs cognitive tasks such as transfer of information to long term memory. Baddeley's original conceptualization of this workbench contains components that specialize in certain types of information: the phonological loop receives, holds and operates on *auditory* information; the visual-spatial sketchpad receives, holds and operates on *visual* information. Using this framework of separable portions of working memory, educators have attempted to identify methods for designing multimedia information that increase learning efficiency by combining auditory and visual streams of information. Research related to instructional design has proposed that controlling the processing demand needed in multimedia learning environments might be achieved by spreading information among these stores (Sweller, 1999).

Sweller and his colleagues have shown that multimodal instruction techniques increase learning efficiency. Investigating the effects of a dual mode presentation on learning, they found that mixing auditory and visual stimuli decreases cognitive load and improves learning (Mousavi et al., 1995). They concluded that dual mode presentation reduces the negative effects of split attention when learners have to use one sensory mechanism to search among and integrate multiple physically distinct information streams. Sweller (1999) also claimed that the presentation of information in two modal forms exploits the multi-component working memory system and functionally increases working memory capacity. However, his explanation and ensuing recommendations focused on perceptual level split attention, and ignored whether learning is influenced by the active processing of the new information in short term memory.

Another example of applying a multi-component short term memory framework to instructional design is the Cognitive Theory of Multimedia Learning (Mayer, 2001). This framework suggests the *modality effect* is an important principal for information design. The modality effect states that, in the presence of other visual stimuli such as diagrammatic representations of information, substituting on-screen text for narration detracts from multimedia learning. The modality effect is consistent with a dual-coding theory of multimedia learning which emphasizes that learners build mental connections between visual and verbal representations (Mayer & Sims, 1994). This depiction of the way dual mode instruction improves learning is dependent upon the reduction of interference between internal representations of information. However, Mayer’s approach to taking advantage of multimodal effects emphasizes controlling external representations to increase the bandwidth of presentation, and deemphasizes the cognitive effects of instructional materials on the way the learner actively processes the new information. Like Sweller, the scope of Mayer’s approach concentrates on the first stage of the three stage of the cognitive theory of learning with inadequate focus on multimodal effects on the second and third stages of the cognitive theory of learning.

1.5 Short term memory operations lead to long term storage

The impetus for this research is that the investigations in the domain of educational psychology such as Sweller's and Mayer's have concentrated on the multimodal effects on learning as benefits realized during the transition between the perceptual and working memory stages of learning. However, knowledge acquisition requires that information undergo further cognitive manipulations to be committed to long term memory. Moreover, working memory representations that endure long enough after perception might include the activation of a subset of long-term memory (Norman, 1968), and modality specific components of working memory could result from the temporary activation of elements within long term-storage (Cowan, 1988). Therefore, it seems possible that effects of multimodal instructions occur while information is held and manipulated in working memory as it awaits transfer to long term memory and is dependent upon modal properties of the *internal* representation. To date, there have been few investigations specific to multimedia learning that target modality effects on cognitive operations made while actively processing information for transfer to a long term memory store; extant research focuses on modality effects during the transition from perceptual encoding to working memory.

Research in other psychological domains has explored whether effects of high workload multimodal tasks occur after sensory information has been transferred to working memory. For instance, research has suggested that modality affects workload at the response stage (Wickens & Liu, 1988), and that interference among cross-modal tasks indicates central interference (DeJong, 1993; Jolicoeur 1999). An understanding of modality effects at these stages of learning might be important for optimal information design in multimedia environments. However, extant research in educational psychology has not investigated this question adequately. To date, research in designing multimedia learning materials has largely focused on improving learning by mixing modality with the aim of realizing benefits at the stage of perception and encoding (e.g., Mayer, 2001; Sweller, 1999). Therefore, in the current study I attempt to elaborate on how modality can affect learning based on enabling or preventing the learner to actively encode and process new information in ways that will result in robust long term storage.

1.6 Bottleneck theories of information loss

Sensory memory is capable of storing more or less complete records of what has been encountered at brief periods of time (Anderson, 1995). However, the environment offers more information than we can attend to and much of what enters the sensory store is lost. When a learner is exposed to instructional information that exceeds their processing capacity, the learner develops the need for selection, and consequently information is lost. The need for selection is often attributed to a bottleneck in cognitive capacity: a point or time when the available information cannot be fully processed (e.g., Broadbent, 1958). The exact locus of this filter has been difficult to pinpoint, but evidence for two stages of sensory storage (Cowan, 1988) indicates the possibility of this bottleneck occurring after perceptual information has been encoded into working memory. Previous attempts to understand the effects of multimedia learning have tried to mitigate loss of information during the transition to working memory, but have not experimentally separated the possibility of a bottleneck at this point from the possibility of a second bottleneck during learning.

Semantic properties of the information determine the processing demands required for cognitive manipulations and possibly influence the proportion of information processed before information loss occurs (Corteen & Wood, 1972). Memories must be encoded and strengthened for knowledge acquisition (Anderson, 1995), imposing active processing demands that might employ cognitive resources that otherwise could be used for storage, and thereby causes a second bottleneck that affects the learning process. In addition to the cognitive demands of learning that might influence this later bottleneck, modality properties and relationships between pieces of information might determine processing demand. Simple bottleneck theories do not adequately explain the phenomena of information loss during complex learning from multimedia instructional materials.

1.7 Interaction between internal representations

Wickens' Multiple Resource theory (MRT) was developed as a means to describe human behavior not explainable by a simple filter theory (Wickens, 2002). Multiple resource theory is a theory of multiple task performance whose practical implications stem from predictions regarding a human operator's ability to perform in high work-load, multi-task environments (Wickens, 2002). The value of the model lies in its ability to predict operationally meaningful differences in performance and prompt changes to workload that can be easily coded by the analyst and designer. It is useful in analyzing 'over-load' situations, although it has not been directly applied to the complicated cognitive process of knowledge acquisition. As such, it is an excellent beginning for an approach to the problem of information design for multimedia learning environments.

In a typical application of MRT, an operator must carry out a number of response patterns based on more than one stream of information input. Multiple resource theory operates under the assumption that resources are inherently limited and are allocated to the information sources and response demands as necessary. As tasks increase along a continuum of resource demand (from automated to effortful; see Anderson**,** 1982; Shiffrin & Dumais, 1981), they necessitate increasing tradeoffs and take available resources from other tasks. Converging measures of 'resource demand' have been developed and task characteristics identified that enable resources to be characterized independently from their influence on dual task performance (Wickens, 2002). Wickens states that an important feature of resources is their structure; an important structural distinction is visual-spatial versus auditory. This conceptualization of resource demand and demand characteristics that depend upon the distinction between visual and auditory information are especially pertinent when considering theories relevant to multimedia learning environments.

Multiple Resource Theory includes a four dimensional model of multiple resources, one dimension being *processing codes* (Wickens, 1988). Processing codes are synonymous with internal representations. Differences in processing codes are defined by the distinction between analogue/spatial processes and categorical/symbolic processes. This distinction between codes is important to functioning during perception, working memory and response stages (Wickens, 2002), and depends on working memory representations of information that influence performance. Understanding how internal representations affect the learners' ability to process information is an important aspect of the current investigation, the goal of which is to facilitate the working memory processes that are integral to robust learning by manipulating the level of interaction among separate information streams. Interaction among information streams is thought to be influenced by the relationship among internal representations of information.

A similar view regarding the relationship among internal representations of information forms Paivio's Dual Code Theory. This theory predicts understanding based on internal representations. These representations are designed specifically for internal cognitive manipulations and contain information gathered by the sensory system (Paivio, 1986). The theory is based on the general view that cognition consists of the activity of symbolic representational systems that are specialized for dealing with environmental information, such as instructional material. Two classes of phenomena handled by separate subsystems are specialized for the representation and processing of information concerning non-verbal and verbal information.

According to Paivio, the two subsystems are separate, integrated subsystems that can function more or less independently. This allows independent storage of structural representations (both verbal and non-verbal) that are semantically meaningful pieces of information held longer than sensory information. It also allows independent cognitive processes (functional activities including activation, organization and elaboration, transformation, manipulation, and retrieval) on internally represented and stored structural information for each subsystem. According to Paivio's theory, each system can be active without the other, and processes in one can trigger activity in the other, but neither form of information depends upon the other (Paivio, 1986). In the context of learning from multimedia instructional materials, both subsystems are likely to be active because the learner is presented with both verbal and nonverbal information simultaneously.

As such, Paivio's dual code theory is another example of a human cognitive theory that separates the external representation of information from its internal representation. It is also another example of a theoretical approach that stresses the importance of the relationship between the internal representations of information for human cognitive manipulations. Dual Code Theory defines these relationships by the interaction between two types of representations: verbal and non-verbal. Multiple Resource Theory predicts performance based on properties of internal representations of environmental information, including visual-spatial versus auditory structure.

1.8 Internal representations from multimedia instructional materials

Non-verbal (or visual-spatial) and verbal (or auditory) internal representations often correspond to diagrammatic and descriptive external representations, respectively. The external modality of information has been observed to be one property that determines the efficiency of cognitive resources required to hold and process information. For instance, diagrammatic, visual presentation allows the learner to process more of the instructional material simultaneously (Larkin & Simon, 1987) and there exists evidence that verbal material (e.g., a list of words) is better recalled if presented via auditory means than if presented visually via text (De Haan, Appels, Aleman, & Postma, 2000; Najjar, 1998). However, research investigating multimedia instructional materials in light of Wickens' and Paivio's theories will define internal representations by more than just materials' external representation. The motivation for this study is that visually and auditorily presented information included in multimedia learning environments might correspond imperfectly to the visual and auditory modalities.

The common definition of information *modality* refers to the physical form of presentation for a specific piece of information. However, after conversion of information from its physical/perceptual form to a representational/processing code form, the modal form of information might not be dependent upon its physical form in the world (Schnotz, 2002). Research on modality effects has been performed with the caveat that verbal material can evoke the construction of visual representations, and visual material can evoke the construction of verbal representations (Mayer & Sims, 1994; Paivio, 1986). Moreover, there is evidence that regardless of physical presentation, there exist circumstances in which a given task may be performed using either a verbal or spatial strategy (Wickens & Liu, 1988).

Therefore, it is possible that the *physical* modality of information (e.g., visual or auditory presentation style) will be transformed to a different *mental* modality for one reason or another. Such a transformation might be manifested in sub-vocalizing text or forming a mental image of a verbal description. The external representation of information, its physical modality, is inherent in the presentation mode. Internal representations of information, its mental modality, are inherent to how the information is stored and processed by the observer. This would be consistent with a two stage process for converting environmental information to an internal representation. The two stages would include a brief initial phase providing continued (physical) sensation for up to several hundred milliseconds, and a second phase retaining more processed (mental) sensory information (Cowan, 1988). Candidate reasons learners might transform information are individual differences, to make use of underused processing capacity, because the information is more easily stored in another manner, or because one storage manner will better preserve the relevant content.

Baddeley (1994), Pavio (1986), and Wickens (2002) all consider the mental modal form and interaction of these internal representations of information in their attempts to explain human performance and behavior. Baddeley (1994) separates two parts of working memory that are closely tied to attention and information acquisition. The phonological loop (auditory) and visio-spatial sketchpad (visual) components control separate processing and storage capacity for two distinct types of information. Each is specialized, and one component can lose information when it is overloaded, while the other remains inactive. In addition, the independent central executive is responsible for integrating information from the two separate stores. According to this approach, preserving information by allocating it to underused working memory components will prevent information loss and enable elaborate processing via the central executive.

Wickens and Lui (1988) claim that performance-demand reciprocity depends upon the match between two tasks' fundamental processing structures. Predicting performance-demand reciprocity includes an analysis of the relationship among information based on the dimension of information modality (auditory vs. visual). According to Multiple Resource Theory, cross modal time sharing is better than intra-modal time sharing, but whether external resource factors or internal resource factors are responsible is uncertain (Wickens, 2002). Moreover, differences in internal representations are defined by the distinction between analogue/spatial processes and categorical/symbolic (usually linguistic or verbal) processes, and performance in multi-task environments is accounted for by the separation of processing resources for these two types of information. According to this approach, an actor in a multitask environment will reallocate information for processing based on efficiency of internal representation storage and manipulation.

Paivio's (1986) dual-coding hypothesis claims that image and verbal representations are independently stored and operated on, but both types of representation contribute to knowledge acquisition (Clark & Paivio, 1991). The non-verbal system is a set of structural representations that are presumed to be derived from and retain functional properties of sensorimotor systems, and therefore retain properties of sensory modalities; the verbal system deals with language and associations among non-verbal representations. Paivio (1986) labels internal representations imagens (non-verbal information) and logogens (verbal information) and defines them as subsystem specific bits of information that can be combined and manipulated during human information processing. According to this approach, verbal and non-verbal information are stored separately and combined for learning.

1.9 Dichotomous representations and dual tasks

Three theories of knowledge assimilation propose a dichotomous view of internal representations. Moreover, the dichotomy for each of Baddelay (1994), Pavio (1986), and Wickens (2002) conceptualization of internal representations separate auditory/verbal information from visual/spatial information. Empirical studies demonstrating the advantage of multimodal instruction design compared to unimodal instruction design (e.g. Gellevij, Van der Meij, de Jong & Peiters, 2002; Mayer & Sims, 1994; Mousavi, Lowe & Sweller, 1995) have focused on manipulating the modal form of presentation (*physical* modality) in order to improve learning, but have not converged on the possibility that there are effects on performance that depend upon the modality of internal representations (*mental* modality). A thorough approach to designing information should include a complete understanding of how the modality of information affects both the encoding and processing of information.

These views regarding internal representations are relevant to an investigation of multimedia instructional materials because the learner's goal is to integrate and encode information from two related but distinct streams of information, one verbal (auditory and descriptive) and one non-verbal (visual and diagrammatic). With these two information streams, multimedia learning is in effect a cross-modal, dual task condition. Therefore, research in cross-modal dual task performance should be informative for understanding how multimedia instructional materials affect learning. Unfortunately, research in this area does not point to a single way that dual mode stimuli benefit or detract from the ability of learners to encode, manipulate, and remember information.

In dual task studies, participants are asked to perform two tasks at the same time. Performance for each of the two tasks can be differentially affected by properties of the concurrent task. The implication is that the overlapping processing capacity used by two particular tasks can vary according to some relationship between them. For example, musicians can shadow an auditory-verbal message while sight-reading and playing a piece of music (Treisman, 1960) because two different input modalities (auditory and visual) are being used, as well as two different output modalities (speech and motor). However, it is very difficult even for skilled audio-typists to audio-type at the same time as shadowing, because both tasks involve the same input modality (auditory). The differences in the way concurrent stimuli affect dual task performance are important in the context of multimedia learning. When humans need to perform more than one task at a time, the human cognitive system often approaches, and might exceed, its limitation, preventing cognitive operations necessary for effective learning.

Research in varied psychological domains has found strong effects on human performance for dual tasks when the tasks have unlike modal forms. These investigations also indicate that a bottleneck in cognitive capacity might exist at various stages of cognition. Significant time costs of intra-modal tasks compared to inter-modal tasks have been found when matching and identifying polygons (Ittyerah, 1983). These findings show interference that depends upon the relationship among tasks, and demonstrate interference manifested in the slowing of real time comparisons. Phonological and visuo-spatial suppression tasks have been found to differentially affect performance depending on different concurrent arithmetic operations (Lee & Kang, 2002). Different arithmetic operation are theorized to use different cognitive resources and these results show interference that depends upon the relationship between the concurrent tasks, and that such interference can affect performance of singular mental operations. Output as a function of input from different modalities has been shown to be slower than output that modally matches input

(Wickens & Lui, 1988), another example of when concurrent tasks of different modalities affect performance, in this case during output.

Each of the preceding examples shows dual task interference that depends upon the relationship between the tasks and the cognition involved, and in combination support the contention that the effects of dual mode presentation occur at various stages of human information processing. They do not seem to indicate a single bottleneck or pattern of effects in cross-modal dual task performance. Encoding and understanding new information during multimedia learning is a cognitively demanding activity, and performance is likely to show a similar range of patterns (i.e., somehow determined by the relationship between concurrent events) to these cross modal dual task examples. The observed patterns are likely to be a function of interference or information loss that could take place during either or both the perceptual and cognitive stages of learning. Therefore, understanding the effects of dual mode presentation must encompass an understanding of how modality affects cognitive operations on both the *physical* and *mental* representations. Likewise, an understanding of the effect of multimodal instructional materials must include an understanding of how diagrammatic and descriptive information affects cognitive operations of both the *physical* and *mental* representations. However, understanding the differences between diagrammatic and descriptive information is not the focus of this project; the focus of this project is to discover how information can be presented to optimize (or reduce suboptimal effects of) cognitive operations performed on the mental representational form of information.

1.10 Individual Differences

To examine the underlying causes of human behavior, and the accuracy of the theories that shape this investigation, an accurate assessment of the internal (to individual observers) factors affecting behavior must be included. It is rare that a large group of untrained participants presented with the same instructions and task will behave in the same way. Therefore, it is important for this investigation to consider individual differences in the way that people use mental representations for comprehension (i.e., *comprehension strategy*) when presented multimedia materials. MacLeod, Hunt and Matthews (1978) found the Sentence Picture Verification (SPV) task to be a suitable method for assessing comprehension strategy.

Using the SPV task for the assessment of verbal abilities, Carpenter and Just (1975) presented a model (the C&J model) of how people form representations of both linguistic and non-linguistic stimuli when making a comparison between the two. They suggested that people represent a sentence with a logical proposition that is a function of the surface structure of that sentence, and that people similarly represent the picture with a simple proposition. Comparing the sentence to the picture involves mentally scanning one representation in reference to the

Table 1: The sentence-picture pairs as a function of trial type.

Trial Type	Sentence	Picture
True Affirmative (TA)	STAR IS ABOVE PLUS PLUS IS BELOW STAR	* +
False Affirmative (FA)	PLUS IS ABOVE STAR STAR IS BELOW PLUS	+ *
True Negative (TN)	PLUS IS NOT ABOVE STAR STAR IS NOT BELOW PLUS	* +
False Negative (FN)	STAR IS NOT ABOVE PLUS PLUS IS NOT BELOW STAR	+ *

other. The theory that people use a linguistic strategy is supported if response times reflect the need to make a predictable number of mental comparisons, and facility in making these comparisons is measured on the difference in time it takes to make additional comparisons (Carpenter & Just, 1975).

In the SPV task, the participant observes a simple sentence, such as PLUS IS ABOVE STAR, and then a picture with a star (*) above or below a plus (+). The task is to indicate as rapidly as possible whether the sentence and picture match. The independent variable is the number of mental comparisons necessary to determine whether the sentence and picture match. This is manipulated in two ways, by increasing the complexity of the sentence, or by having the picture and sentence match or not match. See Table 1 for examples and labels.

MacLeod et al. (1978) explored the use of the C&J model for assessing language skills. They used a pool of untrained participants that were measured with an array of intelligences tests. Their analysis led to the conclusion that testing the fit of each individual to this model produces grouping that reflects preference for comprehension strategy, not the assessment of language skills. It was shown that considering a general approach to a task by people is too simplistic, and that different people can approach the same task in different ways. This suggests that the SPV task should not be used to measure differences in language skills because, in contrast to Carpenter and Just's assumption that all individuals approach the task in the same manner, it appears that other non-linguistic strategies exist. In addition, more recent research has argued that the types of simple negated sentences used in picture verification studies are more ambiguous than their positive counterparts (Glenberg, Robertson, Jansen & Johnson-Glenberg, 1999). It has been suggested that this ambiguity could be what causes the increase in processing time. Comprehension strategies of individuals reflect differences in the way people represent information, and MacLeod et al. (1978) suggested that they could use the SPV task to determine how people compare information.

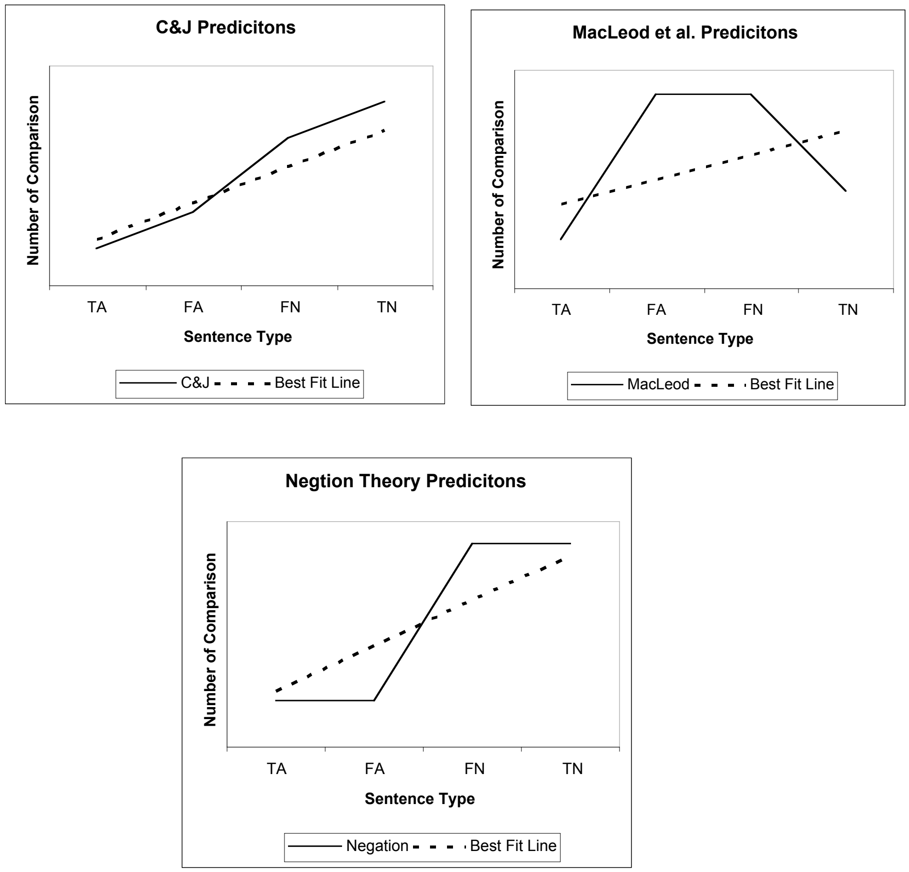


Figure 1: A comparison of example reaction time predictions based on sentence type made by the C&J model (1978), MacLeod et al. alternative model (1978), and negation phenomena theory (Glenberg et al, 1999).

According to the C&J model, the comparison types involved in the SPV task will yield a predictable pattern of performance (Carpenter & Just, 1975). However, according to MacLeod et al. (1978), participants who do not follow the C&J strategy might follow a different pattern of performance. Moreover, considering the effects of negation (Glenberg et al., 1999), response times might follow yet another pattern. Figure 1 shows examples of response patterns that might be predicted by the C&J model (Carpenter & Just, 1975), the negation phenomena (Glenberg et al., 1999), and an alternative proposed by MacLeod et al. (1978).

MacLeod et al. (1978) divided participants into groups of good, medium and poor fit to the C&J model based on the magnitude of the correlation between an individual's data and the predictions provided by the model. They suggested that those individuals who poorly fit the model use a visual-spatial comprehension strategy. Strategies were found to correlate with verbal and spatial skills and influence people's performance on cognitive tasks.

1.11 Choosing the optimal medium

When choosing between multimedia instructional ingredients, both physical and mental properties of the materials under consideration must be recognized. Though there might be individual differences in the way people represent information, the strength of the effect does not appear to be great (Landauer, 1995). Text and pictures should be logically chosen by, and tailored to, their respective strengths (Wickens & Hollands, 2000). Some past psychological and educational research provide insights into characteristics of a variety of multimedia ingredients that might influence the decision of what ingredients to include (e.g. Ainsworth, 1999; Larkin & Simon, 1987; Mayer, 2001, Sweller, 1999) However, the present understanding is far from complete. For instance, "hardly anything is known about the use of audio in instruction" (Aarntzen, 1993, p 354). Suggestions such as "If verbal information is lengthy, it should be visual (text) rather than auditory (speech), because of the greater permanence of visual information and the higher working memory demands of understanding speech" have been posited (Wickens & Hollands, 2000, p. 220). Such ideas are probably based on good intuition and a sample of research, but may intuitively contradict other research (e.g. that some verbal (word lists) material are better remembered if presented auditorily than if presented visually (De Haan, Appels, Aleman & Postma, 2000; Najjar, 1998)). The goal of this research was to begin an empirical investigation into how considering modality in two stages, physical and mental, can help improve predictions about the effectiveness of multimedia learning materials, specifically audio-verbal descriptive materials.

Chapter II

HYPOTHESES

For multimedia learning environments, presentation (physical) modality is a salient attribute of the information design, and can be changed with relative ease. Mayer (2001), following from Baddeley (1994), distinguished between visual/pictorial and auditory/verbal channels of information, but stated that "additional research is needed to clarify the nature of the differences between the two channels" (p.48). While a consideration of the nature and organization of working memory components is beyond the scope of this project, I investigated whether a multi-component working memory system (assumed to have an auditory and visual component) would be shown to contribute to learning in a multimedia environment. I hypothesized that the ability to commit and retain internal representations in working memory and ensuing cognitive operations (using such representations to manipulate, relate and store information for later use) influence this effect. It is proposed that a multi-component memory system contributes to the learners' ability to manipulate and store multimedia information efficiently. These effects would take place during some stage of information processing following perceptual encoding.

For this study I took the view that the *modality* of a bit of information may refer to its perceptual form or internal representation. If information is said to have *visual modality*, it contains properties of a visual or spatial form and corresponds to non-verbal, visual and diagrammatic representations as described above; if information is said to have *auditory modality*, it contains properties of an auditory or temporal form and corresponds to verbal, auditory and descriptive representations as described above. It must be noted that the physical presentation of information does not unequivocally determine its internal representation (mental modal form). Hence, an effective design strategy would derive from a classification of information according to both physical and mental representations. It is clearly important to understand how the modality of information affects learning at capacity limited stages of human information processing.

A framework for understanding how mental and physical properties of information relate to each other could be conceptualized by an understanding of how the information dualities described in this paper are related to each other. The dualities describe information at varying stages and groups of stages throughout the information-processing model. The experiment considers these dualities in an attempt to go beyond split attention effects associated with manipulating the perceptual mode of information and addresses modality effects associated with the processing of internal representations of information. Early and late stages of working memory activities were examined independently. This was accomplished by presenting the diagrammatic and descriptive information separately (via an *asynchronous* presentation). Presenting information separately has been identified as a bad idea for multimedia design (See Mayer, 2001), but was used in this investigation for the purpose of forcing learners to represent and store presented information in working memory for a short time. In this way, I was able to isolate effects that occur in later stages of working memory (after the presentation event) from those that occur in earlier stages (during the presentation event).

The experiment focused on separating the external modal form (perceptual form) and internal modal form (which might be related to information content) of information presented to be learned. I compared dual (multimedia) and single (only visual) mode presentation performance within asynchronous and synchronous conditions. Physical-perceptual form (i.e., visual text versus auditory narration) was manipulated while verbal-descriptive content information remained the same. I did not change the content of the information in an attempt to preserve the mental form of representation. Additionally, I examined the effects of group differences involving representational strategies, and whether they effected learning from these presentation combinations. Inefficiencies in human information processing during extended storage and cognitive operations related to learning (relatively later stages of cognitive operations) were hypothesized to cause the effects predicted. I hope to draw conclusions, some general and some specific, about what this means for designers who wish to optimize cognitive benefits of multimedia instructional materials.

2.1 Hypothesis 1

Synchronous presentation of content will produce better learning than asynchronous presentation of equivalent content, regardless of presentation modality.

If a learner must compare and integrate multiple streams of information, and some information is lost due to the incomplete nature of working memory representations, the learner will benefit from the availability of sensory information while attempting to perform cognitive operations important for learning. The loss of information over time between sensory and working memory representations of information reduces the likelihood that the learner will have the information necessary for producing complete memories. However, the ability to commit the first stream of information to long term memory would ameliorate this effect. Therefore, the ability of the learner to commit new information should be controlled when comparing synchronous and asynchronous presentation. This ability is controlled in all analyses using participants' O-Span score as a covariate.

I suggest that the decay of information in working memory produces the contiguity effect, which has shown that presenting descriptive and diagrammatic information separately is worse than presenting them together (Mayer, 2001). My explanation differs from that of Mayer (2001) because I believe that the benefits are a result of more stable (or replenished) internal representations of the information to be learned. The effect is not solely attributable to the (external) physical or temporal proximity of information.

2.2 Hypothesis 2

Learning from long verbal content is not as good when presented as narration (compared to text) because long auditory descriptions strain the limited capacity of working memory.

The relative permanence of visual information (compared to auditory information) and the higher working memory demands associated with rehearsing speech for maintenance are likely to be the contributing factors (Wickens & Hollands, 2000). Even though research has shown some verbal content to be better remembered after being presented auditorily (Najjar, 1998), the more complicated cognitive operations involved with manipulating and relating elements of speech will add to the demands of storing and processing the information. This will increase the likelihood that working memory capacity is exceeded, and lead to the selection of a subset of the information for processing. Moreover, learning from lengthy narratives will require more cognitive effort on the part of the participants.

2.3 Hypothesis 3

Synchronous presentation of equivalent content will produce better learning when presented with a mixture of modalities than with one; that is when auditory-descriptive material (as opposed to visual-descriptive material) accompanies visual-diagrammatic material.

Past research has emphasized the effects of mixing the mode of presentation modality on learners' ability to encode new information (e.g. Mayer, 2001; Sweller, 1999). In accordance with previous research, data supporting this hypothesis would show improved learning from synchronous, multimodal (perceptual) instructional materials. It would not unequivocally show why. The goal of this investigation is to determine whether effects of multimedia presentation might occur after the learner forms internal representations. In the past, educational psychologists have assumed that the bottleneck that leads to the performance decrement for unimodal instructional materials compared to multimodal instructional materials occurs around the transfer of sensory information to working memory.

However, the poorly constructed multimedia presentation used in this experiment might not show a multimedia effect. Multimedia presentations including long narrations might be relatively inefficient, that is, equally efficient as unimodal presentations. A failure to support Hypothesis 3 (even though the power of the experiment is sufficient to detect other effects) indicates that perceptual effect explanations are incomplete. If a multimedia effect is not found, it is likely to be the result of the cognitive costs of lengthy auditory-verbal material countering the cognitive benefits of spreading information among sensory inputs. The goal of this investigation is to determine whether an influence on learning from multimedia presentation occurs after perception. In light of this and the rest of the hypotheses, I hope to show that multimedia effects cannot be completely explained by perceptual effects. Realizing this will be the first step in developing a more complete explanation of multimedia effects.

2.4 Hypothesis 4

Eliminating the effects of multimodal presentation on information loss due to perceptual effects will not eliminate the total effects of multimodal presentation on learning. That is, the relationship between presenting information asynchronously or synchronously is the same for unimodal and multimodal information.

If this hypothesis is supported, it will suggest that at least some of the effects of multimedia information presentation occur after sensation. Presenting less sensory information at one time will reduce the amount of information awaiting transfer to the working memory system. Hence, asynchronous presentation will reduce information loss due to perceptual effects, i.e. before transfer to the working memory system. The next few sets of hypotheses will attempt to show that multimedia effects might be dependent upon internal representations, and not solely external presentation.

2.5 Hypothesis 5

Differences in performance based on the order of information type will depend on physical-modal combinations.

The order of content (i.e. descriptive content before diagrammatic content, or vice-versa) might have an effect on learners' ability to integrate internal representations and certain types of information might be more easily processed in certain internal modal forms. However, properties of internal representations of information can also create interference. According to my view of modality (based on that of Baddeley (1994), Pavio, (1986), and Wickens (2002)) , the optimal storage and active processing of descriptive (verbal) material takes place in verbal/auditory component of working memory.

2.5.1 Hypothesis 5a

Learning from instructional material presenting visual text (descriptive) information first is not significantly different from learning from instructional material that presents auditory narration (also descriptive) first.

Descriptive material presented auditorily would require fewer cognitive operations for transfer to this store than visually presented descriptive material. However, in asynchronous presentation, capacity limits are less likely to be exceeded, and learning effects are less likely to be observed. The visual perceptual presentation of the descriptive material is converted to an auditory internal representation that behaves the same as the auditory perceptual presentation before a modality effect occurs.

2.5.2 Hypothesis 5b

Learning from instructional material presenting visual text (descriptive) information last is worse than instructional material that presents auditory narration (also descriptive) last.

Visual descriptive material would engage processes required to convert visual descriptive information to a verbal/auditory form and reduce available cognitive capacity in one or both stores necessary for forming robust memories. Learners might experience more interference (overlap between internal and external representations causing cognitive overload that leads to selection and impedes cognitive operations necessary for learning) when attempting to convert a perceptual representation to an internal representation when presented with text than with narration. This effect will result in amplified behavioral outcomes when they are actively storing a visual-spatial internal representation of the diagram in visual memory. The diagram would create more interference with the visual presentation of the description than the auditory presentation of the description.

2.5.3 Hypothesis 5c

Learning from instructional material presenting text first is better than instructional material that presents text last.

Combining Hypotheses 4a and 4b yields Hypothesis 4c. The hypothesized cause is that learners transfer the text to the auditory store and free capacity for operation on the diagram. However, when the order is reversed, the diagram retains a spatial representation that interferes with processing the visual text. Hypotheses 4a – 4c involve modality effects that occur when perceptual effects are reduced or eliminated. The presence of such modality effects might demonstrate that factors at a level deeper than perception affect the benefits derived from the different combinations of information presentation.

2.6 Hypothesis 6

If grouping individuals based on differences in comprehension strategy (the use of internal representations) influence patterns of learning, these differences might be attributed to the formation and storage of internal representations. If this can be shown to be true, then internal representations, in addition to perception, should be considered when designing multimedia materials.

I will use a methodology similar to that of MacLeod et al. (1978) for using the SPV task to assess comprehension strategy. Taking the methods and limitations of this study into consideration, I will use a similar method to group participants, and make a similar, but more conservative claim that is strong enough to support my hypotheses. That is, I will group individuals into those who use a linguistic strategy, those who do not, and those who use neither or a mixture of strategies. I will then determine whether the linguistic and non-linguistic groups experience equal learning effects based the type of information presentation. I will examine the effects of comprehension preference on interference among internal representations of information.

In the past, educational psychologists have assumed that the bottleneck that leads to the performance decrement for unimodal instructional materials compared to multimodal instructional materials occurs around the transfer of sensory information to working memory. I believe that effects of multimedia presentation occur after this point. Showing that there is a difference in learning patterns for those who use a linguistic strategy and those that use a non-linguistic strategy would show that the way people prefer to comprehend linguistic information influences the presence of multimedia effects. In turn, it is concluded that internal representations influence the multimedia effect.

Chapter III

METHOD

3.1 Participants

The participants were 120 undergraduate students at Georgia Institute of Technology who volunteered to participate for course credit. An analysis of previous research in the domain of multimedia learning was conducted to help arrive at this number. Effects sizes were obtained from the work of Mayer (2001), and from the work of Sweller and colleagues (Mousavi, Lowe, & Sweller, 1995). Mayer (2001) reported extremely robust effect sizes measuring recall in a set of experiments using dual mode instruction, ranging from .48 to 1. Since Mayer's materials were used in other learning experiments and known to be very robust, and in the interest of being conservative, the effect size of .48 (From Mayer and Moreno, 1998, experiment 2) were considered here.

Experiment 4 of Mousavi et al., (1995) was conducted to reaffirm the superiority of mixed auditory and visual presentation over single mode presentation due to expanded working memory capacity, adding the constraint of holding study time constant across learning material types. Time spent answering 'test' questions (those most similar to, but not the same as, the examples used during instruction) showed a significant effect, $F(1,36) = 7.31$, $MSE = 2{,}956.58$, and an effect size of .169. This particular experiment was chosen for its similarity in question and dependent measure to the experiment planned here.

Using the G-Power software to compute the necessary N to detect a significant difference in a special comparison of two of six manipulations (previous work has compared only dual mode and single mode presentation, corresponding to two of the six manipulations planned here), the parameters used were and alpha = .05, power = .95, groups = 6 (because the planned experiment has six within subjects manipulations), numerator df = 1 (because the approximated effect sizes are based on one comparison within these six). An eta of .372 was used based on computing the average of the effect sizes in the selected work of Mayer and Sweller. The analysis yielded a recommendation of 96 participants. So that 20 people would see each latin square set of materials, and to increase power because the materials to be used in this experiment are different than the materials used in previous work, I rounded the number of participants to 120.

3.2 Materials

3.2.1 Pre-measures.

Prior to the experimental manipulation, measures of working memory capacity and individual differences in information processing strategy were assessed. Each were measured using tests that have been systematically checked for validity and reliability during other research programs. Both were administered with E-Prime and took 10 to 15 minutes each.

Working memory capacity was measured using Automated O-Span. Automated Ospan is an automated version of a popular working memory capacity task (operation span). This task has been shown to be a reliable and valid indicator of working memory capacity (Unsworth, Heitz, Schrock & Engle, in press). In this task participants are asked to remember letter strings while performing simple arithmetic. The two tasks (arithmetic and letters to be remembered) are presented separately to participants, in an alternating manner. Participants indicate whether a given number is or is not the answer to an arithmetic problem throughout each trial and indicate recall of letter strings by mouse clicking to indicate the letter and order they were presented at the end of the trial. Dependent measures include an absolute span score shown to correlate with other measures of WM. This score will be used to control for working memory capacity in the final analysis.

Individual differences in information processing strategy were measured using the Sentence-Picture Verification described above (SPV) (Macleod et al., 1978). During the SPV task, participants completed 16 practice comparisons and 64 experimental comparisons. During the test, the sentences and picture combinations were manipulated to yield four comparison types as outlined in Table 1. Response patterns were analyzed to group participants into those that follow the Carpenter and Just Model (1975), those who follow another pattern, and those who follow a mixture of patterns (see Results section for more detail). Grouping reflects preference for compression strategy and has also been shown to be related to psychometric measures of verbal and spatial cognitive abilities. Participants were grouped according to their strategies to see if this has an effect on learning from different forms of presentation.

3.2.2 Learning Materials.

The short lessons used to create experimental multimedia lessons were adapted from David McCaulay's (2003) "The New The Way Things Work." This book describes in simple terms the way that everyday objects function. Each explanation included labeled diagrams and explanatory text about a specific object. The topics included as lessons in this experiment were Window Shade, Stapler, Air Filter, Refrigerator, Hot Air Balloon, Metal Detector and an Electric Bell (See Appendix A for the pictures and text). The lessons were chosen to have similar information content in both diagrammatic and textual forms. Each lesson had similarly complex diagrams and labels with similar word length and passage divisions for their descriptive text.

All textual descriptions were recorded with the same female voice to make auditory narrations for trials containing auditory stimuli. Participants were able to listen to the narrations by starting a sound clip with a single mouse click over an on screen button. The button was labeled with the title of the text passage to be narrated, and an indication of how long the narration will last. Each lesson had 2 to 4 buttons depending upon the number of text passages that have been replaced. If a particular button was pressed, other narrations are stopped and the desired narration begins. The buttons were physically positioned in the same region of the screen as the text they replace. As a result of the similar textual length of the total descriptive content for each topic, the total narration length (sum of each included narration) for the lessons were similar (three had M = 53.5 sec; SD = 4 sec; three had M = 81.5 sec; SD = 6 sec).

Each lesson was scanned in color to be presented via an IBM PC using an LCD monitor. Separate presentation screens for each lesson contained only the diagram, only the text, only the buttons, both the text and the diagram, and the diagram with buttons (see Figures 2-5 for examples). Physical positioning within the screen was preserved for descriptive and diagrammatic information in all presentation forms. The slides were combined according to the experimental procedure using Flash. A counter was included to indicate time remaining for the current lesson (see Figure 5). These visual stimuli were presented via Internet Explorer using the full screen setting with no toolbars or menus visible to the participants. Participants listened to the auditory stimulus via Koss R/65B Stereo Headphones.

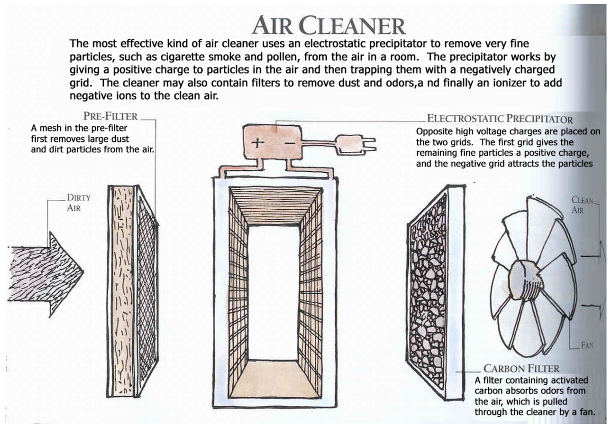


Figure 2: Simultaneous Diagram and Text for the Air Cleaner lesson.

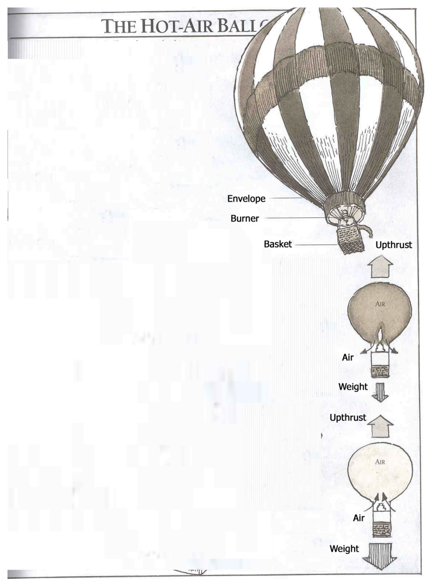


Figure 3: Diagram only for the Hot Air Balloon lesson.

Stapler: A stapler is an everyday device that conceals an ingenious arrangement of springs. It uses both a coil spring and a leaf spring, which feed the staples along the magazine and return the stapler to its original position once it has been used. Pushing down the stapler causes the blade to descend into the magazine, forcing the front staple through the papers. The anvil bends the ends of the staple to clip the papers together. The return spring then raises the magazine and blade, allowing the magazine spring to advance the next staple into position.

A projection on the base plate flattens the return spring when the stapler is used. The spring raises the magazine away from the anvil after use.

The return spring is a leaf spring that raises the blade from the magazine and moves the magazine and anvil apart after use.

A strip of staples is fed into the magazine of the stapler and held there by a coil spring that advances the next staple into position.

Figure 4: Text only for the Stapler lesson.

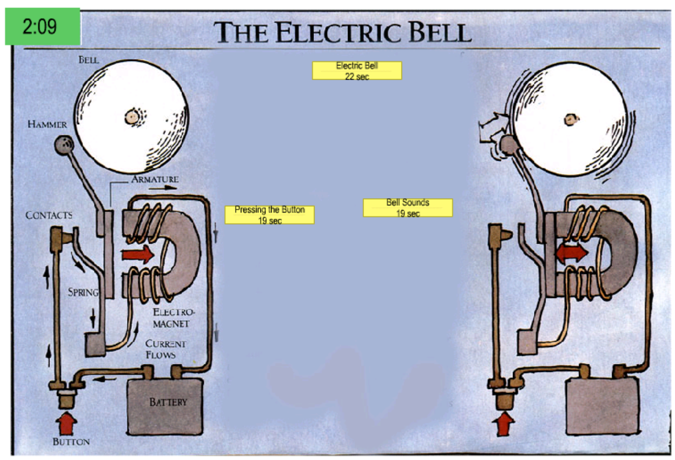


Figure 5. Diagram with narration buttons for the Electric Bell lesson. The counter is displayed in the upper left corner.

3.2.3 Test Materials

Learning was measured with a paper and pencil test. At the start of the experiment, participants were given a packet that contains ten recall questions for each lesson. After each lesson, participants were asked to answer ten selected content questions based on the lesson material. Test questions were developed to assess learning from all parts of the learning material. A majority of the test content was drawn from the descriptive material. The questions from the descriptive material were balanced across the descriptive passages. Additionally, some test questions were developed to be related to diagrammatic material. The number of questions regarding the diagrammatic material approximates that of each descriptive passage. These questions were simple fill-in-the blank statements that could be answered according to the lesson material and were tested in piloting. Selected questions met the criteria of being answered correctly 85% or more with 'open book' materials (that is, participants were able to view the text and diagram while answering the questions) and answered correctly less than 30% of the time with no materials present. Participants who answered less than 85% of all questions correctly while having access to open book materials were treated as outliers and were not considered in piloting analysis. Appendix B contains a legend showing each test question and the part of the content addressed by each question.

Participants also completed a NASA TLX (Hart & Staveland, 1988) regarding the previous lesson. The TLX items subjectively measure such constructs as cognitive load, physical load, and frustration. Summed together, the 5 items yield an overall Workload score. This subjective rating of workload may be informative in situations where increased cognitive load affects participants' ability to learn.

3.3 Procedure

Upon arrival to the lab, participants were briefed in a small group (up to four participants at a time) on the purpose and procedure of the experiment, and given a demonstration of how to use the interface. Subjects were instructed to view the lessons about "The Way Things Work" and to attempt to answer the following questions about the lesson they just viewed. They then each sat down at an individual workstation where they were assessed for both working memory capacity and individual differences in information processing strategy. Each participant completed the O-Span task followed by the Sentence-Picture Verification task.

After completing these measures, participants put on the headphones to listen to the narration of the on-screen lessons. They were given a paper and pencil packet that contains the general demographics form, and the questions and TLX forms for each of the seven lessons. The initial screen contained instructions similar to those that guide them through the rest of the experiment. These instructions directed them to fill out a one page demographics and background knowledge questionnaire, and then click the mouse in the indicated region in the center of the screen to begin their first lesson. This first lesson was a training lesson, though the participants were unaware of this. The purpose of the training lesson was to familiarize participants with the interface and ensure that the equipment is working properly. After the training lesson, the screen directed them to answer content questions and TLX regarding the preceding lesson. To begin the next lesson they were again instructed to click the mouse in the indicated region in the center of the screen. Each lesson was followed by the same instructions to answer content questions and the TLX item regarding the preceding lesson.

Each participant received seven lessons, one training lesson and six experimental lessons. Each of the six experimental lessons had a visual diagram accompanied by either text (visual input, hence single mode) or narration (audio input, hence dual mode). For each combination (i.e., diagram and text vs. diagram and narration), participants received three variations on the presentation order (form of lesson), creating six types of trials:

1. diagram then text (D+T)
2. text then diagram (T+D)

3. text and diagram simultaneously (T&D)

4. diagram then narration (D+N)

5. narration then diagram (N+D)

6. narration and diagram simultaneously (N&D)

Each participant received the same order of lesson content (Microscope, Stapler, Air Cleaner, Hot Air Balloon, Electric Bell, Window Shade, and Metal Detector) with the order of the form of lesson varied by Latin Square. This was done to eliminate order effects for form of lesson, but will hold the order of the lesson content constant across participants. Each lesson lasted three minutes during which participants can attend any available information they chose. For trials D+T, T+D, D+N, N+D, participants are able to view or hear the first set of information for 90 seconds, and the second type of information for 90 seconds immediately following. For trials T&D and N&D, all the information is available for all three minutes. When participants are finished, the screen directed them to call the experimenter. The experimenter debriefed the participants and escorted them out of the lab.

Chapter IV

RESULTS AND DISCUSSION

For the analysis of the results, 12 participants were not used. The unused participants were removed from the analysis due to data coding errors. For these participants, at least one of O-Span, SPV or a single learning score was inaccurately recorded. The number of participants used, 108, still exceeded the minimum number determined by the power analysis, 96.

Assessing Working Memory Capacity. The overall O-Span Score was computed and recorded for each individual. The scores ranged from 33 to 72, $M = 62.91$, $SD = 9.19$. Working memory ability, as measured by the O-Span task was treated as a covariate in all analyses. O-Span significantly correlated to learning performance, $r = .139$, $p < .01$, and therefore met the necessary assumption of a significant covariate. It is likely that the magnitude of the correlation would have been larger if the participants exhibited a wider range of ability on the O-Span task. There was a clear restriction of range in the present sample, compared to the sample used in Unsworth, et al., (in press) which had $M = 39.16$, $SD = 17.41$. It was also confirmed that there was not a violation of the homogeneity of regression slopes (i.e. there is no interaction between experimental groups and O-Span score).

4.1 Assessing Individual Differences in Comprehension Strategy.

Before analysis, response times were screened for efficacy. Participants had to respond correctly on 52 of 64 experimental SPV trials. Reaction times more than two standard deviations from the cell-means for each participant were discarded. The remaining data were used to classify people as good, medium and poor fits to the C&J model. The approach that MacLeod et al. (1978) took to grouping individuals centered on their correlation to the C&J model predictions. Carpenter and Just's (1975) predicted number of comparisons was used as the independent variable and time to response was used as the dependent variable. The fault that I see with the approach of MacLeod et al. (1978) is that other strategies or factors (e.g. those illustrated in Figure 1) might also yield a linear correlation, but not reflect a linguistic strategy.

In an attempt to more completely examine participants' pattern of responses, a univariate ANOVA with linear, quadratic, and cubic contrasts was performed for each individual. Carpenter and Just's (1975) predicted number of comparisons was used as the factor; time to respond was used as the dependent variable. To asses the degree to which performance followed the C&J model and was not influenced by other psychological phenomena, the amount of variance accounted for by the linear contrast was compared to that of the other two contrasts. I also computed a correlation of an individual's data to the C&J model in the same way Macleod et al. (1978) would have, for two reasons. First, I grouped the participants in a way that I believe MacLeod et al. (1978) would have for comparison. Second, a negative correlation between the response pattern and the model clearly reflects a poor fit, and was a criterion for grouping.
The comparison was a ratio of the variance associated with the linear contrasts (SSLinear) and the other contrasts (SSOther). A ratio was used because it is a straightforward way to conceptualize the degree to which performance was influenced by each model, and because it is important to distinguish between people whose behavior primarily followed the linear model. A variant of the Scree Test (Kim, 1978) was used to determine the cutoffs: Ratios greater than 2.5 were labeled good fit; ratios between 1 and 2.5 were labeled medium or mixed fit. This means that in cases where the variance accounted for by a linear model (SSLinear) is more than 2.5

times greater than the variance accounted for by the quadratic and cubic models together (SSOther), it can be said that the individual follows the C&J model well.

Table 2: Relationship between the two Individual Differences testing models.

Present Groups: **MacLeod et al. Groups:**	**Poor**	**Medium**	**Good**
Poor	29	1	1
Medium	14	14	5
Good	1	7	53

Following analysis, the examination of individuals' ratios yielded three groups that reflect the likelihood that they use a linguistic strategy in the SPV task. If the data for an individual follow the C&J model well, it can be presumed that the individual predominately uses a linguistic strategy similar to the one theorized by Carpenter and Just (1975). The dominant effect for people with a medium fit is unclear. For those with a poor fit (ratio < 1), it is likely that a negation or pictorial strategy dominates their behavior. MacLeod et al. (1978) grouped their participants into three groups with similar meaning. Table 2 shows a contingency table illustrating the relationship in findings for my method and the method used by MacLeod et al. (as estimated by the comparison analysis).

For most individuals (96 of 125), both the method I used and the MacLeod et al. comparison analysis method yielded the same grouping. However, the proposed method found more individuals to be poor fits to the C&J model (44 compared to 31). Though not intended by the analysis, I find this increase to be advantageous for using the SPV task as a grouping measure. The ambiguity of the strategy used by people in the medium fit group makes it difficult to predict their behavior. Therefore, categorizing people in the middle group is less informative than

Table 3: Examples of the differences in grouping by the present method compared to that of MacLeod et al. (1978).

Participant #	Correlation	MacLeod et al. (1978) Comparison Group	SSLinear/ SSOther (Ratio)	Present Group
343	.510	Good fit	1547149 / 360973648	Poor fit
327	.156	Poor fit	591502/ 32092	Good fit
232	.201	Medium fit	187549/ 68879	Good fit
306	.203	Medium fit	305210/ 107452	Good fit
304	.212	Medium fit	330650/ 52980	Good fit
300	.227	Medium fit	290338/ 73725	Good fit
415	.310	Medium fit	1292996/ 6558	Good fit
272	.386	Good fit	7146077/ 6163030	Medium fit
421	.392	Good fit	1397108/ 929405	Medium fit
245	.597	Good fit	10936553/ 5771109	Medium fit
224	.179	Medium fit	35148/ 718779	Poor fit
341	.234	Medium fit	1732040/ 3976594	Poor fit
249	.247	Medium fit	1908299/ 6171616	Poor fit
226	.271	Medium fit	736342/ 1359372	Poor fit

categorizing them in good or poor fit groups.

To examine the cases where the two models differed in grouping an individual, Table 3 is provided. I believe that the results from participants 343 and 327 make the case for the proposed method. Participant 343 had a very high correlation with the model. However, using the ANOVA contrasts, I can see that the data do not have a linear correlation with the C&J model; it might not be straight as predicted in the C&J model, but increases (as shown in Figure 1). Participant 327 had a low correlation to the C&J model. However, again examining the ANOVA contrasts, we can see that the data are linear. It is probable that this individual experienced the linear pattern predicted by Carpenter and Just (1975), but the slope was shallow. The C&J model does not make predictions about the magnitude of increase in response times, so even a small slope would theoretically follow their model, but not the MacLeod et al. (1978) interpretation of it.

Examining the patterns of learning for the participants who follow the C&J model (primarily use a linguistic comprehension strategy) and those participants who do not follow the C&J model (do not primarily use a linguistic strategy) shed light on the effects of individual differences and support the general claim that multimedia effects are influenced by internal representations.

4.2 Assessing Workload.

Subjective workload was assessed using the NASA TLX rating scale. The data was gathered as a means to further examine the results of the experimental manipulations. On only a few occasions were there interesting differences in workload ratings. On these occasions (Hypotheses 2 and 4), the differences in workload were helpful in examining and explaining the experimental results. Results involving subject workload ratings are included in the next section.

4.3 Assessing Learning.

Analyses of learning performance (dependent measure) on different experimental trials (independent measures) were conducted with one-way between groups analysis of covariance. O-Span was used as a covariate in all analyses.

Contrasting trials by unimodal vs. multimodal trial, asynchronous vs. synchronous presentation, and by order of information type (descriptive or diagrammatic presented first) allow examination of patterns in the data. Before analysis, outliers were deleted. The removed outliers were those whose total raw score (out of 10) for one lesson (topic) in one presentation condition (presentation type) were more than 2 standard deviations from the mean. Twelve test scores were removed by this method. Table 4 describes the mean, variance and number of observations for each lesson and each experimental condition. Raw Means, and Estimated Marginal Means using O-Span as a covariate, are presented.

The lessons and tests were developed for this experiment. To asses the efficacy of the data gathered, each test was assessed for internal consistency reliability using Cronbach's alpha. This measures how well a set of items measures a single unidimensional latent construct (i.e., learning from the previous lesson). The tests had reliabilities ranging from .555 to .709 (see Table 5). A reliability of 0.7 is commonly regarded as an acceptable reliability coefficient (Nunnaly, 1978), but lower thresholds are sometimes used in the literature. The reliabilities are such that it is not necessary to be overly concerned about the results of the present study, but future studies using these materials should include an improvement of the reliability of its tests.

Hypothesis 1: Synchronous presentation of content produced better learning (*Estimated Marginal Mean (EMM)* = 5.77) than asynchronous presentation (*EMM* = 5.39) of equivalent content, including both presentation modalities, $F(1, 618) = 4.34$, $p < .05$, $MSE = 4.53$. This result supports the contiguity effect, which uses a perceptual basis to explain why presenting descriptive and diagrammatic information together is better than presenting them separately(Mayer, 2001). However, I suggest that the contiguity effect is a result of cognitive benefits that arise at a stage in the information processing model later than perception or

Table 4: Means and Variance of the data collected in each experimental condition and lesson. Number of participants in each cell (N), raw score (Raw Mean) and Standard Deviation (SD) on the ten question quiz, and ANCOVA adjusted scores (Estimated Marginal Mean) with associated Standard Error (SE) are shown.

Lesson: / Condition:		Metal Detector	Window Shade	Electric Bell	Hot Air Balloon	Air Conditioner	Stapler	All
N+D	*N*	18	17	17	17	17	18	104
	Raw Mean (SD)	4.56(1.95)	3.06(1.98)	3.53(1.51)	6.24(1.52)	6.82(0.95)	5.28(2.63)	4.91(2.25)
	Estimated Marginal Mean (SE)	**4.50(0.45)**	**3.12(0.46)**	**3.58(0.46)**	**6.13(0.46)**	**6.79(0.47)**	**5.53(0.46)**	**4.92(0.20)**
N&D	*N*	17	18	18	18	18	18	107
	Raw Mean (SD)	4.35(1.77)	5.06(2.36)	5.39(1.50)	6.00(2.11)	7.39(1.65)	6.44(2.04)	5.79(2.12)
	Estimated Marginal Mean (SE)	**4.41(0.46)**	**5.00(0.45)**	**5.36(0.46)**	**6.05(0.45)**	**7.53(0.46)**	**6.35(0.45)**	**5.78(0.22)**
D+N	*N*	17	17	18	18	17	18	105
	Raw Mean (SD)	4.06(2.05)	4.88(1.17)	5.17(1.43)	6.22(1.69)	5.88(1.93)	5.56(2.28)	5.30(1.88)
	Estimated Marginal Mean (SE)	**4.11(0.46)**	**4.79(0.46)**	**5.11(0.45)**	**6.30(0.46)**	**5.91(0.46)**	**5.36(0.46)**	**5.26(0.20)**
T+D	*N*	17	18	17	17	17	17	103
	Raw Mean (SD)	4.06(1.92)	5.39(1.75)	5.71(1.26)	4.18(1.51)	6.76(1.99)	7.18(2.56)	5.54(2.17)
	Estimated Marginal Mean (SE)	**4.10(0.47)**	**5.30(0.46)**	**5.60(0.46)**	**4.23(0.46)**	**6.81(0.46)**	**7.12(0.46)**	**5.54(0.21)**
T&N	*N*	17	18	18	18	17	18	105
	Raw Mean (SD)	4.94(1.85)	5.35(2.50)	4.50(2.23)	6.06(2.13)	7.94(1.14)	5.50(2.50)	5.70(2.34)
	Estimated Marginal Mean (SE)	**4.86(0.46)**	**5.42(0.47)**	**4.55(0.45)**	**6.30(0.46)**	**7.89(0.46)**	**5.55(0.45)**	**5.75(0.23)**
D+T	*N*	18	17	18	17	18	18	106
	Raw Mean (SD)	5.56(2.01)	4.47(1.42)	5.11(2.11)	6.35(1.80)	6.94(1.51)	6.44(2.48)	5.82(2.07)
	Estimated Marginal Mean (SE)	**5.60(0.46)**	**4.54(0.46)**	**5.11(0.46)**	**6.30(0.46)**	**6.85(0.45)**	**6.49(0.45)**	**5.83(0.21)**
All	*N*	104	104	106	105	104	107	630
	Raw Mean (SD)	4.60(1.95)	4.71(2.04)	4.91(1.82)	5.85(1.92)	6.96(1.66)	6.06(2.46)	5.51(2.16)
	Estimated Marginal Mean (SE)	**4.59(0.19)**	**4.70(0.19)**	**4.89(0.19)**	**5.88(0.19)**	**6.96(0.19)**	**6.07(0.19)**	**5.52(0.09)**

Table 5: Reliability for each test by lesson.

Test	**Reliability** (Chronbach's Alpha)
Stapler	.709
Electric Bell	.570
Air Filter	.589
Hot Air Balloon	.582
Metal Detector	.555
Window Shade	.617

encoding. This interpretation can be considered support for the argument that working memory processes influence the learners' ability to integrate new information. In this case, the working memory process that influences learning is the need for selection due to overloading capacity. Sustained external representations allow re-sampling of the information lost due to the capacity limitations of storing and operating on internal representations. The rest of the analyses from this experiment will help us see whether it is plausible to attribute this difference to working memory processes rather than perceptual contiguity.

Hypothesis 2: Learning from lengthy verbal content was not as good when presented as narration (*EMM* = 5.32), compared to text (*EMM* = 5.71), $F(1, 618) = 5.08$, $p < .05$, $MSE = 4.52$. In addition, subjective workload was higher for trials with narration ($M = 50.03$), compared to text ($M = 46.00$), $F(1, 627) = 4.13$, $p < .05$, $MSE = 622.18$. This supports the suggestion that long auditory descriptions impose a strain on the limited capacity of Working Memory (Wickens & Hollands, 2000). It appears that auditory narrations of the length included in this experiment strain working memory capacity. The phenomenon of higher workload occurs even when the narration is presented separately from other information that might influence the availability of working memory capacity (i.e., asynchronous trials), $F(1, 407) = 4.70$, $p < .05$, $MSE = 585.42$ (Asynchronous Narration $M = 52.79$; Asynchronous Text $M = 47.67$). Designers need to consider the effects that their chosen presentation, such as presenting lengthy verbal content via narration, will have upon learner's working memory processes.

Hypothesis 3: If the multimedia effect improves learners' ability to learn new information by presenting information with two modalities rather than one (Mayer, 2001; Mousavi et al., 1995), then simultaneous presentations including narrations should be better than those including text. However, no effect was found when comparing text (*EMM* = 5.75) versus narration (*EMM* = 5.78) conditions within the simultaneous presentations, $F(1, 208) = .01$, $p > .05$, $MSE = 4.85$. If even a portion of the multimedia effect can be explained by the ability to hold and manipulate information in two working memory stores after perception, then considering the fact that, for the presentation choices in this experiment, simultaneous presentation has cognitive benefits (Hypothesis 1), and long narrations have cognitive costs (Hypothesis 2), it might be expected that the two effects cancel each other out, producing a null multimedia effect. These results contradict previous research on two counts: multimedia presentation is shown not to be better than unimodal presentation, and the totality of multimedia effects are shown not to be attributable to manipulating presentation form. It appears that existing explanations for the multimedia effect on learning are inadequate, and the need for further investigation exists.

Hypothesis 4: If the relationship between presenting information asynchronously or synchronously is the same for unimodal and multimodal information, at least some of the effects of multimedia information presentation occur after sensation. To support this, the relationship between learning from information presented separately and together should be similar for both unimodal and multimodal information. The data did not support Hypothesis 4. For multimedia information combinations (narration and diagram), a comparison of separate (*EMM* = 5.09) and simultaneous (*EMM* = 5.78) trials yielded a significant difference, $F(1, 310) = 7.91$, $p < .05$, $MSE = 4.21$. For unimodal information combinations (text and diagram), a comparison of separate (*EMM* = 5.69) and simultaneous (*EMM* = 5.75) trials yielded no significant difference, $F(1, 310) = .06$, $p > .05$, $MSE = 4.77$.

It seems that eliminating perceptual effects (by presenting the information asynchronously) for unimodal conditions (text and diagram) does not influence learning. This would contradict the contiguity effect and demonstrate that the benefit of presenting information all at once is not influenced solely by perceptual effects involving temporal or physical contiguity. Such a demonstration reaffirms my stance that existing explanations of the multimedia effect on learning are incomplete. On the other hand, it is possible that the participants who received the narrations at the same time as the diagram were able to use the diagram as a way to segment the narrations, thereby reducing the cognitive costs associated with lengthy narrations. This is further supported by the fact that subjective workload was higher for multimedia (narration and text) trials using asynchronous (*M* = 52.79) compared to synchronous (*M* = 44.52) presentation, $F(1, 312) = 7.71$, $p < .05$, $MSE = 619.91$. This would suggest that the multimedia effects produced by relationships between the information that arises during internal and external representations, and their interaction, is something that designers of multimedia instructional materials should consider. In this case, providing learners with a way to reduce the length of the narrations was helpful to them. To further examine this effect, future research might divide the narrations and present the diagram between narration segments to test whether, and to what degree, perceptual effects are involved.

Results related to the remaining hypotheses (Hypotheses 5a, 5b, 5c, and 6) would provide converging evidence that shows that the effects of multimedia information design do not occur at the level of perception, but are results of effects that occur during a later cognitive process. These comparisons will involve trials that present the learning materials asynchronously. By presenting each stream in isolation, perceptual interference among information should be eliminated. Remaining differences observed are the result of effects at some stage of cognition that occurs after perception.

Hypothesis 5a: If learning from the text-then-diagram lesson (T+D) is equal to learning from the narration-then-diagram lesson (N+D), internal representations of the descriptive material (presented first and alone) used to compare and combine with the diagrammatic material (presented second, and alone) was the same regardless of physical presentation. However, judging from the fact that learners were more successful after a text-first presentation (*EMM* = 5.54) than a narration-first presentation (*EMM* = 4.92), $F(1, 203) = 4.11$, $p < .05$, *MSE* = 4.76, it appears that the internal representation that arose from the text presentation was better for learning. This is probably due to the fact that long narrations carry cognitive costs. Still, this supports the idea that the reason text is better for lengthy descriptive content is not influenced by synchronous perceptual events.

Hypothesis 5b: Similarly, learning from the diagram-then-text lesson (trial D+T, *EMM* = 5.83) was better than the diagram-then-narration lesson (trial D+N, *EMM* = 5.26), $F(1, 204) = 4.41$, $p < .05$, *MSE* = 3.84. This further supports the hypothesis that long narrations carry cognitive costs, and further supports the idea that the reason text is better for lengthy descriptive content is not influenced by synchronous perceptual events.

Hypothesis 5c: Learning from the text-then-diagram lesson (trial T+D, *EMM* = 5.54) was not shown to be significantly better than the diagram-then-text lesson (trial D+T, *EMM* = 5.83), $F(1, 204) = .96$, $p > .05$, *MSE* = 4.49. Moreover, learning form the narration-then-diagram lesson (*EMM* = 4.92) was not significantly different from learning from the diagram-then-narration trial (*EMM* = 5.26), $F(1, 204) = 1.44$, $p > .05$, *MSE* = 4.10. Taken together, this set of findings (5a, 5b, and 5c) shows that the cost to learning associated with the lengthy descriptive content overrides the effects of presentation order. This shows that cognitive influences on learning can be strong, and that there is a need for further investigation into ways to reduce the cognitive costs, such as those posited in Hypothesis 2.

Hypothesis 6: If grouping individuals based on differences in comprehension strategy (the use of internal representations) influence patterns of learning, these differences might be attributed to the formation and storage of internal representations. Group differences in comprehension strategy did not show differences in the multimedia effect: Linguistic Group, multimodal *EMM* = 5.13, unimodal *EMM* = 5.61, $F(1, 102) = 2.66$, $p > .05$, *MSE* = 4.36; Non-Linguistic Group, multimodal *EMM* = 5.13, unimodal *EMM* = 5.56, $F(1, 68) = 1.37$, $p > .05$, *MSE* = 4.53. However, after learning from a unimodal synchronous presentation (T&D), learners who use a non-linguistic comprehension strategy (*EMM* = 6.53) performed better than learners who use a linguistic strategy (*EMM* = 5.07), $F(1, 84) = 4.53$, $p < .05$, *MSE* = 5.04. Both groups probably experienced perceptual costs prior to creating their internal representations (i.e., the split attention effect (Mousvai et al., 1995)). However, Carpenter and Just (1975) suggest that learners who depend upon linguistic representations would represent the diagram linguistically. This further taxes their verbal-auditory working memory capacity, leading to an increased likelihood for overload and the need to select information.

These results show that considering the physical form of presentation alone does not enable the information designer to accurately predict learning. Therefore, considering the type of information (e.g., descriptive) in addition to the physical form of presentation (e.g. text versus narration) is important to learning. Modality effects of internal representations seem to exist. Considering solely a perceptual level explanation for the relationship between information design and learning is not adequate to account for multimedia learning effects.

Chapter V

CONCLUSIONS

The results of this study demonstrate that multimedia effects might be influenced by both physical (perceptual) and mental (representational) modality effects. The need to recognize and understand the structure and implications of these effects has been noted in a variety of domains, from computer science (Drommi et. al, 2001; Reeves et. al, 2004), to educational psychology (Schnotz, 2002). On more than one occasion during this experiment, cognitive benefits and cognitive costs associated with multimedia design interacted with each other and produced effects not commonly predicted by the present understanding of multimedia effects. Similar to effects observed when cross-modal dual task situations have been examined in other domains of psychology (e.g., Ittyerah, 1983; Lee & Kang, 2002; Treisman, 1960; Wickens & Lui, 1988), the effects of multimedia seem to be too complex to be explained by a single pattern of effects. Instead, the effects might depend upon both the presentation of task materials (such as shown in other domains by Ittyerah, 1983) and the cognitive requirements of the task (such as shown in other domains by Lee & Kang, 2002)

In this experiment, synchronous presentation was shown to increase the learner's ability to later retrieve the information compared to asynchronous presentation. These results support the view that verbal information and pictorial information kept simultaneously in working memory facilitates elaborative processing of the two different (and complementary) sets of information (Schnotz, 2002), and leads to better learning (Mayer & Sims, 1994). The effects that have led to a perceptual based view of the origins of multimedia effects might actually stem from the fact that visual-diagrammatic and auditory-descriptive materials presented in tandem facilitate elaborative processing by presenting complementary sets of information together. It is possible however, that presentation modality is not the only way to control these associated cognitive benefits. The fact that a multimedia effect was shown not to exist during synchronous presentation, and that during asynchronous presentation unimodal presentation was shown to be *better* than multimodal presentation, is the most salient example of how the present results go against the existing interpretation of multimedia effects. That is, contrary to previous research that has shown that substituting narration for on screen text improves learning (Mayer, 2001; Sweller, 1999), the simultaneous mixed modality condition did not produce better learning than the simultaneous unimodal presentation condition, and during asynchronous presentation, the opposite was in fact true. In turn, these results support the contention that the effects of dual mode presentation are more complicated than previously thought.

It is likely that some property of the descriptive content used in the present study created cognitive costs when presented in the form of a narration as compared to when it was presented via text. It might be that Wickens (2000) was right when he suggested that lengthy information should be visual (text) rather than auditory (speech) because of the greater permanence of visual information and the higher working memory demands of understanding speech (Wickens & Hollands, 2000). However, learning from the lengthy descriptive narration in the present study was not significantly different than learning from a text description when presented in tandem with diagrammatic materials. It is concluded that providing a diagrammatic representation that allowed learners to segment the problematic content might have mitigated the costs of lengthy narrations, but these costs remained during asynchronous presentation.

A consideration of the ease with which a chosen presentation modality is processed is important because, though *physical* properties of both visual and auditory information are perceived quite easily, processing capacity limitations might require the selection of a subset of *semantic* information before it is transferred to a semi-permanent form of storage and rehearsed as it awaits cognitive manipulations (Cherry, 1953; Cowan, 1988). Knowledge acquisition requires that information must undergo cognitive manipulations *after perception* to be committed to long term memory (Anderson, 1995). The impedance to learning during the asynchronous narration conditions (relative to both the synchronous narration condition *and* the asynchronous text conditions) was not one of perceptual interference, because there was no other information present to interfere. Instead it seems to have occurred after perception, while the learner was utilizing an internally represented form of the previously perceived information (akin to a second stage of sensory storage). The quality of this representation was affected by its external representation, but not perceptual interference.

The possibility that some effects of multimodal instructional materials occur while information is held and manipulated in working memory as it awaits transfer to long term memory is supported by the results of this experiment. Past research has shown that working memory representations that endure long enough after perception might include the activation of a subset of long-term memory (Norman, 1968), and modality specific components of working memory could result from the temporary activation of elements within long term-storage (Cowan, 1988). The scope of Sweller and Mayer's approach to describing the Multimedia effect concentrates on multimodal effects of the physical properties of presentation alternatives, with inadequate focus on multimodal effects of the semantic properties of information. The filter relevant to the effects observed in the present experiment seem to occur after perceptual information (physical properties) has been encoded into working memory, consistent with evidence for two stages of sensory storage (Cowan, 1988). Multimedia effects might be dependent upon modal properties of the *internal* representation.

Semantic properties of the information determine the processing demands required for cognitive manipulations and possibly influence the proportion of information processed before information loss occurs (Corteen & Wood, 1972). Therefore, we must consider these properties of information content in addition to the physical properties of their presentation. The use of cognitive resources to rehearse and maintain these complex representations seem to have reduced the amount of cognitive resources that remain for encoding and actively processing the information for committal to long-term memory. Therefore, the various presentations of the information influenced learning in different ways. However, due to the semantic properties of the information presented, this influence was not as simple as impedance from one presentation combination or facilitation from another. To the contrary, the presence or absence of complimentary materials, allowing the learner to control the rate at which they perceived and encoded the information, caused a change in the pattern of effects. That is, when complimentary materials were present at the same time (synchronous presentation conditions), there was no significant difference in recall performance after unimodal and multimodal presentation, but without the presence of these other materials (asynchronous conditions), unimodal presentation was actually better than multimodal presentation.

One additional piece of evidence for the position that internal representations affect learning from materials such as those used in this study is that the preferred comprehension strategy of individuals influenced performance after the synchronous text presentation. A group of untrained participants are not likely to approach a task in the same way (MacLeaod et al., 1978). Merely the fact that people can approach the same multimodal task in various ways supports the position that the presentation of information does not alone influence the efficiency of learning. Based on past research (Carpenter & Just, 1975; Glenberg et al., 1999; MacLeaod et al., 1978), individuals in this experiment were grouped into those who use linguistic and non-linguistic strategies to compare verbal and visual materials. Individual differences reflected in comprehension strategy and differential performance in learning from these materials further support the position that the presentation of information does not influence the efficiency of learning alone. Though the evidence from this experiment does not demonstrate that this occurs across all learning types (significant differences were found for only one experimental trial type: synchronous text), it does show that the possibility of differences may exist. It seems probable that those using a linguistic strategy represented the diagrammatic material linguistically (Carpenter & Just, 1975) and this interfered with representing the descriptive material linguistically. This, in addition to perceptual interference of presenting both types of information visually, may have added up to the effects observed. The effects observed in other conditions might have been reduced by tow factors: small sample sizes after grouping, and the possibility that all presentation types might not experience this effect. It is possible that some smaller effects were present, but not significantly significant within this sample. Further research is needed to examine how comprehension style influences learning from a variety of multimedia instructional materials.

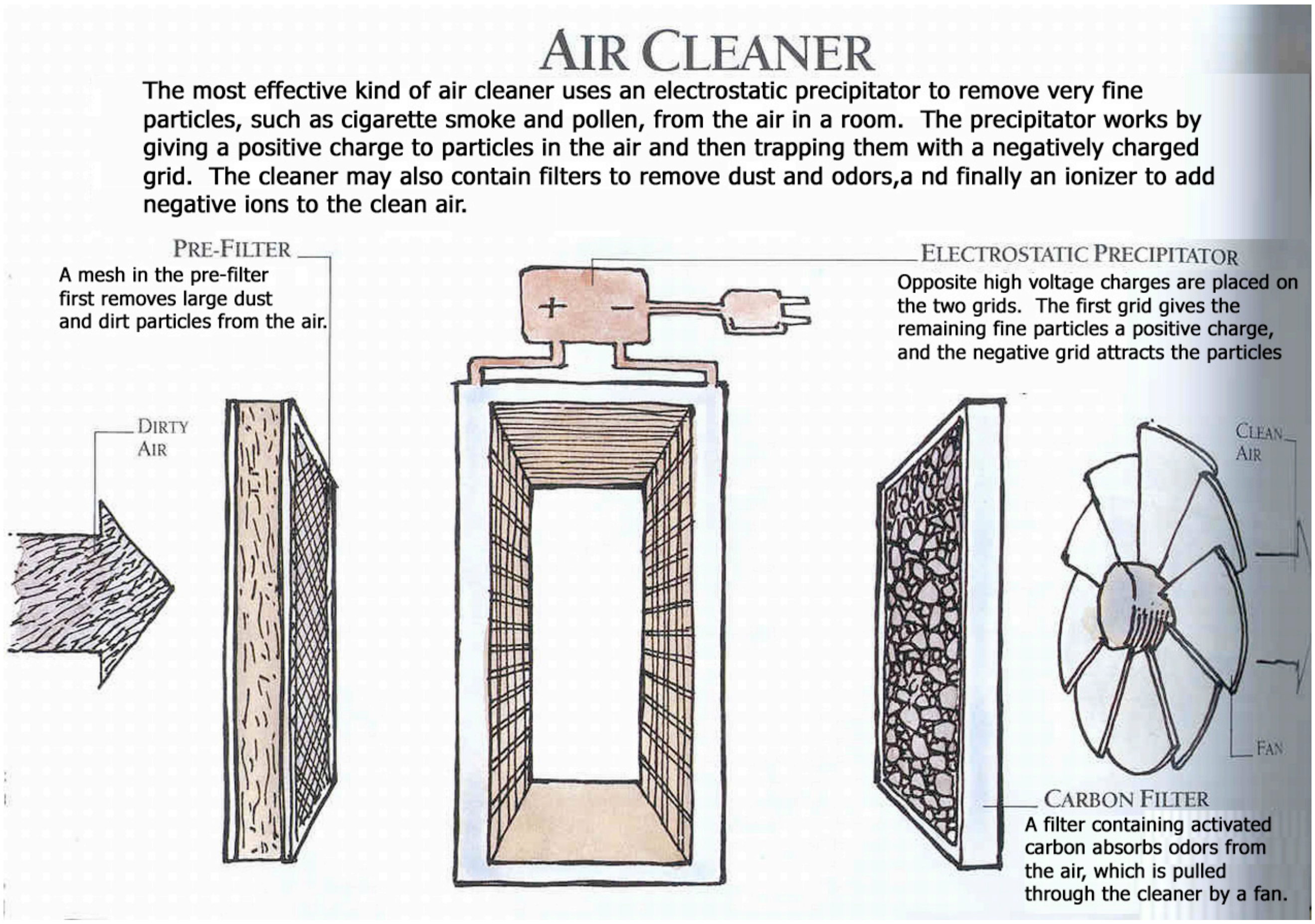
The results of this study suggest that design guidelines for the presentation of multimedia information need to consider the content and circumstances of the information to be presented, in addition to what is known about implications of its presentation form. One way to conceptualize these post-perceptual influences on the multimedia effect is to consider the internal representations of information that are formed while an observer attempts to perform mental operations on to-be-learned information. Other researchers (e.g. Baddeley, 1994; Pavio, 1986; Wickens, 2002) investigating the influence of internal representations on human behavior define internal representations separately from materials' external representation, but still as a dichotomy that separates auditory/verbal information from visual/spatial information. The motivation for this study was that, from these perspectives, visually and auditorily presented information included in multimedia learning environments might correspond imperfectly to visual and auditory modalities after perception. Future research on the multimedia effects in instructional materials should also consider the conceptualization of internal representations as a dichotomy that is not equal to, or a simple function of, the external representation.

It is time that developers of multimedia instructional material expand their understanding of multimedia effects to include a consideration of mental modality effects. The specific lesson to draw from this experiment is that descriptive content should not be presented via lengthy narration. This is direct support for Wickens' suggestion that lengthy information should be visual (text) rather than auditory (speech) because of the greater permanence of visual information and the higher working memory demands of understanding speech (Wickens & Hollands, 2000). However, it is still unclear whether the cause is the greater permanence of visual information or higher working memory demands of understanding speech. Designers of multimedia information might create better information presentation by offloading some visual-text material to auditory-narration presentation. However, it is likely that any benefits gained will be negated if designers are not careful to consider the limitations of the learner's working memory capacity. The need remains for a better understanding of these effects from an external and internal representation point of view. This will enable us to know when and why benefits or costs occur. Providing narrations with reduced content, or a means by which observers might divide lengthy narration into segments, is one way to cater to working memory capacity and thereby improve learning from multimedia instructional materials. It is my belief that any such effects that are manifested in the complex domain of learning can be transferred to other domains of information design, especially those that aim to foster non-trivial interaction with, or examination by, its observers.

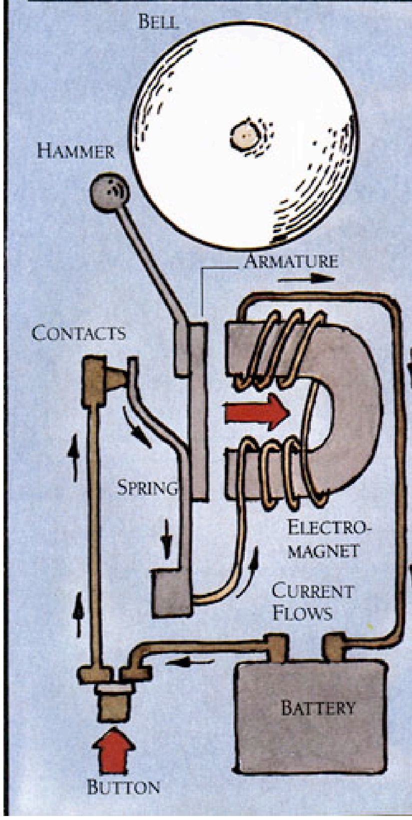
The next steps for this research area are to directly test the effects of providing shorter narrative subsections, and to expand the investigation to diagrammatic, including animated, materials. Including the factors that research currently posits about physical modality, and adding to it new discoveries about internal modality, I hope to work towards a basis for developing a taxonomy of multimedia effects. The vision of this taxonomy is an understanding of human information processing based on the effects of modality for both internal and external representations of information that can be used to provide guidelines for designing multimedia information.

APPENDIX A:

LESSON MATERIALS



THE ELECTRIC BELL



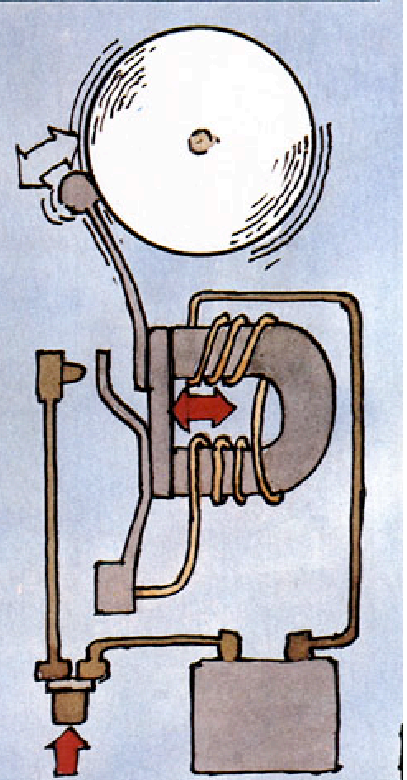
One of the many everyday uses of electromagnetism is the electric bell. The button at the door is an electric switch that sends current from a power source such as a battery to the striking mechanism. This makes a hammer move back and forth several times a second, sounding a metal bell An electromagnet and a spring alternately pull the hammer.

PRESSING THE BUTTON

When the button is pressed, the contacts are first closed. Current flows through the contacts and the spring to the electromagnet, which produces a magnetic field. This field attracts the iron armature, which moves toward the electromagnet against the spring and makes the hammer strike the bell.

THE BELL SOUNDS

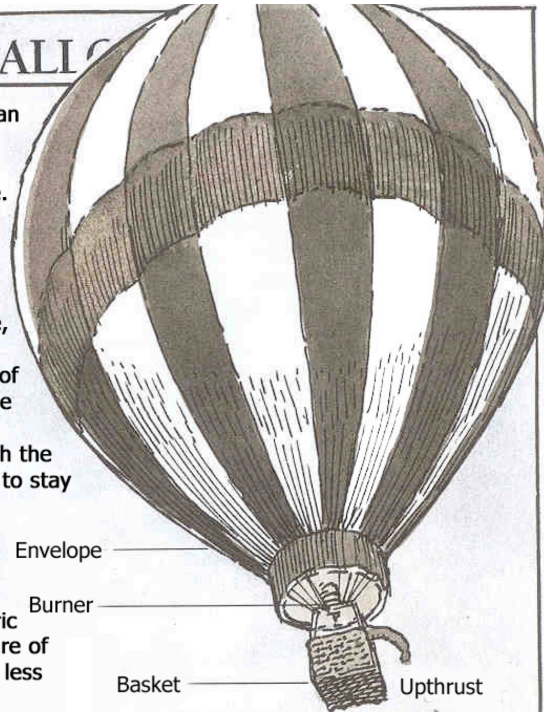
As the hammer strikes the bell, the movement of the armature opens the contacts. The current stops flowing to the electromagnet, which loses its magnetism. The spring pulls the armature back, and the hammer moves away from the bell. The contacts then close again, and the cycle repeats itself for as long as the button is pressed.



THE HOT-AIR BALL

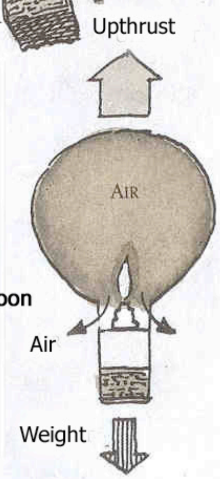
The envelope of a hot-air balloon has to be big so that it can displace a large amount of air, thereby creating sufficient upthrust to float tha basket and its occupants through the air. The balloon works like and underwater craft in reverse. Operating the burner heats the air into the envelope; the air expands and some escapes form the envelope. The overall weight decreases, and the upthrust carries the balloon upward. When the burner cuts out, the air in the envelope cools and contracts. Air now enters the envelope, increasing the balloons weight and causing it to descend. Fast descent can be achieved by opening a port in the top of the envelope. This partially deflates the envelope to reduce the upthrust.

A hot-air balloon has no means of propulsion and drifts with the wind. Intermittent blasts of the burner enable the balloon to stay at a constant height.



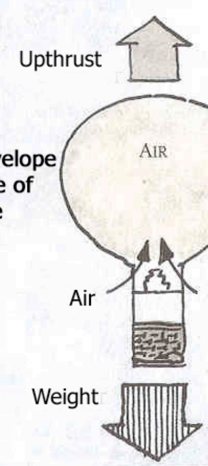
Envelope

The envelope of an airship is made of synthetic fabric and is not rigid, maintaining its shape by the pressure of gas inside. The gas is helium, which is seven times less dense than air and is non-flamable



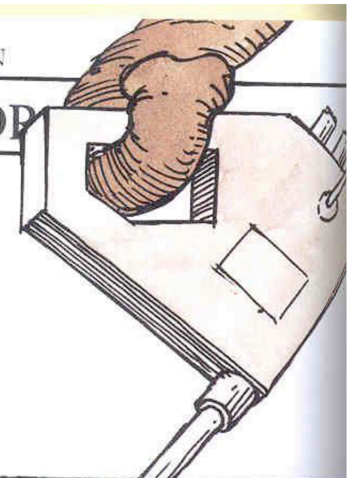
Ascent

The burner, which uses propane for fuel, heats the air in the envelope to a temperture of about 100 C. the air expands, and about a quarter of the hot air leaves the base of the envelope. The weight of the whole baloon is reduced to less than its upthrust, and the balloon rises.



Descent

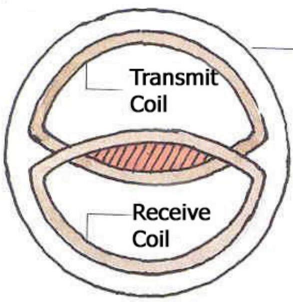
The burner cuts out and the air in the envelope cools. It contracts and air enters the base of the envelope, increasing the weight of the balloon to exceed the upthrust



METAL DETECTOR

The technology that enables us to discover buried treaasure also tests coins in ticket machines or vending machines, invisibly frisks people at airports and controls traffic lights. All these machines are basically metal detectors, and they work by electromagnetic induction.

When a peice of metal passes thourgh a magnetic field or the field passesthrough the metal, the field produces electric eddy currents that circulate in the metal. The eddy currents in turn produce their own magnetic field, and metal detectors work by detecting this field.

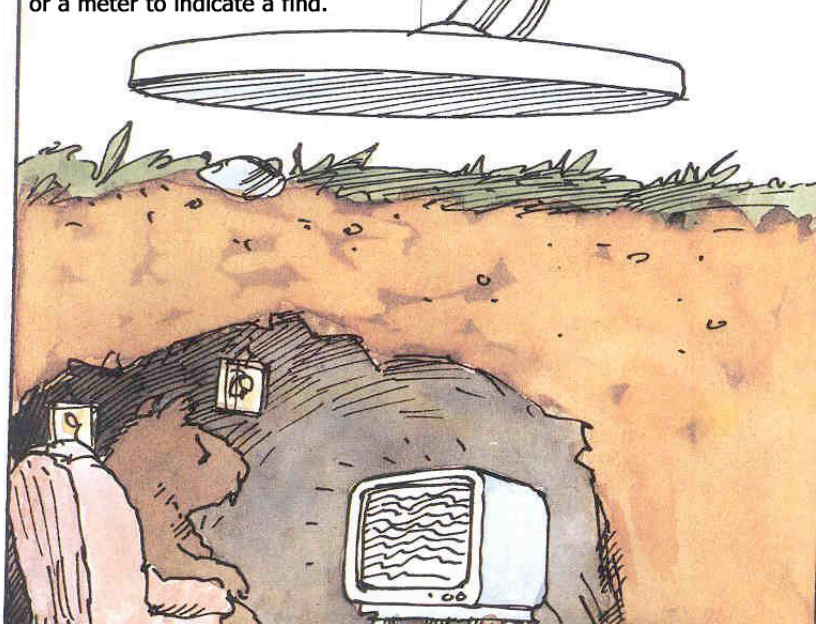


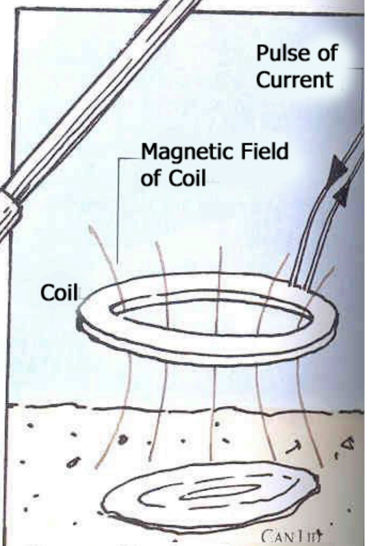
DETECTOR COILS

The transmit and recieve coils overlap so that each induces a current in the other. Normally, the two currents cancel out, but the magnetic field of a metal object distorts this balance and a low current appears in the receive coil.

DETECTOR HEAD

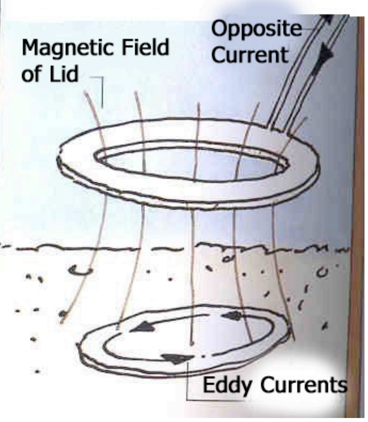
In the metal detector head, one coil usually transmits a magnetic field and another coil picks up the magnetic field produced by a metal object below. The recieve coil produces an electric signal that goes to a light, earphones or a meter to indicate a find.



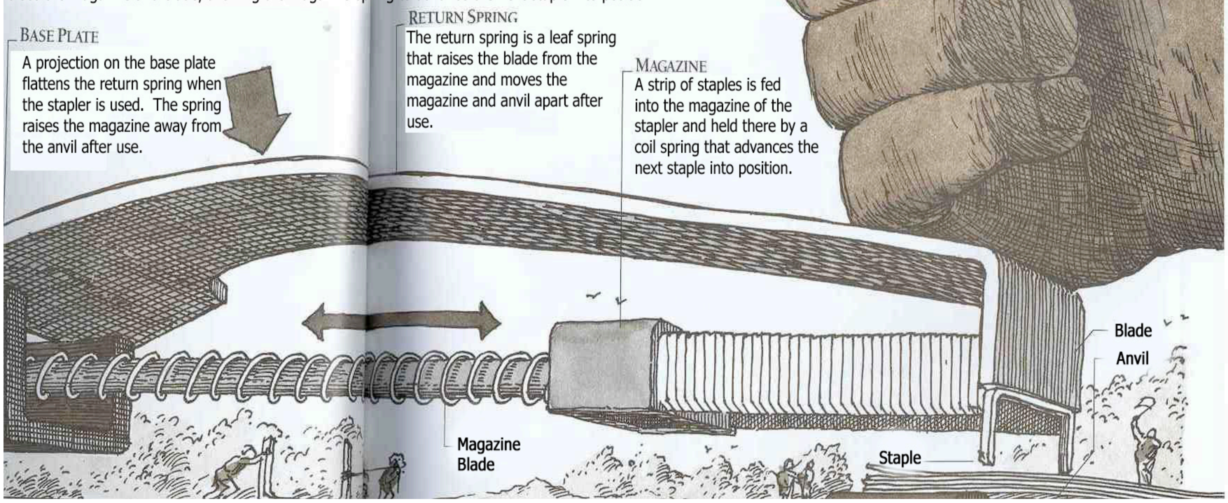


SINGLE-COIL DETECTOR

A brief pulse of current fed to the coil produces a magnetic field around the coil. The field enters the buried can lid. Eddy currents induced in the metal generate a magnetic field that in turn induces an opposite current in the now incative coil. Sensitve electronic components detect this current.



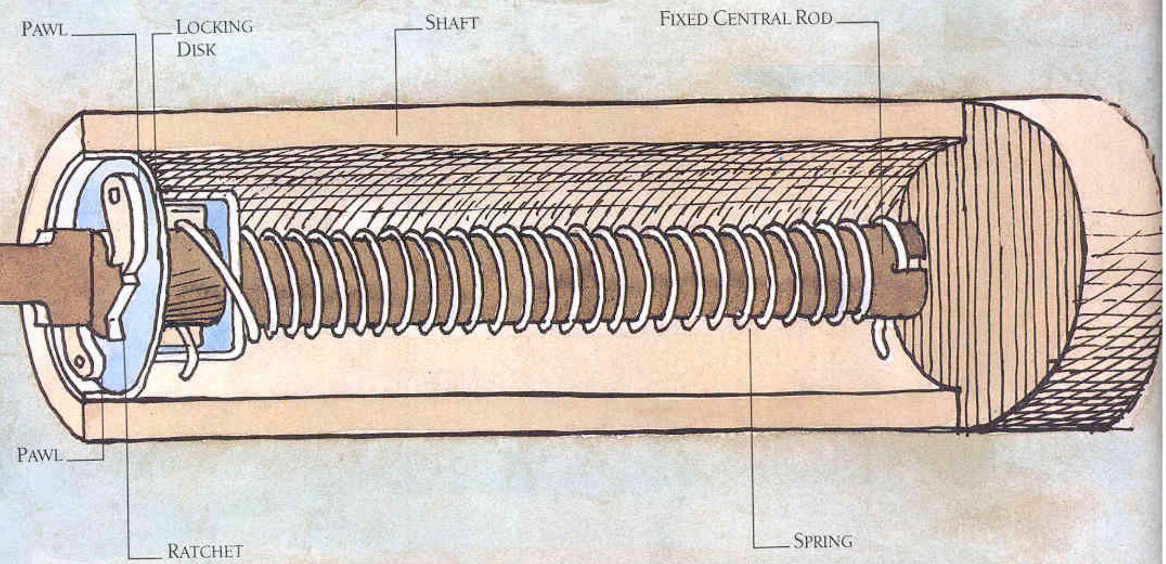
Stapler: A stapler is an everyday device that conceals an ingenious arrangement of springs. It uses both a coil spring and a leaf spring, which feed the staples along the magazine and return the stapler to its original position once it has been used. Pushing down the stapler causes the blade to descend into the magazine, forcing the front staple through the papers. The anvil bends the ends of the staple to clip the papers together. The return spring then raises the magazine and blade, allowing the magazine spring to advance the next staple into position.

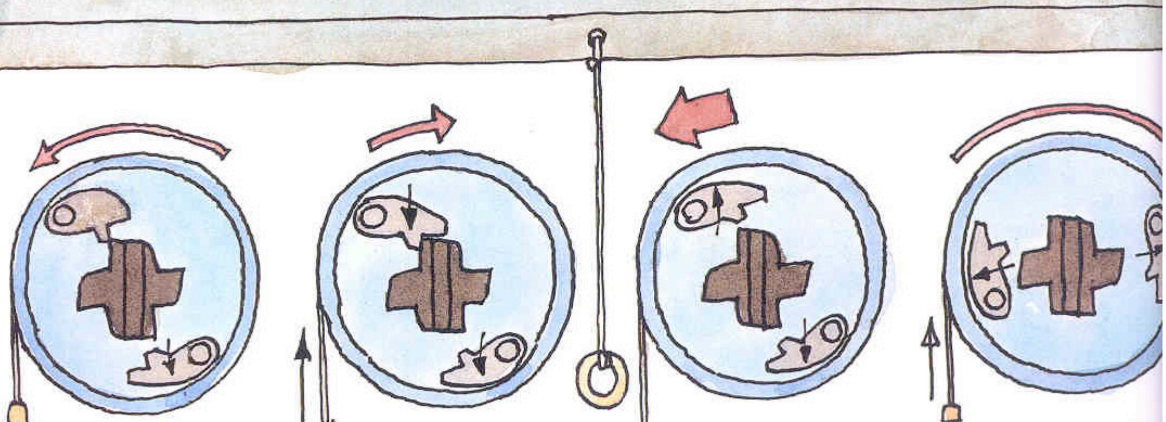


WINDOW SHADE

A window shade is lowered simply by pulling it down; the shade unrolls and remains in any position. To raise it, all that is needed is a sharp tug and the whole shade will roll up. But how can it tell a gentle pull from a sharp tug?

The shaft on which the shade is rolled contains a powerful spring. This winds up as the shade is lowered. A locking mechanism - a simple ratchet - prevents the spring unwinding if it is released gently. But when the shade is pulled suddenly, the ratchet no longer holds the shade in position. The motion makes a centrifugal device in the locking mechanism release the spring: the spring unwinds, releasing the energy that it has stored up, and up goes the shade.





Lowering the shade: As the shaft rotates, it turns the locking disk to wind up the spring. The pawls are hinged and move under the ratchet, which is fixed to the central rod and does not move.

Securing the shade: When the shaft stops, the spring pulls the locking disck back slightly. One of the pawls falls to engage the ratchet, securing the locking disk.

Freeing the shade: A tug on the shade rotates the shaft sharply, making the locking pawl move back and disengage the ratchet. The locking disk is now free to move.

Raising the shad The spring unwinds rotating the locking disk rapidly. Centrifugal force holds the pawls away from the ratchet, and the shade rolls up

APPENDIX B:

TEST QUESTIONS

Window Shade

The locking mechanism is a simple _____________.	Descriptive passage 1
The mechanism that controls the shade is housed in what is called the _____________.	Diagram
The spring is wrapped around the _____________.	Diagram
Centrifugal force holds the _____________ away from the ratchet.	Descriptive passage 5
As the _____________ rotates it turns the locking disk.	Descriptive passage 2
When the shade goes up, the spring _____________ its energy.	Descriptive passage 1
When _____________ the shade, the spring winds up.	Descriptive passage 1
When the shaft stops, the spring pulls _____________ back slightly.	Descriptive passage 3
A sharp tug makes the _____________ disengage the ratchet.	Descriptive passage 4
The _____________ is fixed and does not move.	Descriptive passage 1

Metal Detector

A low current appears in the _____________ coil in the presence of metal.	Descriptive passage 2
Electric _____________ currents produce the field detected.	Descriptive passage 1
The coils are positioned so that they _____________ and each induces a current in the other.	Diagram and Descriptive passage 2
The _____________ coils sends a signal to the light.	Descriptive passage 3
The detector coils are housed in the detector _____________.	Diagram and Descriptive passage 3
A single coil detector uses a _____________ of current.	Descriptive passage 4
Metal detectors work with _____________ induction.	Descriptive passage 1
Metal objects _____________ balance in detector coils.	Descriptive passage 1
In a dual coil detector, the _____________ coil creates a magnetic field.	Descriptive passage 3
In a single coil detector, metal generates a _____________ that induces an opposite currents in the inactive coil.	Descriptive passage 4

Air Cleaner

The Air Cleaner traps _____________ with a grid.	Descriptive passage 1
The ionizer adds _____________ to the clean air.	Descriptive passage 1
The pair of _____________ are connected to the power source.	Diagram
The air is pulled by a _____________.	Descriptive passage 4
The particles have a _____________ charge before being trapped.	Descriptive passage 3
The carbon filter absorbs _____________.	Descriptive passage 4
According to the lesson, the most effective air cleaners use an electrostatic _____________.	Descriptive passage 1
Opposite _____________ are placed on the grids.	Diagram Descriptive passage 3
The _____________ in the pre-filter removes dust and dirt particles.	Descriptive passage 2
In the example used in the lesson, the carbon filter is after the grids and adjacent to the _____________.	Diagram

Electric Bell

The magnetic field pulls the _____________ toward the electromagnet.	Descriptive passage 2
The movement of the armature _____________ the contacts.	Descriptive passage 3
The armature needs to be made of _____________ for the electric bell to function.	Descriptive passage 2
The _____________ is an electric switch.	Descriptive passage 1
The _____________ strikes the bell.	Descriptive passage 3
The force of the _____________ opposes the electromagnet.	Diagram
Current flows through contact and _____________ to electromagnet.	Descriptive passage 2
Electricity and _____________ produces magnetic field.	Descriptive passage 2
Ringing _____________ while the button is pressed.	Descriptive passage 3
The armature pushes against the _____________.	Diagram and Descriptive passage 1

Stapler

The _____________ spring brings the stapler back into position.	Descriptive passage 1
The _____________ holds a strip of staples.	Diagram and Descriptive passage 4
The strip of staples is held by a _____________ spring that advances the next staple into position.	Diagram Descriptive passage 4
The _____________ bends the end of the staples.	Descriptive passage 1
The _____________ descends, forcing the staple through the papers.	Diagram
The _____________ spring is flattened when the stapler is used.	Descriptive passage 2
The return spring is a _____________ type of spring.	Descriptive passage 3
The return spring is a projection of the _____________.	Descriptive passage 2
The _____________ is pushed so that it hits the anvil.	Diagram
Pushing down the stapler causes the _____________ to descend into the magazine.	Diagram and Descriptive passage 1

Hot Air Balloon

The pressure from the ____________(type) gas inside allows the Balloon to maintain its shape.	Descriptive passage 2
The burner of the Hot Air Balloon heats the air stored in the ____________.	Diagram and Descriptive passage 3
According to the lesson, air ____________ when it is heated.	Descriptive passage 1
The envelope made out of ____________.	Descriptive passage 2
The gas used in the balloon is ____________ (#) times less dense than air.	Descriptive passage 2
Constant ____________ is achieved through intermittent blasts of the burner.	Descriptive passage 1
The air heated to ____________ (temperature).	Descriptive passage 3
The HAB uses _____ a means of propulsion.	Descriptive passage 1
The burner uses ____________ (gas) for fuel.	Descriptive passage 3
Cooled air causes the balloon to ____________.	Descriptive passage 4

REFERENCES

Aarntzen, D. (1993). Audio in courseware: Design knowledge issues. *Educational and Training Technology International, 30*(4), 354-356.

Ainsworth, S.E., (1999) A functional taxonomy of multiple representations. *Computers and Education, 33*(2/3), 131-152.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.

Anderson, J. R. (1995). Learning and Memory: An integrated approach. John Wiley & Sons, Inc. New York, NY.

Baddeley, A., & Hitch, G.J. (1994). Developments in the concept of Working Memory. *Neurosychology*, *8*(4), 485-493.

Broadbent, D.E. (1958). *Perception and communication*. Pergamon Press, London.

Busey, T. A., and Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review, 101*: 446-469.

Carpenter, P. A., and Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review,* ***82***, 45-73.

Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25*(5):975-979.

Clark, H. H., and Chase, W. G. (1972). On the process of comparing sentences and pictures. *Cognitive Psychology,* ***3***, 472-517.

Cohen, A. (in press). *Selective attention*. Encyclopedia of Cognitive Science. Nature Publishing Group (Macmillan).

Corteen, R.S. & Wood, B. (1972). Electrodermal responses to shock-associated words in an unattended channel. *Journal of Experimental Psychology 94*: 308-13.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constrains within the human information processing system. *Psychological Bulletin, 104,* 163-191.

De Haan, E. H. F., Appels B., Aleman A., Postma A. (2000). Inter- and intramodal encoding of auditory and visual presented material: effects on memory performance. *Psychological Record, 50*, 577-586.

DeJong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance, 19*(5), 965-980.

Drommi, A., Ulferts, G. W., Shoemaker, D. Interface Design: A Focus on Cognitive Science. In *The Proceedings of ISECON 2001*, v 18 (Cincinnati).

Gellevij, M., van der Meij, H., de Jong, T., Pieters, J. (2002). Mulitmodal versus unimodal instruction in a complex learning context. *The Journal of Experimental Education, 70*(3), 215-239.

Glenberg, A. M., Robertson, D. A., Jansen, J.L, and Johnson-Glenberg, M. C. (1999). Not propositions. *Journal of Cognitive Systems Research, **1***, 19-33.

Good, T. L. and Brophy, J. E. (1990). Educational Psychology: A Realistic Approach. New York: Longman.

Gully, S. M., Payne, S. C., Kiechel, K. L., & Whiteman, J. K. (1999, April). *Affective reactions and performance outcomes of error-based training.* Paper presented at the meeting of the Society for Industrial and Organizational Psychology, Atlanta, GA.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.) Human mental workload (pp.139-183). Amsterdam: North-Holland.

Ittyerah, M, Broota, K. D. (1983). Inter and Intra modal processing of sensory-specific stimuli. *Perceptual and Motor Skills*. Vol 56(2), 507-517

Jolicoeur, P. (1999). Dual task interference and visual encoding. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(3), 596-616.

Jolicoeur, P., and Dell'Acqua, R. (1999). Attentional and structural constraints on visual encoding. *Psychological Research, 62*: 154-164.

Jonassen, D. (1999). Designing Constructivist Learning Environments. In C.M. Reigeluth

Kim, J. (1978). Introduction to Factor Analysis: What it is and How to Do It. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-013. Newbury Park, CA: Sage.

Kroll, N. E. A. (1975). Short-term memory. Eds. Deutsch, D. and Deutsch, J. A. Academic Press: New York, NY. Pp. 154-179

Landauer, T. K. (1995). *The trouble with computers: Usefulness, usability and productivity.* Cambridge, MA: MIT Press.

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

Lee, K., & Kang, S. (2002). Arithmetic operation and working memory: Differential suppression tasks. *Cognition, 83*(3), B63-B68.

MacLeod, C. M., Hunt, E. B., and Matthews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior 17*, 493-507.

Mayer, R. (2001) *Multimedia Learning.* Boston: Cambridge University Press.

Mayer, R. E. & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*, 389-401.

Mayer, R.E., Heiser, J., & Lonn, S. (2001). Cognitive constraints in multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*(1), 187-198.

McCaulay, D. (2003). *The New Way Things Work*. Boston: Houghton Mifflin Company

Mousavi, S.Y., Low, R, & Sweller, J. (1995). Reducing Cognitive load by Mixing Auditory and Visual Presentation Modes. *Journal of Educational Psychology*, *87*(2) 319-334.

Najjar, L. J. (1998). Prinicpals of educational multimedia user interface design. *Human Factors, 40*(2): 211-323.

Norman, D.A. (1968). Toward a theory of memory and attention. *Psychological Review, 75*(6), 522-536

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw Hill.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Potter, M. C. (1976). Short term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 509 - 522.

Reeves, L. M., Lai, J., Larson, J. A., Oviatt, S. (2004). Guidelines for multimodal interface design. *Communications of the ACM*, *47*(1): 57-59.

Schmidt, F. L. (1988). The problem of group differences ability test scores in employment selection. *Journal of Vocational Behavior, 33,* 272-292.

Schnotz, W. (2002). Towards an Integrated View of Learning from Text and Visual Displays. *Educational Psychology Review, 14*(2), 101-120.

Shiffrin, R. M., & Dumais, S. (1981). The development of automatism. In J. A. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 111-140). Hillsdale, NJ: Erlbaum

Sweller, J. (1999). *Instructional Design.* Melbourne: ACER Press.

Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12*, 244-248.

Wickens, C. D. (2002). Multiple resources and performance prediction. Theoretical Issues in Ergonomic Science, 3(2), 159-177.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance*. Upper Saddle River, NJ: Prentice Hall.

Wickens, C., & Lui, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors, 30*(5), 599-616.