The Implications of Working Memory on Symbolic Comparison

Ramona Eryuzlu

A thesis submitted to

the Faculty of Graduate Studies and Research

in partial fulfillment of

the requirements for the degree of

Master of Arts

Carleton University

Ottawa, Ontario

June, 2005

Your file Votre référence
ISBN: 0-494-10043-5
Our file Notre référence
ISBN: 0-494-10043-5

Abstract

The purpose of the present research was to examine which working memory resources (phonological, visuo-spatial, central executive) are accessed when responding to symbolic comparison problems, where the symbolic information has just been learned. Participants responded as quickly and as accurately as possible to either "Taller?" or "Shorter?" probe with respect to pairs of fictional individuals and their learned associated relative heights. Each participant also participated in a condition in which they had to concurrently perform one of three working memory secondary tasks (articulatory suppression, visuo-spatial tapping, or random letter generation). All three of the symbolic comparison effects of distance, end, and semantic congruity generally occurred in terms of latencies, accuracy, or both. Furthermore, although overall latencies increased and accuracy decreased when secondary tasks were performed, there was an interaction between the distance effect and the random generation condition in terms of accuracy only. The present results can only suggest that individuals require general working memory resources to solve symbolic comparison problems with stimuli, which were just learned within the experiment. The results, however, do not support the use of specific resources required by models that assert either semantic codes or spatial arrays in symbolic comparisons.

Acknowledgments

I would like to first thank my supervisor, Dr. Craig Leth-Steensen, for the opportunity of working in his lab and introducing me to the whole aspect of reasoning. I sincerely appreciate his efforts, guidance, insight, patience and sense of humour throughout my Master's program and for his financial support. All of his help has definitely made attaining my Masters degree in cognitive psychology an enjoyable experience.

I would like to thank my two lab-mates, Hossein Samavatyan and Jordan Schoenherr for making data collecting and waiting for our "employees" to show up (when and if they did) a much less lonely experience, and for being so easy to compromise with regarding the scheduling of lab times for our experiments.

I would also like to thank Christian Maillet for always listening to me, and for imparting his thesis experiences on me whenever I was in doubt of my capabilities. I have a feeling that we will be ruminating about experiments over Tim Horton's for a long long time to come! (Wink!)

Last and certainly not least, I would like to thank my family: Mom, Dad, Tina, and Crystal for your support and encouragement and continually asking, "How is your thesis going?" as what it seemed a constant reminder to not procrastinate.

Table of Contents

List of Tables

List of Figures

INTRODUCTION

As human beings, we are always reasoning to make sense of what has just occurred in our daily interactions in order to plan or predict what the next event will be. Usually, we reason using information that we already know and are familiar with (such as a horse is larger than a cat, or 1 PM occurs before 3 PM). Hence, an interesting issue involves what happens when we have just learned the material with which we must reason with. In this case, it would also be interesting to investigate if, and subsequently, how we use our working memory to keep track of such just-learned material. The present study examined the role of working memory in symbolic reasoning, wherein the relational stimuli have been learned within a laboratory setting.

The Three Common Reasoning Paradigms

Conditional Reasoning

Conditional relationships between propositions are represented in the form of rules such as "if p *then* q". The first part, "p", is termed the antecedent, and the second part, "q", as the consequent. A conditional reasoning paradigm typically contains a conditional sentence involving two propositions (the major premise) and a second premise involving the assertion or denial of one of these propositions (the minor premise; Evans, 1982).

If p then q

p

There are four types of conditional inferences. In the first type, *modus ponens*, the minor premise states the antecedent and the conclusion states the consequent. In the second type, *modus tollens*, the minor premise denies the consequent and the conclusion denies the antecedent (Evans, 1982). The two other possible inference types, *denial of the antecedent* and *affirmation of the consequent*, are only valid if the relationship between the antecedent and consequent are biconditional. Studies investigating conditional reasoning usually require participants to state a conclusion that may be made on the basis of such premises, or have them assess whether a given conclusion is valid or invalid (Evans, 1982; Toms, Morris, & Ward, 1993).

Syllogistic Reasoning

Syllogistic reasoning entails thinking logically about category relationships and involves two premises that are assumed to be true, for example, "All Dalmatians are dogs" and "All dogs are mammals." One premise relates the subject of the argument (Dalmatians) to the shared or middle term (dogs), and the other premise relates the middle term to the predicate (mammals). In syllogistic arguments, the nature of relationships between the subject, predicate, and middle terms include those of set inclusion, overlap, and exclusion, namely, "all", "some", "none", and "some not" (Gilhooly, Logie, Wetherick, & Wynn, 1993; Copeland & Radvansky, 2004). Studies investigating syllogistic reasoning paradigms usually require participants to state a conclusion that may be made on the basis of such premises, or have them assess whether a given conclusion is valid or invalid.

Relational Reasoning

A relational reasoning problem, also called a linear syllogism or a transitive inference problem, comprises two or more premises, each describing a relation between two items. A participant's task would be to determine the relationship between two (or more) items not occurring in the same premise given the observation that that at least one of the items overlaps between premises (Sternberg, 1980). For example, "Mighty Joe Young is mightier than King Kong". "King Kong is mightier than Magilla Gorilla". Who is mightiest?" The overlapping item is King Kong, and the participant must synthesize the information in the premises in order to determine the relationship between the two non-adjacent items (Sternberg, 1980). Studies investigating relational reasoning paradigms usually require participants to state a conclusion that may be made on the basis of such premises, or have them assess whether a given conclusion is valid or invalid.

It is hypothesized that one can make an inference in linear reasoning problems either by applying rules on the linguistic representations of the premises or by constructing a spatial array (Evans, Newstead, & Byrne, 1993; Vandierendonck & De Vooght, 1997). According to a linguistic representation position, individuals try to reveal the underlying meaning of the premises. For example Clark (1969) proposed that when people are presented with the premise

"Alan is better than Bob,"

They form a linguistic representation, derived from a dimension of goodness:

A is more good, B is less good.

"Better" is an adjective that describes the relative goodness of two items. However, if the adjective "worse" was used instead of better in the above premise, it would not only describe the relative badness of the two items, but it would also give the conclusive information that they are both bad, and thus require additional information to be stored in comparison to the premise involving the "better" adjective (Evans et al., 1993).

Adjectives such as "better," "taller," and "bigger" are called *unmarked* terms, whereas adjectives such as "worse," "shorter," and "smaller" are called *marked* terms.

Psycholinguistic evidence has revealed that marked terms are somewhat more difficult to grasp than their associated "unmarked" terms (Johnson-Laird & Byrne, 1991). Therefore, response times are often faster in relative judgment tasks involving unmarked adjective pairs (Holyoak & Mah, 1981; Petrusic, 1992).

According to the spatial array position, individuals form a spatial array of the terms used in the premises. In this view, the top of the array represents the highest or most positive end of a scale. Participants usually prefer to work through premises from the top down, and to construct an array from an *end-anchored* premise, in which the first noun phrase refers to an item at one end of the array (Johnson-Laird & Byrne, 1991). A fundamental difference between the spatial array position and the linguistic rule-based position is whether or not marked or unmarked adjectives are represented in an identical format. From the spatial array position, participants are unaffected by whether marked or unmarked adjectives were used to create the array. Although a bottom-up array is more difficult to construct than a top-down array, once the array has been formed, the difference between the marked and unmarked adjectives used to create the array disappears in that they have been represented in exactly the same format. However, the linguistic view suggests that marked and unmarked adjectives are represented in different formats. For example, "A is worse than B" is represented on the linguistic dimension of badness in that A is more bad, and B is less bad, while the premise "B is better than A" is represented on the linguistic dimension of goodness in that B is more good, and A is less good. If people retain this linguistic information, then the markedness of the adjectives should affect them even after they have constructed the representation. Potts and Scholz (1975) found that the time to answer a question was not affected by whether the relations in the premises were marked or unmarked, suggesting that individuals represent marked and unmarked adjectives in an identical manner, which supports the spatial array theory.

Johnson-Laird (1983) proposed the mental models theory, a variant of the spatial array theory, where the mental models that reasoning individuals create are regarded as integrated structural representations of the relations given in the premises. It is hypothesized that these structural representations are constructed in working memory. Therefore, problems that entail the (reasoning) individual to construct and keep in memory more than one model are supposed to be more difficult than problems that entail the construction of only one model. An important implication of the mental models theory is that the load on working memory causes errors in reasoning performance. Consequently, reasoning individuals are likely to make more errors and have increased RTs when the number of models that they have to entertain is increased (Evans, Newstead, & Byrne, 1993; Johnson-Laird, 1983; Vandierendonck & De Vooght, 1997; Vandierendonck, Dierckx, & De Vooght, 2004). Hence, the present study investigates the role of working memory on relational symbolic reasoning, a paradigm which will be introduced in the next sections of this Introduction in terms of the task, the three major symbolic comparison effects, and the major types of symbolic comparison models.

The General Symbolic Comparison Paradigm

The general symbolic comparison paradigm (essentially a form of the relational reasoning paradigm) involves the study of comparative judgments of symbolic stimuli according to attribute information that must be retrieved from memory. Symbolic comparison stimuli usually consist of sets of four or more items whose ordering has been learned (or artificially induced) within a laboratory setting. Such learning can involve training participants either to associate symbolic items (such as names or nonsense syllables) with a set of pre-experimentally ordered stimuli such as circle sizes (Moyer & Bayer, 1976) or personality adjectives (Birnbaum & Jou, 1990). Alternatively, it can involve training participants to make (correct) comparative judgments for the full set of adjacent pairs of symbolic items that come from some a priori, but arbitrarily derived, linear ordering (Leth-Steensen & Marley, 2000). Symbolic stimuli from naturally occurring attribute continua in semantic memory such as digit magnitudes (Banks, Fujii, & Kayra-Stuart, 1976; Poltrock, 1989), color similarity (Te Linde & Paivio, 1979) and animal sizes (Jamieson & Petrusic, 1975; Marschark, Azmitia & Paivio, 1985; Shoben, Cech, Schwanenfugel, & Sailor, 1989) have also been investigated.

A typical symbolic comparison experimental task involves asking participants to compare pairs of either learned or natural stimuli. For example, in a symbolic comparison experiment using several animal sizes, a participant might indicate that a mouse is smaller than a cat when asked which animal is smaller. The participants' reaction times and accuracy are usually measured to study how easily individuals compare particular items differing in size or any other dimension.

Symbolic Comparison Effects

Symbolic comparison studies usually result in three major psychological effects: the distance effect, the end effect, and the semantic congruity effect. It is assumed that the study of these effects will ultimately lead to a fundamental understanding of the way in which people represent and process relational order information, in addition to the way in which people store and retrieve memorial magnitude information.

The distance effect occurs when, as the subjective difference between the stimuli increases, comparison response times decrease and accuracy increases (De Rammelaere & Vandierendonck, 2003; Leth-Steensen & Marley, 2000; Moyer, 1973). The end effect is characterized by faster comparison response times and higher accuracy for comparisons containing stimuli at the outermost ends of a particular range of stimuli that are being compared than for comparisons containing stimuli in the middle of that range (Shoben et al., 1989; Leth-Steensen & Marley, 2000). Semantic congruity effects are characterized by faster comparison response times (and sometimes higher response accuracy) for comparisons in which the polarity of the instructions corresponds with the relative correct position of the stimuli within the set of experimental stimuli (i.e., to select the smaller of two smaller stimuli) than for comparisons in which the polarity of the instructions is incongruent (i.e., to select the larger of two small stimuli; Cech & Shoben, 1985, Leth-Steensen & Marley, 2000; Petrusic, 1992; Petrusic & Baranski, 1989, Shaki & Algom, 2002). The following section will describe a number of models that have been proposed to explain the processes occurring in symbolic comparisons.

Symbolic Comparison Models

The Random Walk Model

Random walk models are composed of a system of assumptions that have various justifications. Some assumptions are central features of random walk models, others are approximations intended to facilitate computation of functional relationships, and still others are justified by analysis of the task and stimuli. A typical random walk model for the symbolic comparison task would transpire as follows. When individuals are presented with a stimulus pair, the magnitudes of the stimuli are represented in a noisy environment (the brain) and consequently the internal analog representations of the magnitude are considered random variables with unknown distributions. Since the analog representations are random variables, reliable comparisons of the two stimulus magnitudes involve an accumulation of information on the differences in magnitudes for some time, with the likelihood of a correct response increasing with time (Poltrock, 1989).

Participants would press a response key if the value of their mental counter surpasses a preset criterion, called a "boundary". If the resulting value does not exceed the boundary, the participant would then repeat the process of accumulating information until a boundary is reached. Fewer repetitions are typically required to surpass a boundary when the separation of the distributions of the two random variables is large, which consequently results in a smaller RT. Given that the level of separation is directly related to the difference between the logarithms of the presented stimuli, RT should be inversely related to the level of separation between the presented stimuli (i.e., the distance effect; Poltrock, 1989). Variations of the random walk model have been described and investigated by psychophysicists such as Poltrock (1989), Link (1990), Birnbaum and Jou (1990), as well as Buckley and Gillman (1974).

Poltrock (1989), for example, proposed a random walk model of digit comparison that describes how individuals choose which of two digits is greater. His model describes only how the participant decides between stimuli, and how the time required and the decision errors depend on properties of the stimuli and instructional manipulations. In each trial of a digit comparison task, participants encode the two digits presented, decide which is greater, and make their response based on their decision. Thus, the response time RT_{ij} for a pair of digits *i* and *j* can be expressed as

$$RT_{ij} = DT_{ij} + R,$$

where DT_{ij} is the decision time and R is the duration of all other processing. According to this model, errors are caused by incorrect decisions and not by response execution errors.

The main characteristic of the random walk model is that a random variable that is associated with the choice dimension is added up until one of two boundaries is reached. In each time interval Δt, the decision process computes the difference (d_{ij}) between the two digits, which are transformed into analog magnitudes (*i* and *j*), and adds this difference to an accumulated value, D_{ij}. Therefore, when d_{ij} is added to $D_{ij,}$ it is actually the analog magnitude of the left-hand digit *i* minus the analog magnitude of the right-hand digit *j*. The mean value of d_{ij} is μ_{ij}

$$\mu_{ij} = g\ (i) - g\ (j).$$

In the above equation, g (*i*) represent the mean value of the analog magnitude for the digit *i*. The difference, d_{ij}, is considered a step in the random walk, and is assumed to be equivalent to an independent random sample from some difference distribution $f\ (d_{ij})$ with mean μ_{ij}. The source of variability in d_{ij} may arise from both the establishing of the analog magnitudes and the computation of the difference in magnitudes.

Poltrock (1989) gives an example of two possible paths followed by D_{ij}. In one path D_{ij} starts at an initial value of zero and adds successive values of d_{ij} until it reaches boundary *A*. The random walk stops, and the decision is made when the absolute added difference exceeds that boundary. The decision time DT_{ij} is Δt multiplied by the number of differences that are computed. The random walk is supposed to end at a boundary and not pass it. However, the initial value C of D_{ij} may be a non-zero value, resulting in bias for one of the response choices. A response bias reduces the time to reach one decision but increases the decision time for a decision going the opposite direction. Remarkably, in each of Poltrock's (1989) experiments, the estimated response bias for comparisons involving the digit one was closer to the boundary than to the analog magnitudes of any other digit. Therefore the random walk boundary is reached in a single step when the digit one is presented, leading to very fast and accurate responses.

Birnbaum and Jou (1990) developed a biased random walk model comprised of a set of equations to describe the distance effect, end effect, and semantic congruity effect, using a single scale of subjective magnitude. The authors developed the model upon examination of their experimental results (1990; Experiment 1). In their study, participants evaluated the difference in likeableness between pairs of fictitious persons. First, their participants were required to memorize the association of each name with a personality. For example, they were required to memorize that Carl is malicious and Bill is loyal. The participants were later asked to rate how much more they would like Bill than Carl (in the format of a subtraction problem such as, Bill - Carl = ?), using integers from -9 to 9. A difference of 9 indicated that the first "person" was much more likeable than the second person in the equation, a difference of zero meant that the two persons being compared were "equal in likeableness", and negative numbers indicated that the first person presented was less likeable than the second. Participants also generated response times by pressing a key to indicate which person was "more" likeable, and in a separate block of trials, they indicated which person was "less" likeable. The authors then used their data to develop their theory of a biased random walk. The authors developed three key equations as follows:

$$1.\ D_{ij} = a\,|\,s_i - s_j\,| + b$$

$$2.\ TM_{ij} = m_o + m_i m_j / \,|\,s_i - s_j\,|$$

$$3.\ TL_{ij} = l_o + l_i l_j / \,|\,s_i - s_j\,|$$

In all three equations, s_i stands for a subjective scale of stimulus magnitude. In the first equation, D_{ij} stands for the predicted "difference" in likeableness between person *i* and person *j*; in the second equation, TM_{ij} is the predicted response time to choose which of the two presented names is more likeable; in the third equation, TL_{ij} stands for the predicted response time to choose which name is less likeable. Also, *a* and *b* are linear constants that express the linear relationship between "difference" ratings and subjective difference; m_o and l_o are additive constants that represent the time required to read the stimuli and to press the key. Finally, m_i, m_j, l_i, and l_j are bias parameters that should account for both the end effect and the semantic congruity effect.

End effects and semantic congruity effects may occur when the bias parameters are not constants. The end effect occurs when the values of *m* and *l* are smaller for end stimuli than for middle stimuli and the semantic congruity effect occurs when the bias values of *m* are relatively smaller for more likeable[1] stimuli than for less likeable stimuli compared with the relative values of *l*. Since the bias values are multiplied, the time to compare two stimuli will be faster if *either* stimulus is biased to yield a fast time, and

[1] Or any other term on the high end of a scale, such as "larger" or "better," etc.

responses will be slower to stimuli that are relatively unbiased (e.g., when both stimuli are toward the middle of the series). The extra assumption in this model is that the starting point for the random walk is the product of the bias parameters. The product of the bias parameters suggests that when either one of the stimuli is near the end of the series, the participant begins the random walk closer to the correct choice boundary, which produces a faster (correct) response, requiring less accumulation of evidence from the stimulus difference (s_i - s_j). Therefore, the latter two equations provide the average times for *correct* responses in a random walk model, assuming the product of bias parameters ($l_i l_j$ *or* $m_i m_j$) signifies the starting point of the walk.

The Leth-Steensen and Marley (2000) Model of Symbolic Comparison

Leth-Steensen and Marley (2000) have developed a related connectionist, evidence accrual model that successfully simulated the distance, end, and semantic congruity effects. The model includes a connectionist learning component (that is used to generate internal analog-based representations for each member of a pair of discrete input stimuli) and dual evidence accumulation decision-making components. It assumes that comparison responses can be based either upon information concerning the positional difference between the presented stimulus items or upon information concerning their endpoint status. Both forms of information are derived from noisy strength values generated within the model that are accumulated over time to produce explicit response time and accuracy predictions.

In the connectionist, evidence accrual model, two distinct vector name representations are used as input units to the network to simulate the paired comparison task. Activation progresses in a "feed-forward manner" along the connections departing from each of the input units to their corresponding left and right stimulus representation units, from each of these two representation units to the output unit, and then along the separate "Shorter?" and "Taller?" instructional pathways to the left and right response units. Simulated response latencies are obtained by assuming that an overt response is made as soon as the activation accumulated from one of the response units reaches a positive threshold (T). The "decision latency" in this model is the number of simulation steps required for an overt response to be made. It is assumed that incorrect responses occur with the addition of noise to the accrual process.

According to this model, the distance effect occurs because the absolute value of the level of activation on the output unit increases as the stimuli in a simulated comparison pair become further apart in the ordering. The increase of the activation level results in the passing of stronger signals to the response units, which then lead to faster simulated responses.

End effects are modeled based on the end anchor strategy suggested by Potts (1972, 1974). Hence, the model includes end anchor units that are activated according to the distance (or similarity) of the stimuli to the ends of the ordering. High anchor unit activations provide evidence for the presence of an end term within the presented comparison pair. Furthermore, left and right response units are connected to each of the anchor units by "Shorter?" and "Taller?" instructional pathways that are given weight values that are set in such a manner as to activate the correct response and inhibit the wrong response. To obtain response time simulations, the two evidence accrual processes (analog comparison and end anchor) are assumed to be independent and to run at the same time. Therefore, simulated responses occur once the integrated activation from either route surpasses the response threshold *T*. The reason why analog similarity route is also assumed is because the end anchor similarity decision process alone is unable to give a description for the response times of the middle Split 1, 2, and 3 pairs (the term "split" refers to the subjective distance between two stimuli).

In order to model the semantic congruity effect, Leth-Steesen and Marley (2000) took an analogous approach to that of Cohen, Dunbar, and McClelland (1990), who conceptualized the Stroop effect as a competition between the continuous processing of information within two separate pathways (i.e., a word-naming pathway and a color-naming pathway) that meet at a common output (i.e., response) stage. Thus, Leth-Steensen and Marley (2000) made two computational assumptions concerning the role of semantic interference in generating semantic congruity effects using the concept of competing processes proposed by Cohen et al. (1990). The first assumption is that the information processing within each of the instructional pathways is not all or none. Rather, the processing within the irrelevant pathway competes with the processing within the relevant instructional pathway. The second assumption is that the amount to which the two instructional pathways compete with each other hinges on the comparative "shortness" or "tallness" of the given stimuli. Semantic interference effects occur whenever anchor unit activations supply information that is incompatible with the task instruction. Hence, interference effects occur as such information leads to a type of implicit response competition that then has to be overcome in order to respond correctly or possibly because such information interferes with the formation and maintenance of the internal representation of the context supplied by the comparative instruction.

The Semantic Coding Model

The semantic coding model proposed by Banks (1977) has two stages. The first is an encoding stage, whose purpose is to produce a linguistically based, propositional semantic description of the stimuli. The second comparison stage uses the semantic codes to decide upon the correct response. The total RT for performing the task is expected to be the sum of the times taken up by both stages. In the case of comparing digits, this model presumes that a digit is initially coded as either larger (L+) or smaller (S+) than a criterion mid-point on the numerical continuum. The criterion point varies from trial to trial and has a skewed distribution where its mean is below the mid-point of the digits used in the task. Thus, the skewed distribution results from the fact that smaller numbers are spaced further apart on the subjective continuum than the larger numbers, thus making the criterion mid-point more likely a smaller digit. In this model, the same criterion is always used for both digits on a given trial. If both fall above it, the codes will be L+/L+, if both fall below it, the codes will be S+/S+; and if they are on both sides of the criterion, they will be coded as either L+/S+ or S+/L+.

The comparison stage determines the correct response by comparing the previously stored instructional codes with the stimulus codes generated by the encoding stage. The instructional codes are processed as L+ for "choose larger" and S+ for "choose smaller." The comparison stage uses the earliest codes available for the stimuli. If the codes are L+/S+ or S+/L+, there is an immediate match between the instructional code and one of the two stimulus codes, resulting in a relatively small RT. However, if the stimulus codes are L+/L+ or S+/S+, the comparison stage consequently takes longer to search for more detailed codes in order to tell the digits apart. The overall RT is presumed to be a mix of latencies of the comparison stage resulting from the coding of the stimulus pair. The model predicts RT by showing how the probabilities of these codes are affected by various factors. First, the distance effect emerges because the greater the split or separation between the digits, the more likely are they to be on different sides of the criterion (i.e., S+/L+ or L+/S+). A "min effect" (i.e., RT increases with the size of the minimum digit) occurs due to the fact that the criterion is usually located among the smaller numbers. In other words, there is a greater probability of a S+/L+ coding and a decreased RT if min is a small value.

According to the semantic coding model, processing should be faster when the codes for the digits match the instructional codes than when they do not, which explains the semantic congruity effect. For example, under the instruction "choose larger" (coded as L+), participants would select the correct member of the pair 8,9 (coded L+/L+, and eventually coded as L/L+) faster than if the instruction were "choose smaller" (coded as S+). The S+ instruction would take longer for stimuli coded L/L+, as their codes must be transformed to S+/S. Thus, the model predicts an interaction between the instructions and the general size of the digit pair. Banks et al. (1976) also hypothesize that, for digit stimuli, there is less of an effect of congruity for the smaller pairs. Since there is more S+/L+ coding for smaller pairs, "choose smaller" and "choose larger" function will yield similar results for smaller pairs, leading to a funnel-shaped cross-over effect. Banks et al. (1976) further hypothesized that the size of the congruity effect will weaken as the split increases because the larger the split, the more likely the occurrence of L+/S+ coding where a difference between stimulus and instructional codes is unlikely. All of these effects were found in the results of the experiments presented by Banks et al. (1976).

The differential coding model (Shoben et al., 1989) supplements the semantic coding model with the assumption that the time required to generate a binary code for each stimulus depends on the magnitude of the presented stimulus. That is, individuals can code items of extreme magnitude easier and faster than items of intermediate magnitude. The differential coding model differs from the semantic coding model in that it predicts that the greater difficulty in coding items of intermediate magnitude results in a bowed serial position effect. The bowed serial position effect is simply the tendency for pairs of extreme magnitude to be discriminated more readily than pairs of intermediate magnitude (i.e., the end effect). Therefore, individuals code items of extreme magnitude quickly, so that these codes can be passed on to the comparison stage relatively quickly, resulting in faster response times. Pairs of intermediate magnitude require more time in the code generation stage and thus lead to slower response times.

The differential coding model can also explain interactions between symbolic distance and serial position. At minimum symbolic distance, pairs in the middle of a series are composed of items that are themselves of intermediate magnitude, such as "*pelican-weasel*". According to the semantic coding model, both of the objects in this pair are difficult to code, thus requiring a long response time in terms of a comparison to occur. As symbolic distance increases between the middle pairs, they will involve items of more extreme magnitude. Thus, as pairs involve items of more extreme magnitudes, codes for these pairs are easier to obtain. This reasoning also applies to pairs of extreme magnitude. The combination of the effects leads to an attenuation of the bowed serial position effect at high symbolic distance.

The Spatial Scanning Model

In a study by Moyer and Bayer (1976), participants associated an artificial class of concrete CVC nouns with referents (circles) that varied only in size. This procedure leads to what is called an "induced linear ordering" (Bower, 1971). A separate group of participants learned the same "nouns" with the size difference between the adjacent circles increased by two millimeters. Regarding the circle size differences between the sets, the two sets of circles were called "small range" and "large range". In the experimental condition, both groups had to decide which of two symbols stood for the larger referent, when prompted. Furthermore, within each series, a second group of participants made perceptual comparisons between the circles themselves, so that the effect of the range manipulation on the perceptual comparison time could be assessed.

Moyer and Bayer (1976) found that for both the small and large range circle groups, RT decreased as the ordinal position difference between the CVCs increased. They also found that at each ordinal position difference, CVCs were compared faster when they represented the large range circles. Furthermore, the large range circles were also more rapidly compared than the small range circles in the perceptual condition. However, although there was a main effect of ordinal position difference and of circle size range, there was no interaction of the two. In addition, the effect of circle size range in the memorial groups was consistent across individual comparison pairs. Similar results were found for the perceptual groups. Finally, for each circle size range within an ordinal position difference all pairs differed significantly from one another in RT (i.e., it decreased as the pairs became larger in accord with the semantic congruity effect). These differences were not characteristic of the perceptual groups.

Based on their findings, Moyer and Bayer (1976) developed a model. They proposed that the participants store an internal representation that maintains the absolute size of the referent circle with each allotted nonsense syllable. When the participants later have to decide which of two CVCs stands for the larger circle they locate each CVC in memory, retrieve the absolute size value associated with it, compare the two values, and select the syllable associated with the greater value.

Moyer and Bayer proposed that their participants stored these CVCs in memory ordered according to size along an imaginary axis in mental space. When asked which of two syllables stands for the larger circle the participants enter this ordered array from the "larger" end and successively examine each entry until one of the two displayed CVCs is found. They then draw a sample from the associated distribution of absolute values, similarly locate the remaining CVC and retrieve its value, and then compare the values according to a "random walk" model such as that proposed by Buckley and Gillman (1974).

In the related ends-inward scanning model, Woocher (1976, as cited in Woocher, Glass, & Holyoak, 1978) proposed that a longer search would be required to access the central items in a relatively long list. Therefore, the bowed serial position effect is expected to be increasingly apparent as list length increases. Woocher, Glass, and Holyoak (1978) tested this model, among others, by having participants merge the orderings of two eight-term learned series (e.g., A > B > C > D > E > F > G > H and I > J > K > L > M > N > O > P). Each ordering consisted of eight two-syllable occupation names. The relation used in forming the ordering was always "taller than", and each participant received two different lists. In the first part of the experiment, the participants had to judge whether two items presented from within an eight-term series were taller; they later had to determine whether two items within the list were adjacent. In the second part of the experiment, the participants were told that the shortest person on one of the lists was taller than the tallest on the other list, and they had to subsequently make adjacency judgments with given pairs from the newly joint series.

For the eight-term orders, RT decreased as the distance between the items increased from one item to seven items. The distance effect was also significant when just the terms without the end items were taken into account. For the adjacent pairs, the serial position curve resembled a bow shape in that RT was significantly smaller for comparisons involving the tallest and shortest ends of the distribution, rather than the middle of the distribution. Furthermore, RT was smallest at the tallest end of the distribution, presumably reflecting the semantic congruity effect that occurs when the instructions (i.e., "taller") match the end item being considered. Interestingly, reaction times between pairs of various size differences often overlapped, even when the pairs that included end terms were ignored, which suggested to the authors that serial position effects are relatively continuous and not due to the end items alone.

For the combined 16-term orders, the mean RT was longer than for the separate 8-term lists. Although distance was confounded with the relative number of between- vs. within-list comparisons, there were significant distance effects for both between-list and within-list comparisons when they were considered separately. End effects were found in that reaction times to pairs containing either the first or last term of the 16-term ordering were consistently faster than those containing no end items. Once again, this effect was mainly owing to pairs containing the "tallest" term rather than the "shortest" term.

Because the serial position effects for the 16-term orders were comparable to those obtained with the separate 8-term orders, these data suggested that participants in fact merged the two lists into one "mental array." Generally, the RT data reflected the relative positions of terms within the entire list, with little residual effect resulting from the fact that the sequence was derived from two previously separate sequences.

The serial position effects obtained for adjacency judgments were similar to the pattern found for the "taller" judgments about the same pairs. Moreover, RT was faster overall when pairs were presented in the correct order rather than in the reverse order from tallest to shortest. However, the advantage of the correct order disappeared for pairs near the "shorter" end of the list. These results are consistent with the hypothesis that participants used an internal array to perform the adjacency judgments. The authors proposed that adjacency judgments involved a search component; the left item (which is most likely read first) is first located in the array, initiating a "scan" from that array position to locate and identify adjacent items. It is possible that such a scan may have a directional bias (Holyoak & Patterson, 1981). Since the list was to be learned in the order "tallest" to "shortest," the scan would generally move from the "tall" items in toward the center of the list. According to these findings, Woocher's (1976) ends-inward scanning model seems most plausible.

As the present study explores the role of working memory on symbolic comparisons, the next sections of this Introduction will give some background on the development of the working memory model by Baddeley and Hitch (1974) and present some research findings involving working memory and reasoning abilities.

The Working Memory Model

Working memory consists of several specialized components of cognition that allow humans to understand and mentally model their immediate environment, to retain information about very recent cognitive events, to maintain new knowledge, to solve problems, and to begin, connect, and perform current goals. Working memory communicates with long-term memory the outcome of its operations (Baddeley & Logie, 1999). The original model proposed by Baddeley and Hitch (1974) includes a central executive controlling mechanism and two subsidiary or "slave" systems, called the phonological, or articulatory, loop and the visuo-spatial sketchpad, which are dedicated for the processing and temporary maintenance of material within a particular domain (i.e., verbally coded information and visual and/or spatial information, respectively).

The Central Executive

The central executive is theorized to have the capacity to coordinate performance on two separate tasks, to switch retrieval strategies, to attend selectively to one stimulus and inhibit the disrupting effect of others, and to hold and manipulate information in long-term memory (Baddeley, 1986, 1996, 1998). The central executive is also believed to manage and coordinate the activities of the phonological loop and visuo-spatial sketchpad (Baddeley & Logie, 1999). Because disorders that involve a decrement in central executive control are associated with damage to the frontal lobes, one approach to understanding central executive processes is to study the function of the frontal lobes (Baddeley, 1996; 1998).

Several methods have been employed to analyse the central executive, consequently leading to its present theorized capabilities. One method of analysing the central executive is by studying the effects of dual tasks. A common subject group for studying the central executive is Alzheimer disease (AD) patients, as they have a significant deficit in executive control, as well as in episodic memory (Baddeley, 2002). In a study of attention control, Logie, Cocchini, Della Sala and Baddeley (2000) first manipulated only the difficulty of a single task and found that performance of AD patients did not differ from that of young and elderly controls. However, when performing a dual task, say a visuospatial pursuit-tracking task[2] and a digit span task together, AD patients performed significantly worse than elderly and young participants (Baddeley, 2002). Furthermore, Baddeley, Logie, Bressi, Della Sala, and Spinnler (1986, as cited in Baddeley, 1996) examined the central executive's capacity to co-ordinate information from the two slave systems. The visuospatial pursuit-tracking task was combined with each of the following tasks: articulatory suppression, reaction time to a tone, and a digit span task. The authors found that although the tracking task did not impair articulatory suppression ability, it did impair performance on the concurrent digit span and reaction time tone tasks in AD patients, but not in the elderly or young controls. Since the capacity to combine performance on two tasks has been established as a necessary function of the central executive, the results suggested that AD patients have pronounced executive impairments in comparison to both elderly and young controls.

A common task that is associated with the central executive is random generation, and it will be utilized as a secondary task in the current study. Random generation has been useful as a secondary task, disrupting the operation of the central executive component of working memory in tasks such selecting the next move in chess or

---

[2] A visuospatial working memory paradigm in which a participant is asked to keep a pointer on a moving spot of light; performance is measured by the total amount of time they keep the pointer on the target.

acquiring artificial grammar (Baddeley, Emslie, Kolodny, & Duncan, 1998; Towse, 1998). Baddeley (1986) applied Norman and Shallice's (1980) Supervisory Attentional System (SAS) to his analysis of the effects of random generation on the central executive. The SAS system is proposed to handle novel events and unexpected emergencies by intervening to modify ongoing behaviour. When participants are asked to generate a series of random letters or digits, the SAS is proposed to intervene in order to avoid the tendency of participants to produce stereotyped output such as the alphabet, common acronyms, or counting (Baddeley et al., 1998). Random generation requires attention, thus the nature of it could be considered the opposite of what is currently being accepted as being automatic (Baddeley, 1996).

The Phonological Loop

The phonological loop is hypothesized to consist of two components: the *phonological store* and an *articulatory control process* (Baddeley, 1986; Baddeley & Logie, 1999). Articulatory suppression, the word length effect, the phonological similarity effect, and irrelevant speech effect are all areas of research involving the phonological loop that support the hypothesis that verbal short-term memory comprises a temporary storage system, the phonological store, and an active rehearsal system, the articulatory control process.

Articulatory suppression is the prevention of articulation of to-be-remembered words, by repeating a simple syllable aloud (such as "*the*"), that impairs short-term memory for verbal material but not visual material. This suggests that it does not only distract general resources away from the memory task but rather that it prevents maintenance of the information in a speech-based store. The word length effect refers to the fact that short-term memory is better for shorter words than longer words. This phenomenon suggests that short-term memory for verbal information depends on the limited rate at which we can rehearse information to prevent it from decaying (Andrade, 2001). The phonological similarity effect is characterized by poorer short-term memory for similar-sounding stimuli (Conrad & Hull, 1964), and it supports the notion of a temporary storage system specifically for speech-based or phonological material. The memory traces for similar-sounding items are assumed to be harder to discriminate at recall (Andrade, 2001). Murray (1967) found a remarkable effect when participants were required to recall sequences of phonologically similar consonants presented visually, namely, that participants showed no indication of a phonological similarity effect when they were required to suppress articulation (Baddeley, 1986). The irrelevant speech effect refers to the finding that memory is better when items are presented against a quiet background than against one with irrelevant speech (Andrade, 2001). In the present study, articulatory suppression will also be utilized as a secondary task.

The Visuo-Spatial Sketchpad

The other slave system of working memory is the visuo-spatial sketchpad. The purpose of the sketchpad is to provide temporary storage of visual and/or spatial information (Logie, 1995). Logie (1995) fractionated the visuo-spatial sketchpad into a passive visual working memory, called the *visual cache* and an active spatially based system called the *inner scribe*. The visual cache is thought to be passive and contain information about static visual patterns. McConnell and Quinn (2000; as cited in McConnell & Quinn, 2004) have established that, in order to cause interference within the visual cache, there has to be a dynamic aspect to the interfering display, provided by the dynamic visual noise technique. Dynamic visual noise involves the presentation of a visual noise field while participants are presented with input that they are required to encode as a visual image. The noise field comprises an array of black and white dots in which the change in rate between black and white throughout the extent of the array has been shown to interfere with visual images.

The inner scribe is the spatial working memory that maintains dynamic information about movement and movement sequences, and is linked with the control of physical actions. Logie (1995) states "the scribe provides a means of 'redrawing' the contents of the visual cache, offering a service of visual and spatial rehearsal, manipulation, and transformation". Tasks that draw on the inner scribe resources include asking participants to move their arm to follow a moving target or tap a series of keys laid out on a table or keyboard. In the present study, visuo-spatial tapping will be utilized as a secondary task.

Previous Studies Examining the Role of Working Memory in Reasoning

Conditional Reasoning and Working Memory

Toms et al. (1993) administered a series of experiments to establish the role of working memory on conditional reasoning by exploring the involvement of the visuo-spatial sketchpad, the phonological loop, and the central executive components on conditional reasoning. Toms et al. (1993) developed a series of conditional reasoning problems wherein participants were required to evaluate each of the four inferences for each of four conditional reasoning rules (again, modus ponens, modus tollens, denial of the antecedent, and affirmation of the consequent). Each problem consisted of two premises, together with either a valid or logically wrong conclusion. The primary task in both experiments required that participants indicate whether or not the conclusion necessarily followed logically from the premises and to signal their response by pressing a mouse button to indicate "yes" or "no".

The experiment exploring the effects of concurrent visuo-spatial sketchpad tasks on conditional reasoning included both a tapping task and a tracking task. In the tapping task, participants were required to tap the centre key of a Moar Box (the Moar Box comprises a keypad with 25 keys laid out in a 5 x 5 array). In the tracking condition, participants had to tap around the border of the Moar box at a minimum rate of two taps per second, beginning just before the start of a trial. The difference between these two tasks essentially involves the amount of spatial monitoring needed. The experiment specifically examined the possibility that conditional reasoning might entail visuo-spatial resources. If conditional reasoning requires the visuo-spatial sketchpad, a concurrent spatial task should interfere with reasoning performance and should further produce a marked effect when the more demanding modus tollens inferences are presented. An example of a modus tollens inference used in the study is:

Premises: If it is a circle, then it is not red

It is red

Conclusion: It is not a circle

The authors examined both the participants' "willingness to accept inferences" and response latency; however, they did not examine participants accuracy on the conditional reasoning problems. The results revealed that although there are the typical patterns of differences between inference type in terms of both willingness to accept inferences and response latency, the visuo-spatial tasks had no effect on reasoning performance. These results suggest then that the visuo-spatial sketchpad does not appear to be implicated in conditional reasoning.

Toms et al. (1993) then repeated the above experiment, except that the participants were under articulatory suppression and memory load secondary task conditions. In the articulatory suppression condition, participants were required to repeat, in order, the digits 1-6 at a rate of at least two digits per second. In the memory load condition, an experimenter orally read a list of the digits 1-6 prior to the starting of a conditional reasoning trial. The participants were required to repeat this list over and over at a rate of as least two digits per second, until they had made their response to a particular trial. On each memory load trial a different random ordering of the six digits was presented. Although the articulation load is the same in the two conditions, the memory load condition differs from the articulation condition because the participant cannot predict the ordering of the digits. The results from this experiment indicated that memory load only affected the frequency of acceptance of modus tollens trials in that it reduced the likelihood that participants would accept modus tollens inferences. Furthermore, it was shown that only memory load increased response latencies. There was no detrimental effect of concurrent articulation.

Given these results, Toms et al. (1993) asserted that the lack of an effect of tracking or articulation on conditional reasoning implies that the two working memory slave systems are not required, and that the most probable level at which interference occurs is the central executive level. It was suggested that the memory load condition, which also involves the phonological loop, loaded on the central executive that carried the excess memory load from the phonological loop.

A recent study done by Duyck, Vandierendonck, and De Vooght (2003), however did find an effect of concurrent tapping on conditional reasoning problems involving spatial content. For example,

> If Pete lives to the right of Paul then Stan does not live to the right of Kurt
>
> Pete lives to the right of Paul
>
> Conclusion: Stan does not live to the right of Kurt.

The problem started with a representation of the first premise centered on the screen. Participants were able to take all the time they needed to read and understand this premise and then initiate the presentation of the next premise. The participants also initiated the solution presentation, and pressed one mouse button if they thought the conclusion was valid and the other mouse button otherwise. Duyck et al. (2003) found that when the premises in the reasoning problems were based on spatial propositions (as opposed to non-spatial propositions), the first premise was processed approximately 3.5 s slower in the dual task condition compared to the single task condition, resulting in an interaction between premise content and the secondary task condition. These results suggested that when the relational content is spatial, visuo-spatial working memory is involved for the construction of a model of the premise. Visuo-spatial working memory resources, then, seem to be automatically involved if the content calls for a visuo-spatial representation, since the interaction only occurred for the spatial content problems.

Syllogistic Reasoning and Working Memory

Gilhooly et al. (1993) explored the effects of working memory on syllogistic reasoning performance. The participants in their study were instructed to attempt to draw their own conclusions from an argument with 2 premises previously presented, and then choose from five possible conclusions. Measures were taken on each trial of the times the participants spent viewing the pairs of premises, of the times taken to indicate conclusions after the response alternatives were displayed, and of the responses made. The participants were instructed to perform one of three working memory dual tasks in time to a metronome beat. The central executive task was composed of random number generation of the numbers 1, 2, 3, 4, and 5; the phonological loop task was to repeat the numbers 1, 2, 3, 4, and 5 over and over again; the visuo-spatial task was for the participant to tap four switches on a board in a clockwise direction.

Gilhooly et al. (1993) compared each dual-task condition with its own control. The authors also tried to identify the probable strategy used by each participant in the control and dual-task conditions and to assess the distribution of strategies over conditions. Their results showed that a matching strategy gave the best fit to the data in all conditions. The matching strategy is characterized by individuals who simply match the logical form in the conclusion to the premise that is more conservative (the degree of conservatism, from most to least, is 'No,' 'Some not,' 'Some,' and 'All'). For example, for the syllogism "Some dogs are fierce dogs; No corgis are guard dogs," an individual using the matching rule would respond with "No corgis are fierce dogs" (Gilhooly et al., 1993).

In terms of accuracy, random generation was the only secondary task that impeded syllogistic reasoning. Furthermore, participants in this study seemed to shift toward a guessing strategy under random generation. In terms of response time only random generation had a significant effect on the premise processing times. Conclusion response times were not affected by any of the secondary tasks. The authors assert that the lack of effect for random generation on conclusion reporting time signifies that the effect of the random generation is specific to the premise processing phase and not just a general slowing-down effect.

Recently, Copeland and Radvansky (2004) found that individuals with relatively larger working memory capacities[3] reasoned better on syllogistic reasoning problems than individuals with smaller working memory capacities, in that they were faster and more accurate in their responses. Furthermore, individuals with larger capacities were more likely to use mental models than individuals with smaller capacities.

Relational Reasoning and Working Memory

Vandierendonk and De Vooght (1997) conducted two experiments testing the use of working memory components during reasoning with temporal and spatial relations in four-term series problems. The first experiment tested reasoning accuracy under different dual-task conditions derived from the hypothesis that reasoning relies on working memory resources. Participants in the first experiment were randomly and evenly assigned to control, articulatory suppression, visuo-spatial interference, and central executive interference dual-task conditions. Each participant solved four-term linear syllogisms problems that varied with respect to content (temporal or spatial relations) and problem type. Five different problem types were used: two one-mental-model problems and three two-mental-model problems. One-model problems consisted of a series of three premises that described a linear ordering of elements in a connected way, for example:

---

[3] As measured by the operation span (Turner & Engle, 1989)

The guitar is to the right of the violin.

The guitar is to the left of the drum kit.

The drum kit is to the left of the piano.

In half of the one-model problems, the question asked for the position of the second and the third terms, so that it could be answered by recalling the second premise. The other half of the problems asked for the position of the first and the third term, which required a transitive inference. The two-model problems contained three premises that gave indeterminate descriptions of the positions of the four elements. For example:

The guitar is to the right of the violin.

The piano is to the left of the guitar.

The guitar is to the left of the drum kit.

Three types of questions were asked with respects to the premises. Relating to this latter example, these questions were about either the position of the piano and the guitar, the position of the piano and the drum kit (transitive inference), or the position of the piano and the violin (no conclusion possible). In all five problem types, the question was in the form, "Where is *x* relative to *y*?" and three possible answers were provided. The answer alternatives were: "*x* is to the left of *y*," "y is to the left of *x,*" and *"there is no correct solution."* The premises in the spatial form consisted of relations among concrete objects as in the examples given above, whereas premises in the temporal form illustrated actions performed by a person, such as "Charles went to the play before going to the movies".

Within each problem, the three premises appeared one by one in the centre of the screen, each for 7 s, and then the problem statement was shown for a further 30 s. The participants had to answer the question within the 30 s period by pressing the corresponding button on a three-button computer mouse with their left hand.

Participants in the three secondary task conditions started the secondary task 5 s before the presentation of the first premise and continued until they responded to the problem. In the articulatory suppression condition, the participants were instructed to repeat aloud a sequence of four digits (1, 3, 6, 8), at a rate of about two to three digits per second. In the visuo-spatial interference condition, participants were instructed to tap a sequence of four keys (1, 3, 6, 8) continuously on the numerical keypad of the computer (with their right hand) at a rate of 2-3 taps per second. Participants in the central executive interference condition were requested to keep track of a random pattern of short and long intervals before a click was presented. Participants were required to respond to the click as quickly as possible by pressing the zero key of the numerical keypad.

Vandierendonk and De Vooght's (1997) first experiment resulted in all three secondary tasks having comparable inhibitory effects on performance accuracy. Furthermore, problems that required the participants to construct two models were significantly more difficult than those requiring one model. Since all the secondary tasks significantly interfered with the participants' accuracy with solving the problems, it is possible that any secondary task may impair performance. Considering the possibility of nonspecific interference, the authors attempted to clarify which aspect of information processing during the reasoning task is responsible for the interference with a second experiment.

The findings of Experiment 1 were consistent with the hypothesis that the participants construct one or more models representing the premise information. The materials and procedure in Experiment 2 was similar to Experiment 1, except for a few changes. Vandierendonk and De Vooght (1997) assumed that constructing models require time. Hence the authors assumed that most of the 7 s interval during which each premise was visible in Experiment 1 was used to construct and to elaborate these models. They hypothesized that the participants might find forming models more difficult and thus abandon that method of solving the problems within a shorter time interval. In order to test this possibility, half the participants in Experiment 2 were required to perform the reasoning tasks with speeded premise presentation, while the other half received self-paced presentation of the premises. Furthermore, the authors opted to use three-model problems over two-model problems because the load on working memory may be greater when three models have to be considered before making an inference. The three-model problems consisted of a series of three premises that were consistent with three possible orders of four objects, for example:

The guitar is to the right of the violin.

The violin is to the right of the piano.

The drum kit is to the right of the piano.

Other than choosing three-model problems over two-model problems, the same types of problems were used as in Experiment 1.

Vandierendonk and De Vooght (1997) also decided to implement the secondary task conditions as a within-subjects variable, instead of a between-subjects variable. The secondary tasks were similar to those used in Experiment 1. The results showed that participants again made significantly more errors in the three dual-task conditions than in the control condition. Interestingly, participants made more errors in reasoning under the articulatory suppression condition, which was presumably due to the rather complex verbal nature of the task. Moreover, the visuo-spatial and central executive secondary tasks produced comparably greater amounts of errors among the participants' reasoning ability. This finding reflects the possibility that mental models that require visuo-spatial resources are constructed in linear relational reasoning.

Furthermore, only secondary tasks loading the VSSP and the central executive slowed information processing during premise presentation for the self-paced premise presentation group. This finding supports the hypothesis that during premise processing, the verbal information given in the premises is converted into a visuo-spatial code by the central executive. Therefore, the data suggest that primarily visuo-spatial information is kept in memory during relational reasoning. The existence of central executive suppression effects suggests the possibility of maintenance of abstractly coded information by the central executive.

Vandierendonk and De Vooght (1997) asserted that the effects of the secondary tasks on premise reading times are due to interference rather than to the relative difficulty of the secondary tasks. Random-interval repetition task interfered more with the premise reading task when the relations in the premises were temporal than when they were spatial; this is an indication that the results cannot be explained by a simple difficulty factor.

Experiment 2 also varied the premise presentation conditions. Solution accuracy was poorer under speeded than under self-paced presentation conditions. Interestingly, the participants in the self-paced condition needed 7.12 s to read the premises, which is approximately the same as the time given in the first experiment (7 s). Assuming that a visuo-spatial representation is constructed as the participants read the premise, it is implicit that the construction of the representation needs a minimum amount of time. The shorter the time available, the poorer the resulting representation will be which should consequently result in impaired performance.

Symbolic Comparison and Working Memory

Theberge (2003) recently explored the effects of working memory secondary tasks on symbolic comparison. Theberge took four six-item lists from normative size rating results in the Shoben et al. (1989) study. One of the lists was shown before the experiment started, and the participants were asked whether they were familiar with the items. The items on the list were ranked from smallest to largest and the participants approved their size ordering on their list. Each participant responded to only one of the lists (used equally often across participants) throughout the experimental session. A second experiment conducted in this study had a similar procedure except that a different list of words was used for each block.

On each trial, the instruction for comparison ("Larger?" or "Smaller?") was first presented at the top of the screen for 1000 ms. After the pair of stimuli appeared side-by-side in the center of a computer screen, participants had to decide which stimulus (the one on the left or the right) was either smaller or larger by pressing the appropriate response key with the index and middle fingers of their left hand. The next trial started 1000 ms after the previous response. Furthermore, each participant participated in four conditions (counterbalanced for order across the 24 participants), three of which involved a secondary task. In the articulatory suppression condition, participants were instructed to repeat aloud four letters, F, J, N, W, at a rate of 2-3 letters per second. In the visuo-spatial interference condition, the participants were asked to tap out a sequence of four number keys (1, 2, 6, and 8) with the index finger of their right hand (on the number pad of the keyboard, at a rate of 2-3 numbers per second). In the random generation condition, the participants were instructed to repeat aloud a sequence of four letters F, J, N, and W, in random order.

The analyses of the response time data led to several important findings. First, with few exceptions, all three of the symbolic comparison effects of distance, end, and semantic congruity occurred in both experiments. Furthermore, overall symbolic comparisons within the three secondary task conditions were significantly slowed down compared to the control condition. The three secondary tasks delayed primary task response times at different rates. For instance, the random letter generation condition delayed symbolic comparison performance the most. Both the articulatory suppression and the visuo-spatial interference conditions slowed down response times at about the same rate on average.

In almost all cases, however, there were no interactions between the three symbolic comparison effects and the secondary task conditions. This result would appear to suggest that the three secondary tasks are delaying response time in general, and not affecting the decision process per se. These results indicate that perhaps working memory is not involved in symbolic comparison decisions; therefore, it would seem to suggest that symbolic comparison is a highly automated process (at least for correct responses with the kinds of stimuli used in this experiment).

In the analysis of the accuracy data, in both Experiments 1 and 2, there was a decrease in accuracy both as split (i.e., distance) decreased and with the addition of a secondary task. Moreover, unlike the analysis of the reaction time data, the distance effect was enhanced when people had to perform a secondary task. This enhanced split effect was most prominent for the random letter generation secondary task condition, although a larger split effect also existed for both of the other two secondary task conditions (especially in Experiment 2). In terms of accuracy measures, then, the secondary tasks did seem to have an effect on the symbolic comparison decision process (i.e., that it might not be as automatic as concluded earlier). Therefore, the fact that such an enhanced distance effect did not also occur in the reaction time measure is puzzling because increases in decisional processing difficulty that increases the degree of error responding across splits usually should also result in corresponding significant enhanced delays in reaction times. One way to explain the finding of increased errors without a delay in reaction time is to note that this situation occurs mainly under conditions in which people trade off accuracy for speed by reducing the amount of evidence they require to make a decision. Hence, a major consequence of sharing cognitive resources between two tasks might be that people spend less time on the actual decision process to compensate for any potential delay in responding.

The Present Study

The present thesis was a replication of the study conducted by Theberge (2003), with the exception of a couple of significant changes. The first change is that in the present study, the symbolic comparison stimuli were learned within the experiment setting itself (i.e., an artificially induced ordering), rather than occurring naturally. The second change is that the participants in this study were subject to a control condition and only one secondary task condition, instead of all three secondary task conditions.

The present study consisted of two parts. Participants first underwent a learning phase of determining the relative heights of six imaginary individuals. In this phase, each participant was presented with each of the five comparison pairs consisting of the stimuli that were adjacent to each other in the ordering (Leth-Steensen & Marley, 2000). In the experimental phase, the participants underwent *one of three* randomly assigned secondary task experimental conditions while they responded to the comparison probes; each participant also partook in the control condition, either before or after the experimental condition (i.e., in a counter-balanced fashion). In the control condition, the participants responded to the "Taller ?" or "Shorter?" probes by indicating which of two presented names, previously learned from an ordered series of names, was the correct choice. The participants were also required to respond to either the "Taller ?" or "Shorter?" probe in the experimental condition (in the same manner as in the control condition) but performed this task concurrently with a working memory dual task. The three randomly assigned experimental conditions differed in the type of working memory resources involved: articulatory suppression, visual-spatial tapping, and random letter generation.

This present study aimed to extend Theberge's (2003) findings specifically by using just-learned stimuli; symbolic comparisons with stimuli that have just been learned were assumed to tap into participants' working memory resources more effectively than with stimuli occurring naturally, thus giving the opportunity to go a step further in the exploration of which aspect of working memory is involved in symbolic comparisons.

Predictions

First, all three of the symbolic comparison effects of distance, end, and semantic congruity were expected to occur. Furthermore, it was assumed that overall symbolic comparisons within the three secondary task conditions would be significantly slowed down and accuracy would decrease in comparison to the control condition. It was also assumed that some interactions would occur between the three symbolic comparison effects and the secondary task conditions in terms of both response time and accuracy since participants would have to presumably use their working memory to a greater extent than in Theberge's (2003) study after learning the associated names and heights of the fictional individuals. In Theberge's (2003) study, the participants had a pre-existing notion of which animal was larger or smaller than another before the experiment even began, and hence, the retrieval of magnitude information in that study could be regarded as being more highly automaticized.

Of the three working memory tasks, the random generation task was hypothesized to interact with the symbolic comparison effects more than articulatory suppression or visuo-spatial tapping. The proposed greater effect of the random generation task in comparison to the other two tasks is because random generation is theorized to require a greater amount of planning and attention. The extent to which the other two tasks interfere with the symbolic comparison effect would help to determine the degree to which either verbal-based or visuo-spatial-based resources are required by the comparison of recently learned symbolic items.

METHOD

Participants

Eighty-eight psychology participants were recruited to participate in the experiment (45 male and 43 female). Eighty-five of these participants were introductory psychology students at Carleton University that took part in the experiment for extra course credit; the other 3 participants were volunteers.

Stimuli and Apparatus

The experiment was programmed with Micro Experimental Laboratory (MEL V.2.0). On all trials, the instruction for comparison (i.e., "Taller?" or "Shorter?") was first presented at the top of the screen. During each comparison trial, participants were presented with pairs of three-letter names, presented side by side (horizontally) in the centre of a computer screen. The font size for the stimuli (as determined by MEL System48 font) was somewhat larger than a regular DOS font size.

Each name was a label that stood for an imaginary "person". There were six "people" in this experiment, all of who were assumed to differ in height. Two different orderings of these names were derived a priori for use in this experiment. The two orderings were (from tallest to shortest): Pat > Bob > Ted > Dan > Jim > Mel and Dan > Ted > Jim > Pat > Mel > Bob, and were assigned randomly between participants. Participants made responses by pressing either the "z" or "x" response keys on the computer keyboard, each covered with a circular yellow sticker.

In the experimental condition, participants had to concurrently perform one of the three secondary tasks. In the articulatory suppression task, the participants were required to repeat out loud the letter sequence "F, J, N, W". During the spatial tapping task, the participants had to use their right index finger to tap four keys (1, 2, 6, and 8) on the numerical keypad (of the same keyboard used in the symbolic comparison task). Each of these keys was covered with a green circular sticker, and individually numbered 1, 2, 3, and 4, respectively, indicating the order that the participants had to press them. During the random letter generation task, the participants were required to repeat out loud the letters “F”, “J”, “N”, and “W” in a random order.

Procedure

The experiment consisted of a learning phase and an experimental phase. In both phases, each participant was given instructions (written and orally) but not given any additional strategies for doing the tasks, or any hint of what was expected in the results. The written instructions are shown in Appendix A.

Learning Phase

In the learning phase, each participant was presented with each of the five comparison pairs consisting of the stimuli that were adjacent to each other in the relevant ordering (i.e., the Split 1 pairs). Each comparison trial began with a blank screen for 1000 ms. Before the pair of names appeared, the relevant comparative instruction (“Taller?” or “Shorter?”) was displayed on the computer screen for 1000 ms just above a plus sign, which acted as a temporary fixation point for the location of the name stimuli. On each comparison trial, the participants were presented with two of the six names, side by side in the center of the computer screen, and had to choose which “person” (the name presented on the left or the name presented on the right) who was the shorter (or, respectively, the taller) of the two individuals in the comparison pair. Participants made their response by pressing the left response key (formally the “z” key on the keyboard) if they felt that the "person" on the left was the correct response, and pressing the right response key (formally the "x" key on the keyboard) if they feel that the "person" on the right was the correct response. The participants only used their index and middle fingers of their left hand to make their responses and kept these fingers on the keys at the beginning of each trial.

The participants were not expected to know the relative heights of the stimuli at the very start of the experiment (hence, they were initially guessing which name was either shorter or taller than the other). After each response, feedback was immediately provided on the computer screen. This feedback indicated whether the response had been correct or incorrect and also provided the correct name for that comparison trial, in either case. The participant was able to examine this feedback for as long as they would like and initiated the next trial with a press of the space bar on the keyboard.

Both the comparative instruction and the pair of names remained on the screen throughout the trial (including during the presentation of the feedback). The Split 1 comparison pairs were shown repeatedly throughout the learning phase of the experiment and the feedback helped the participants eventually learn which person is the "taller" and which person is "shorter" in each pair.

The learning phase took anywhere from 10 to 40 minutes to perform and was completed when the participant was correct for a full block of 20 learning trials (i.e., five pairs with each of the two instructions in each of the two left-right spatial presentation of the pairs). The experimenter was in the room at the beginning of the learning phase (for approximately two or three trials) in case the participants had any pressing questions regarding the learning phase. The experimental program indicated when the learning phase had been completed; when this occurred, the participants indicated this to the experimenter, who was just outside the lab. Participants who did not reach a learning criterion in 10 blocks of learning trials had a shorter experimental period, and their results were not included in the analysis. As a result of not passing the learning criterion, data from 28 participants were excluded from the data analyses (leaving a total of 60 participants).

Experimental Phase

The procedure for the second phase of the experiment was similar to the learning phase, except for a few differences. First, all of the 15 possible pairs of names were now included in the stimulus set. Second, no feedback was provided (hence, each new trial was now initiated immediately after each response). Finally, during this part of the experiment, there was four experimental blocks that consisted of 120 randomly presented comparison trials: 15 (pairs) x 2 (instruction types) x 2 (left-right stimulus presentation order) x 2. The participants were also given an optional rest period between each block (as prompted by the computer program).

In two out of the four blocks, participants were required to perform one of three possible working memory secondary tasks while they made the comparison responses. Participants were randomly assigned to the three working memory conditions, resulting in 20 participants per working memory condition.

Instructions were provided to the participants as to the nature of the secondary task, as well as a small set of practice trials. For the primary symbolic comparison task, participants were asked to be accurate with each decision without taking too much time to respond. The experimenter tape-recorded the secondary task performance of 30 participants (10 participants in each secondary task condition) for a total of 1.5 min. Specifically, three 30-s intervals were recorded at the approximate beginning, middle, and end of a block. Half of the participants were recorded on the first secondary task block, and half were recorded on the second secondary task block. At the end of the experiment, an additional 1.5 min of secondary task performance was recorded while it was being performed alone (i.e., without the symbolic comparison task). The experimenter was in the experimental room throughout the entire experimental phase to make sure that participants were always keeping "on track" with both tasks (to maintain consistent experimental conditions the experimenter also remained in the room during the two control blocks of trials).

Secondary Tasks

Articulatory Suppression Secondary Task

In these two blocks of comparison trials, participants were required to simultaneously perform a letter-vocalizing task in which they had to repeat out loud the letter sequence "F, J, N, W". Participants were asked to repeat these letters as quickly as possible, however at a constant pace.

Visual-Spatial Tapping Task

For two blocks of comparison trials, participants were required to simultaneously perform a task in which they had to use their right index finger to tap four marked keys (in a counter clock-wise order) on the numerical keypad of the keyboard. Participants were asked to try to tap these keys as quickly as possible, however at a constant pace.

Random Generation Task

For two blocks of comparison trials, participants were required to simultaneously perform a task that involved repeating out loud the letters "F", "J", "N", and "W" in a random order. The participants were asked to try to generate the random letters as quickly as possible, however at a constant pace.

RESULTS

Results of the Learning Phase

The data from the learning phase, in terms of the number of blocks of training trials that it took each of the participants to reach the learning criterion are as follows. It took an average of 6.45 blocks for the articulatory suppression group to learn the associative name-height relations, it took the visuo-spatial tapping group an average of 5.70 blocks to learn them; and it took the random letter generation group an average of 5.35 blocks to learn them.

Primary Task Results

Mean response times and proportion correct for the symbolic comparisons will be the dependent measures in the following analyses. The distance effect and its interaction with the respective working memory condition will be analyzed by means of a repeated measures ANOVA for each group, where level of split, instruction type, and secondary task condition are the relevant independent variables. The end and semantic congruity effects, and their interaction with the respective working memory condition will be analyzed separately at each split.

Although it is the conventional degrees of freedom that are provided for all of the statistical results reported here, all of the corresponding statistical tests were actually based on significance levels that were determined by the (more conservative)

Greenhouse-Geisser epsilon adjusted degrees of freedom. Only correct response times between 200 and 6000 ms were used in the response time analyses. The arcsine transformation of the proportion correct measure was used in the accuracy analyses. Also, all participant mean measures involving Splits 1, 2, 3, 4, and 5 were based on samples of 80, 64, 48, 32, and 16 observations per participant, respectively. The results of these tests are summarized in Tables 1 to 4. The accompanying response time and percent correct data are given in Figures 1 to 7.

Response Times

Articulatory Suppression

The Distance Effect

The repeated measures ANOVA included level of split (1, 2, 3, 4, and 5), instruction type ("Shorter?" and "Taller"), and secondary task condition (control and dual task) as the relevant independent variables. The main effect of the articulatory suppression condition on response time was significant, F (1, 19) = 8.46, MSE = 820674, p < .01 (see Figure 1A). The presence of a substantial distance effect was marked in this analysis by a significant overall main effect of split, F(4, 76) = 18.94; MSE = 209246, p < .001, that was also accompanied by a reliable linear trend F(1, 19) = 71.30, p < .001. The marginal response time means for each (increasing) level of split that characterize this distance effect were 1889, 1778, 1686, 1575, and 1470 ms, respectively. The interaction of secondary task condition and split was not significant, F(4, 76) = .113, MSE = 134644, p > .50.

End and Semantic Congruity Effects

A set of ANOVAs was performed separately for each of the splits of 1, 2, 3, and 4 that included secondary task condition (control and dual task), pair (e.g., [1, 2], [2, 3], [3, 4], [4, 5], and [5, 6] for a split of 1), and instruction type ("Shorter?" and "Taller?") as the relevant independent variables. There was a significant main effect of the articulatory suppression condition at a split of 1, $F(1, 19) = 10.67$, $MSE = 679449$, $p < .01$; at a split 2, $F(1, 19) = 7.22$, $MSE = 774665$, $p < .05$; at a split of 3, $F(1, 19) = 7.42$, $MSE = 473136$, $p < .05$; and marginally at split of 4, $F(1, 19) = 4.03$, $MSE = 665425$, $p < .06$. The main effect of instruction type in each of these analyses was significant only for a split of 1, $F(1, 19) = 5.26$, $MSE = 49974$, $p < .05$ (see Figure 2). There was also a Condition x Instruction interaction at Split 3, $F(1, 19) = 5.19$, $MSE = 72282$, $p < .05$. The effect of the secondary task condition, however, did not interact with any other factor at any split. The presence of inverted U-shaped end effects was marked in these analyses by significant main effects of pair that were accompanied by reliable quadratic trends at each of a split of 1, $F(4, 76) = 22.64$, $MSE = 471774$ and $F(1, 19) = 55.48$, $MSE = 535027$, respectively, ps $< .001$; a split of 2, $F(3, 57) = 15.49$, $MSE = 590663$ and $F(1, 19) = 43.57$, $MSE = 345526$, respectively, ps $< .001$; and at a split of 3, $F(2, 38) = 8.77$, $MSE = 365068$, $p < .001$ and $F(1, 19) = 6.92$, $MSE = 377617$, $p < .02$, respectively.

The presence of semantic congruity effects was marked by a marginally significant Pair X Instruction Type interaction at a split of 1, $F(4, 76) = 2.40$, $MSE = 109173$, $p < .08$, accompanied by a significant linear trend interaction of the pair effect with instruction type (i.e., the cross-over effect), $F(1, 19) = 4.85$, $MSE = 239750$, $p < .05$; a significant interaction at a split of 2, $F(3, 57) = 4.06$, $MSE = 152558$, $p < .05$, accompanied by a significant linear trend interaction, $F(1, 19) = 5.30$, $MSE = 107627$, $p <$

.05; a significant interaction at a split of 3, $F(2, 38) = 3.92$, $MSE = 162940$, $p < .05$, accompanied by a significant linear trend interaction, $F(1, 19) = 7.51$, $MSE = 146984$, $p < .05$; and a significant interaction at a split of 4, $F(1, 19) = 22.38$, $MSE = 77461$, $p < .001$.

Visuo-Spatial Tapping

The Distance Effect

The repeated measures ANOVA included level of split (1, 2, 3, 4, and 5), instruction type ("Shorter?" and "Taller"), and secondary task condition (control and dual task) as the relevant independent variables. Although, the response times in Figure 1A were slower in the visuo-spatial tapping condition, there was only a marginally significant effect of the visuo-spatial tapping condition on response times, $F(1, 19) = 3.23$, $MSE = 949674$, $p < .10$. The overall main effect of instruction type was significant, $F(1, 19) = 18.82$, $MSE = 802136$, $p < .001$, reflecting the fact that the mean response times were 89 ms faster overall with the instruction "Taller?" than with the instruction "Shorter?" (i.e., the lexical markedness effect). The presence of a substantial distance effect was marked in this analysis by a significant overall main effect of split, $F(4, 76) = 39.30$; $MSE = 173638$, $p < .001$, that was also accompanied by a reliable linear trend $F(1, 19) = 69.09$, $p < .001$. The marginal response time means for each (increasing) level of split that characterize this distance effect were 1823, 1682, 1558, 1423, and 1268 ms, respectively. As well, the interaction of secondary task condition and split was not significant, $F(4, 76) = .43$, $MSE = 45542$, $p > .50$.

End and Semantic Congruity Effects

A set of ANOVAs was performed separately for each of the splits of 1, 2, 3, and 4 that included secondary task condition (control and dual-task), pair (e.g., [1, 2], [2, 3], [3, 4], [4, 5], and [5, 6] for a split of 1), and instruction type ("Shorter?" and "Taller?") as the relevant independent variables. There was a significant main effect of the visuo-tapping condition only at a split of 3, $F(1, 19) = 4.70$, $MSE = 643535$, $p < .05$, and a marginal effect at a split of 4, $F(1, 19) = 3.86$, $MSE = 428376$, $p < .07$. The effect of secondary task condition did not interact with any of the other factors at any split.

The main effect of instruction type was significant at a split of 1, $F(1, 19) = 5.72$, $MSE = 97246$, $p < .05$; a split of 2, $F(1, 19) = 8.07$, $MSE = 101833$, $p < .01$; and a split of 3, $F(1, 19) = 5.78$, $MSE = 112534$, $p < .05$ (see Figure 3). The presence of inverted U-shaped end effects was marked in these analyses by significant main effects of pair that were accompanied by reliable quadratic trends at each of a split of 1, $F(4, 76) = 22.39$, $MSE = 720705$ and $F(1, 19) = 54.590$, $MSE = 632482$, respectively, $ps < .001$; and a split of 2, $F(3, 57) = 19.60$, $MSE = 518827$ and $F(1, 19) = 27.50$, $MSE = 418739$, respectively, $ps < .001$. The end effect was significant at a split of 3, $F(2, 38) = 10.39$, $MSE = 4988524$, $p < .001$, however the quadratic trend was only marginally significant (due to the fact that although the responses to the tallest Split 3 pair were fairly fast, the response to the shortest Split 3 pair were not), $F(1, 19) = 3.58$, $MSE = 564184$, $p < .08$. There was also a significant pair difference at the split of 4, $F(1, 19) = 18.22$, $MSE = 174074$, $p < .001$. The presence of semantic congruity effects was marked by a significant Pair X Instruction Type interaction at a split of 1, $F(4, 76) = 10.42$, $MSE = 130859$, $p < .001$, accompanied by a significant linear trend interaction of the pair effect with instruction type, $F(1, 19) = 31.68$, $MSE = 99846$, $p < .001$; and a significant interaction at a split of 2, $F(3, 57) = 4.38$, $MSE = 165014.04$, $p < .05$, accompanied by a significant linear trend interaction, $F(1, 19) = 10.07$, $MSE = 134591$, $p < .01$. These interactions did not occur at Split 3, $F(2, 38) = .65$, $MSE = 129706$, and $F(1, 19) = 1.07$, $MSE = 126640$, and the interaction and linear trend interactions, respectively, and were marginally significant at Split 4, $F(1, 19) = 3.91$, $MSE = 107229$, $p < .07$.

Random Generation

The Distance Effect

The repeated measures ANOVA included level of split (1, 2, 3, 4, and 5), instruction type ("Shorter?" and "Taller"), and secondary task condition (control and dual task) as the relevant independent variables. The main effect of the random generation of letters condition on response time was significant, $F(1, 19) = 83.76$, $MSE = 763702$, $p < .001$ (see Figure 1A). The presence of a substantial distance effect was marked in this analysis by a significant overall main effect of split, $F(4, 76) = 14.74$; $MSE = 196795$, $p < .001$, that was also accompanied by a reliable linear trend $F(1, 19) = 20.21$, $p < .001$. The marginal response time means for each (increasing) level of split that characterize this distance effect were 2382, 2258, 2220, 1965, and 1949 ms, respectively. The interaction of secondary task condition and split was not significant, $F(4, 76) = 1.48$, $MSE = 333074$, $p > .20$.

End and Semantic Congruity Effects

A set of ANOVAs was performed separately for each of the splits of 1, 2, 3, and 4 that included secondary task condition (control and dual-task), pair (e.g., [1, 2], [2, 3], [3, 4], [4, 5], and [5, 6] for a split of 1), and instruction type ("Shorter?" and "Taller?") as the relevant independent variables. There was a significant main effect of random generation condition at a split of 1, $F(1, 19) = 55.10$, $MSE = 1024752$, $p < .001$; split of 2, $F(1, 19) = 66.91$, $MSE = 675385$, $p < .001$; split of 3, $F(1, 19) = 85.97$, $MSE = 549846$, $p < .001$;

and split of 4, $F(1, 19) = 50.97$, $MSE = 402044$, $p < .001$. The effect of the secondary task condition, however, did not interact with any other factors at any split.

The main effect of instruction type in each of these analyses was significant only for a split of 1, $F(1, 19) = 5.49$, $MSE = 417097$, $p < .05$. The presence of inverted U-shaped end effects was marked in these analyses by significant main effects of pair that were accompanied by reliable quadratic trends at each of a split of 1, $F(4, 76) = 15.00$, $MSE = 822788$ and $F(1, 19) = 59.16$, $MSE = 442939$, respectively, $ps < .001$; a split of 2, $F(3, 57) = 13.62$, $MSE = 687146$ and $F(1, 19) = 38.80$, $MSE = 322134$, respectively, $ps < .001$; and at a split of 3, $F(2, 38) = 9.01$, $MSE = 587666$, $p < .001$ (the quadratic trend for Split 3 was not significant). The presence of semantic congruity effects was marked by a significant Pair X Instruction Type interaction that only occurred at a split of 3, $F(2, 38) = 3.92$, $MSE = 260828$, $p < .05$, but that was accompanied by a significant linear trend interaction of the pair effect with instruction type, $F(1, 19) = 11.10$, $MSE = 137678$, $p < .01$.

Accuracy

Articulatory Suppression

The Distance Effect

The repeated measures ANOVA included level of split (1, 2, 3, 4, and 5), instruction type ("Shorter?" and "Taller"), and secondary task condition (control and dual task) as the relevant independent variables. The main effect of the articulatory suppression condition on accuracy (i.e., in terms of arcsine-transformed proportion correct) was significant, $F = 5.35$, $MSE = .016$, $p < .05$ (see Figure 1B). The presence of a substantial distance effect was marked in this analysis by a significant overall main effect of split, F

(4, 76) = 8.23, *MSE* = .086, $p < .001$, that was also accompanied by a reliable linear trend, $F(1, 19) = 18.33$, *MSE* = .085, $p < .001$. The marginal mean proportion correct for each (increasing) level of split that characterize this distance effect were .877, .883, .897, .900, and .943, respectively. The effect of the secondary task condition did not interact with split, $F(4, 76) = 1.17$, *MSE* = .019, $p > .25$.

End and Semantic Congruity Effects

A set of ANOVAs was performed separately for each of the splits of 1, 2, 3, and 4 that included secondary task condition (control and dual-task), pair (e.g., [1, 2], [2, 3], [3, 4], [4, 5], and [5, 6] for a split of 1), and instruction type ("Shorter?" and "Taller?") as the relevant independent variables. There was a significant main effect of articulatory suppression condition only at a split of 4, $F(1, 19) = 4.53$, *MSE* = .026, $p < .05$. Although the main effects of pair were only marginally significant overall, the presence of U-shaped end effects was marked in these analyses by reliable quadratic trends at a split of 1, $F(4, 76) = 2.56$, *MSE* = .437, $p < .08$ and $F(1, 19) = 13.26$, *MSE* = .129, $p < .01$, respectively; and at a split of 2 $F(3, 57) = 1.51$, *MSE* = .312, $p < .08$ and $F(1, 19) = 4.81$, *MSE* = .160, $p < .05$, respectively (see Figure 5). The presence of semantic congruity effects was marked by a significant Pair X instruction Type interaction only at a split 4, $F(1, 19) = 5.70$, *MSE* = .021, $p < .05$. However, although there was no main effect of instruction, there was a significant Condition x Instruction interaction at a split 4, $F(1, 19) = 5.370$, *MSE* = .036, $p < .05$; the mean percent accuracy for the "Shorter?" and "Taller?" instructions were .926 and .908, respectively in the control condition and .865 and .915 in the articulatory suppression condition, respectively. The effect of the secondary task did not interact with any other factor at any split.

Visuo-Spatial Tapping

The Distance Effect

The repeated measures ANOVA included level of split (1, 2, 3, 4, and 5), instruction type ("Shorter?" and "Taller"), and secondary task condition (control and dual task) as the relevant independent variables. The main effect of the visuo-spatial tapping condition on accuracy was significant, $F(1, 19) = 12.22$, $MSE = .043$, $p < .01$ (see Figure 1B). The presence of a substantial distance effect was marked in this analysis by a significant overall main effect of split, $F(4, 76) = 18.86$, $MSE = .046$, $p < .001$, that was also accompanied by a reliable linear trend, $F(1, 19) = 2.171$, $MSE = .055$, $p < .001$. The marginal mean proportions correct for each (increasing) level of split that characterize this distance effect were .876, .924, .934, .938, and .956, respectively. The effect of the secondary task did not interact with split, $F(4, 76) = .65$, $MSE = .012$, $p > .50$.

End and Semantic Congruity Effects

A set of ANOVAs was performed separately for each of the splits of 1, 2, 3, and 4 that included secondary task condition (control and dual task), pair (e.g., [1, 2], [2, 3], [3, 4], [4, 5], and [5, 6] for a split of 1), and instruction type ("Shorter?" and "Taller?") as the relevant independent variables. There was a significant main effect of visuo-spatial tapping condition at a split of 1, $F(1, 19) = 8.83$, $MSE = .096$, $p < .01$; a split of 2, $F(1, 19) = 7.56$, $MSE = .051$, $p < .05$; a split of 3, $F(1, 19) = 10.38$, $MSE = .033$, $p < .01$; and at a split 4, $F(1, 19) = 4.84$, $MSE = .030$, $p < .05$. At a split of 1, the presence of a U-shaped end effect was marked only by a reliable quadratic trend, $F(1, 19) = 5.87$, $MSE = .222$, $p < .05$; at a split of 2, there was both a reliable main effect of pair and quadratic trend, $F(3, 57) = 3.51$, $MSE = .146$ and $F(1, 19) = 4.23$, $MSE = .094$, respectively, $ps \leq$

.05 (see Figure 6). Although there were no significant main effects of instruction type or Instruction x Pair effects at any of the splits, there was a marginally significant Secondary Task Condition x Instruction Type x Pair interaction at a split of 4, $F(1, 19) = 4.17$, $MSE = .008$, $p < .06$.

Random Generation

The Distance Effect

The repeated measures ANOVA included level of split (1, 2, 3, 4, and 5), instruction type ("Shorter?" and "Taller"), and secondary task condition (control and dual task) as the relevant independent variables. The main effect of the random generation condition on accuracy was significant, $F(1, 19) = 9.99$, $MSE = .064$, $p < .01$ (see Figure 1B). The presence of a substantial distance effect was marked in this analysis by a significant overall main effect of split, $F(4, 76) = 13.15$, $MSE = .079$, $p < .001$ that was also accompanied by a reliable linear trend $F(1, 19) = 16.56$, $p < .001$. The marginal mean proportion correct for each (increasing) level of split that characterized this distance effect were .850, .885, .931, .938, and .933. Although the Secondary Task Condition x Split interaction was not significant $F(4, 76) = 1.68$, $MSE = .025$, $p > .15$, this interaction was significant, $F(3, 57) = 3.01$, $MSE = .014$, $p < .05$ when Split 5 was omitted from the analysis (i.e., for Splits 1 to 4 inclusive).

End and Semantic Congruity Effects

A set of ANOVAs was performed separately for each of the splits of 1, 2, 3, and 4 that included secondary task condition (control and dual task), pair (e.g., [1, 2], [2, 3], [3, 4], [4, 5], and [5, 6] for a split of 1), and instruction type ("Shorter?" and "Taller?") as the relevant independent variables. The main effect of random generation condition was significant at a split of 1, $F(1, 19) = 19.18$, $MSE = .097$, $p < .001$, and marginally significant at a Split 3, $F(1, 19) = 4.10$, $MSE = .058$, $p < .06$. There were no significant end effects or semantic congruity effects at any split (see Figure 7). Although the effect of secondary task condition did not interact with either pair or instruction type by themselves, there was a marginally significant Secondary Task Condition x Instruction Type x Pair interaction at a split of 2 $F(3, 57) = 2.76$, $MSE = .036$, $p < .06$.

Working Memory Secondary Task Performance

Paired two-tailed, t-tests were conducted to compare the participants' performance on each of the working memory tasks when they were performed as a secondary task with when they were performed alone. These measures were taken and this analysis was performed for the last 10 participants in each working memory secondary task condition group. In both the articulatory suppression and random generation tasks, the number of letters uttered in a 1.5-min time span was recorded. In the visuo-spatial tapping condition, the number of taps made on the numerical keyboard in a 1.5-min time span was recorded.

The results indicated that the rate at which the participants performed the working memory task alone was significantly faster than when performed along with the symbolic comparison tasks: $t(9) = 2.77$, $t(9) = 2.90$, $t(9) = 5.48$ (for comparisons of articulatory suppression, visuo-spatial tapping, and random letter generation performances, respectively), $ps < .05$. In the articulatory suppression condition, the mean number of letters uttered alone was 233 per 1.5 min, whereas when it was performed as a secondary task, the mean was 197.5 per 1.5 minutes. In the visuo-spatial tapping condition, the mean number of keys tapped alone was 281.6 per 1.5 min, whereas when it was performed as a secondary task, the mean was 248.2 per 1.5 min. In the random generation condition, the mean number of letters uttered alone was 142.4 per 1.5 min, whereas when it was performed as a secondary task, the mean was 92.3 per 1.5 min. A one-way single factor ANOVA indicated that when the working memory tasks were performed with the symbolic comparison tasks, the rate of performance differed between conditions, $F(2, 27) = 48.18$, $MSE = 1312$, $p < .001$.

DISCUSSION

The purpose of this study was to explore the role of working memory resources on symbolic comparisons. To that end, the effects of three secondary tasks that involved the three components of working memory (respectively) on relative symbolic reasoning were investigated. The results showed an overall effect of the articulatory suppression on response times, as well as on accuracy. Upon inspection of separate analyses at each split, however, there was a significant decrease in accuracy only for the Split 4 stimuli. There was also an overall effect of visuo-spatial tapping on accuracy, but this task affected response times in only a marginally significant fashion (except at Split 3). The random letter generation condition had substantial overall effects on both response times and accuracy. Upon inspection of separate split analyses in terms of proportion correct however, there was a significant effect of this task only at Split 1 and (marginally) at Split 3.

These findings suggest that overall, the random generation condition was certainly much more difficult for the participants to perform than the other two secondary task conditions. Although both articulatory suppression and random generation impeded responses in terms of both response time and accuracy, the effect of random generation was much larger than that of articulatory suppression which supports the contention that these two tasks do indeed individually encompass different resources, even though they both involved uttering the same letters. However, even in the control condition, the participants in the random generation condition were much slower at responding to the symbolic comparisons than the participants in the articulatory suppression condition (see Figure 1A).

The Distance Effect

The distance effect was robust for all conditions in terms of both response time and accuracy measures. That is, response times decreased and accuracy increased as the distance between the stimuli increased. However, none of the secondary task conditions interacted with the distance effect in terms of response times. The random letter generation condition, nevertheless, did significantly interact with the distance effect for participants' accuracy measures. However, this interaction was not significant until the data from the fifth split was removed, and therefore it was based on Splits 1 to 4. Removing Split 5 seemed necessary because there was an unexpected decrease in accuracy at this level (that was accompanied by an increase in RT; see Figure 1). The interaction at Splits 1 to 4 clearly is evident in Figure 1B, where there is a fairly dramatic difference in accuracy in the secondary task condition compared to the control condition at Split 1, relative to the differences between the control and secondary task condition accuracy measures at Splits 2 to 4.

The End Effect

The end effect was quite marked in this study, especially in terms of response times, in that responses in all conditions were generally faster when the participants were presented with either the shortest or tallest stimuli. This trend was consistently significant at the smaller splits. The participants in the articulatory suppression and visuo-spatial tapping groups, specifically at Splits 1 and 2, also showed an end effect with respect to the amount of errors they made. The end effect was not as robust in terms of accuracy for the participants in the random letter generation group, in that accuracy was, statistically, quite similar regardless of which pair they were presented. In all cases, there were no interactions between the end effect and secondary task condition. This finding suggests that for symbolic comparisons with learned orderings, secondary working memory tasks generally do not either enhance or diminish the end effect.

The Semantic Congruity Effect

The semantic congruity effect was evident in terms of response times, although mainly for the participants that were in the articulatory suppression and visuo-spatial tapping groups. The participants that were in the random letter generation group demonstrated the semantic congruity effect only for Split 3 stimuli. The semantic congruity effect generally did not show up in the accuracy results for any condition, in that the participants were almost never significantly more accurate for stimuli that had the same polarity as the instructions. In all cases, there was no interaction between the semantic congruity effect and secondary task condition. One caveat is that there was lack of a semantic congruity response time effect in the control condition in the random generation group (which made the interaction with secondary task condition unlikely anyway). These findings suggest that in symbolic comparison paradigms with learned orderings, phonological and visuo-spatial secondary tasks generally do not either enhance or diminish the semantic congruity effect.

Lexical Markedness

The lexical markedness effect was also apparent in this study. Participants that were in the visuo-spatial tapping group exhibited this effect quite clearly, in that the "Taller?" instruction elicited significantly faster responses than the "Shorter?" instruction. In the other two groups, the main effect of instruction type was apparent only in the Split 1 data where responses were approximately 51 ms and 151 ms faster for the articulatory suppression and random letter generation conditions, respectively, when the participants were asked to choose the "Taller?" of the two stimuli.

Secondary Task Performance

It is evident from the results involving secondary task performance measures that the participants were slower on the secondary tasks when they were performed in conjunction with the symbolic comparisons, than when they were performed alone. This is to be expected however, because of the nature of dual tasks. That the response times of the symbolic comparisons were significantly slowed and that accuracy significantly declined in comparison to the control condition under the dual-task conditions should eliminate any concern that the secondary task was not being performed, or that there was no effort by the participants to perform the secondary tasks (See Vandierendonck et al., 2004 for a similar results and arguments, in terms of dual-task trade offs). Also, consistent with this supposition is the fact that the three working memory tasks were generally performed at different rates from each other, which is also to be expected given the qualitative difficulty differences between each type of secondary task. Furthermore, although there were significant single- versus dual- task differences, the performance rates were still fairly substantial in the dual-task conditions ($M = 2.2$ utterances per second in the dual-task condition in the articulatory suppression condition; $M = 2.7$ taps per second in the dual-task condition in the visuo-spatial tapping condition; and $M = 1.0$ utterance per second in the dual-task condition in the random generation condition).

GENERAL DISCUSSION

The sections that follow discuss in greater detail the implications of these findings with respect to the nature of the representations presumed to be used in the symbolic comparison process and to the models discussed in the Introduction.

Overall Secondary Task Interference Effects

As mentioned earlier, all three secondary tasks had some detrimental overall effect on symbolic comparisons in terms of either response times or accuracy. The articulatory suppression condition slowed responses overall, and participants made a significantly larger amount of errors (particularly on the Split 4 stimuli) while performing this secondary task. Since the articulatory suppression condition did not interact with any of the symbolic comparison effects, however, it suggests that the phonological loop is not involved in the decision stage of the symbolic comparison process. Given the strictly additive nature of the effect of the articulatory suppression condition on symbolic comparisons performance measures (e.g., they increased or decreased without altering the size of the symbolic comparison effects themselves), one could speculate that the phonological loop is involved in non-decisional (or residual) processing aspects of the symbolic comparison task, such as rehearsal of the instruction or the encoding of the stimuli (although in both cases note that neither the instructional nor the stimulus representations used in the comparison process itself would have been degraded by such interference).

The visuo-spatial tapping condition did not significantly slow the responses, but did decrease accuracy in an additive manner. Since the visuo-spatial tapping condition also did not interact with any of the symbolic comparison effects, however, it suggests that the visuo-spatial sketchpad is also not involved in the decision stage of the symbolic comparison process. The additive effect of the visuo-spatial tapping condition on symbolic comparison in terms of accuracy suggests that the visuo-spatial tapping also interferes with non-decisional processing, a likely candidate being the elicitation of the motor responses (i.e., after the decision has been made).

In the random generation condition, responses were significantly slowed and less accurate in comparison to responses in the control condition. Not surprisingly, this indicates that central executive interference has a substantial overall effect on performance in this task.

Secondary Task Interference and the Symbolic Comparison Process

The Distance Effect

The distance effect was robust in this study. Although all three secondary tasks significantly increased the overall latencies, this increase occurred in an additive manner with split. The only evidence that a secondary task interfered with the symbolic comparison process itself occurred in the accuracy data at Splits 1 to 4 for the random letter generation task (where the distance effect was significantly enhanced under secondary task conditions). This finding suggests that general executive-type attentional resources are involved in comparing the relative magnitudes of the symbolic items in this experiment. Note that it was a differential impairment of the comparisons involving the more difficult Split 1 pairs that contributed heavily to this interaction (see Table 4 and Figure 1B).

In terms of latency, these results coincide with those of De Rammelaere and Vandierendonck (2003), who also found an additive effect of random interval generation on the distance effect. Participants in their study were required to indicate, as fast and as accurately as possible, whether the numbers 1, 2, 3, 4, 6, 7, 8, or 9 were either smaller or larger than a standard of 5 while performing a random interval generation task. In terms of accuracy, however, the present results do not fully agree with those of De Rammelaere and Vandierendonck (2003), as they did not find any interaction between the RIG load and distance.

Nonetheless, these results do suggest that the symbolic comparison process does seem, at the very least, to require the use of central executive resources (i.e., that it is not completely automatic). Given this fact, the part that is most likely not fully automated is the retrieval of magnitude information. This assertion comes from the contention that the Split 1 stimuli are less automatic in terms of retrieval than stimuli that are, for example, at the opposite ends of the continuum (which may involve much more automatic responses).

However, since only the distance effect in accuracy was enhanced, and not in response time, it is not clear whether this effect occurred due to the fact that the symbolic comparison process itself (and especially the more difficult Split 1 comparisons) was degraded by having to share central executive resources with the random letter generation task or because such a resource decrement simply led to a speed-accuracy-type trade-off within the decision process itself (which would have then attenuated any potential enhancements of the response time distance effects and exaggerated the distance effect in accuracy; Poltrock, 1989). For example, in relation to Poltrock's (1989) random walk model, the participants' random walk may have stopped before the decision boundary was reached under the random generation condition (particularly at Split 1). This could be because the participants' attention resources were taken up primarily by the random generation task, and therefore, they made a motor response before they made an actual decision, resulting in more errors for the more difficult to decipher Split 1 stimuli. If the random walk is assumed to be an accurate description of the symbolic comparison process, the results suggest that the random walk requires general attention resources.

Although the random generation condition did not interact with the distance effect in terms of response times, response times were dramatically higher under this condition, compared to the other two working memory conditions, which was consistent across all splits. This finding could also be readily explained with the random walk model. That is, the bias value (again, a response bias reduces the time to reach one decision but increases the decision time for a decision going the opposite direction) at all splits was probably closer to zero under the random generation condition, perhaps because the participants were not able to process the stimuli at full capacity. The lack of being able to process the stimuli fully would in turn result in no bias for any of the response choices; this would consequently lead to a longer time to reach a decision. Note that a reduction in bias would also serve to attenuate end effects (which did occur in the random generation condition accuracy measures).

The End Effect

The end effect was clearly evident for two out of the three groups this study, particularly in terms of response time. However, it was not differentially affected by single- and dual- task conditions, providing further evidence that neither phonological loop not visuo-spatial sketchpad resources are used in the symbolic comparison process itself. It would have been telling, however, if the visuo-spatial tapping condition had interacted with the end effect. The interaction would have implicated the use of visuo-spatial resources in this effect, which would have given support to Woocher's (1976) ends-inward scanning model. Note that this finding is particularly important given the nature of the stimuli used here (i.e., that come from a learned linear ordering which is conducive to being represented as a spatial array).

The Semantic Congruity Effect

The semantic congruity effect was clearly evident for two out of the three groups this study, particularly in terms of response time. However, it was also not differentially affected by single- and dual- task conditions in those groups providing further evidence that neither phonological loop nor visuo-spatial sketchpad resources are used in the symbolic comparison process itself. It is telling that the articulatory suppression condition did not even marginally interact with the semantic congruity effect. If it had, then the implication of verbal-based semantic codes in the symbolic comparison process would have become quite plausible.

The additivity of articulatory suppression is reminiscent of a study by Shaki and Algom (2002). They presented their participants with names of animals printed within pictures of animals for comparative judgments of size based on either the pictures or the names. Stroop congruent stimuli involved the name of an animal embedded a picture of the same animal, and Stroop incongruent stimuli involved the name of an animal embedded in a picture of an animal whose real size is of the opposite polarity (e.g., the name cat embedded within the picture of an elephant). Although there were main effects of both semantic congruity and the Stroop effect, the authors found that Stroop congruity and semantic congruity did not interact, leading instead to additive effects. Since the Stroop-like stimuli for comparison introduced semantic processing at the encoding stage (as asserted by Shaki & Algom, 2002), the additive effects they found suggest that the semantic congruity effect occurs at the decision stage rather than the encoding stage. This reasoning could be applied to the present data as well, since the articulatory suppression condition did not interact with the semantic congruity effect although it did lead to some overall interference.

Implications for Models of the Symbolic Comparison Process

The results of this study do not give much hint, however, as to what resources are required in either the semantic coding or spatial array models (assuming that either of these models accurately depict what is really going on when people engage in symbolic comparisons) as neither the phonological loop nor the visuo-spatial sketchpad resources seem to be used in the decision stage of the symbolic comparison process. The lack of effect of the articulatory suppression condition on the decision process suggests then that the phonological loop is not involved in forming of semantic codes if the semantic coding model is correct in terms of explaining the symbolic comparison process. Likewise, the lack of effect of the visuo-spatial tapping condition on the decision process suggests then that the visuo-spatial sketchpad is not involved in forming spatial arrays if spatial array models are correct in terms of explaining the symbolic comparison process.

However, these results do not rule out this possibility that the representational processes proposed by either the semantic congruity or the spatial array models occur at a more abstract level. In addition, it was rather surprising that neither the phonological loop nor the visuo-spatial sketchpad were implicated in the comparison tasks of these recently learned stimulus associations given that participants often spontaneously self-reported either verbally or spatially rehearsing the order of the names in this kind of experiment.

With respect to the other models discussed in the introduction, the random walk model is supported by default, as it does not invoke any notions of verbal or spatial representations. However, the evidence in the present study suggests that the processing occurring within such a model might not be fully automatic. For the same reasons, the Leth-Steensen and Marley (2000) connectionist model is not contradicted by these results. With respect to the Leth-Steensen and Marley (2000) model, though, an important issue that was not fully addressed by these results concerns the effect of secondary task interference on the semantic congruity effect. While it seems clear that neither phonological nor visuo-spatial interference change the nature of this effect, the effect of the, fairly high-load, central executive interference could not be fully examined given the (surprising) lack of a semantic congruity effect for the random generation group even for comparisons not performed under dual-task conditions. Note that within the Leth-Steensen and Marley model, this effect is deemed to arise due to an enhanced competition between instructional pathways when the stimuli and instructions are incongruent. Further, it could be assumed that the magnitude of such competition on any one trial should depend on the clarity of the context provided by the comparative instruction (i.e., more competition, and more semantic congruity related slowing, for noisy instructional contexts), which itself should likely be diminished by attentional resource-demanding central executive interference. This issue is also quite relevant to the view of the semantic congruity effect proposed by Petrusic (1992), namely that it arises because the rate of each evidence accrual event is slowed when the stimuli and instructions are incongruent. If such differential slowing is occurring then one might expect it to be enhanced by the concurrent performance of a resource-demanding central executive secondary task.

Further Investigations

The results also provide an opportunity to speculate on how the role of working memory on symbolic reasoning might be further investigated given similar yet modified methodology. For instance, given that working mental models require memory, increasing the mental models to two (similar to what was done in the Vandierendonk and De Vooght, 1997 study) in terms of the stimuli might give further insight, particularly under a working memory task. An example of increasing the mental models to two is to increase the number of the names to be remembered to 7, and then making two of the names the same relative height. Therefore, although there are still six different heights, participants will have learned that two of seven "individuals" are the same heights, and will therefore have two mental models of the complete stimuli (i.e, two sets of stimuli).

Furthermore, it is also possible that the load provided by the working memory tasks (especially with respect to articulatory suppression and visuo-spatial tapping) were not large enough to affect the decision process themselves. Perhaps if the phonological loop task had been more involved, say, if the utterances to be repeated rhymed with the names to be remembered (hence, combining the effect of articulatory suppression and phonological similarity), there might then be an obvious effect of this condition as it would presumably involve a greater amount of verbal memory. Similarly, it would be interesting to compare the visuo-spatial tapping condition with a more involved visuo-spatial secondary condition in which, say, the participants have to envision tracing a Z-shape in their minds eye. Furthermore, exploring the random-interval tapping task might also give further insight as to the resources needed in symbolic comparison that is not confounded by the use of phonological processes. Comparing several working memory tasks under the respective sub-components of working memory might tease out the specific resources involved in symbolic comparison, as different tasks should obviously tap into different specific resources.

REFERENCES

Andrade, J. (2001). Introduction. In A. Andrade (Ed.), *Working Memory in Perspective* (3-30). New York, NY: Taylor & Francis Inc.

Baddeley, A. D. (2002). Fractioning the central executive. In D. T. Struss & R. T. Knight (Eds.), Principles of Frontal Lobe Function (pp. 246-260). New York, NY: Oxford University Press.

Baddeley, A. (1998). Recent developments in working memory. *Current Opinion in Neurobiology, 8,* 234-238.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A* (1), 5-28.

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

Baddeley, A., Emslie, H., Kolodny, J., & Duncan, J. (1998). Random generation and the executive control of working memory. *The Quarterly Journal of Experimental Psychology, 51A,* 819-852.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of active maintenance and executive control,* pp. 28-61. New York: Cambridge University Press.

Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgments. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 11, pp. 101-159). New York: Academic Press.

Banks, W. P., Fujii, M., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 435-447.

Birnbaum, M. H., & Jou, J. (1990). A theory of comparative response times and "difference" judgements. *Cognitive Psychology, 22.* 184-210.

Bower, G. H. (1971). Adaptation-level coding of stimuli and serial position effects. In M. H. Appley (Ed.), *Adaptation-level Theory* (pp. 175-201). New York: Academic Press.

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology, 103*, 1131-1136.

Čech, C. G., & Shoben, E. J. (1985). Context effects in symbolic magnitude comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11* (2), 299-315.

Cohen, J. D., McClelland, J. L., & Dunbar, K. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97* (3), 332-361.

Conrad, R. & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55* (4), 429-432.

Copeland, D. E., & Radvansky, G. A. (2004). Working memory and syllogistic reasoning. *The Quarterly Journal of Experimental Psychology, 57* (8), 1437-1457.

De Rammelaere, S., & Vandierendonck, A. (2003). Number comparison under executive dual-task. *Psychologica Belgica, 43* (4), 259-269.

Duyck, W., Vandierendonck, A., & De Vooght, G. (2003). Conditional reasoning with a spatial content requires visuo-spatial working memory. *Thinking and Reasoning, 9* (3), 267-287.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning.* London: Routledge.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning. The psychology of deduction.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Holyoak, K. J. & Mah, W. A. (1981). Semantic congruity in symbolic comparisons: Evidence against an expectancy hypothesis. *Memory & Cognition, 9* (2), 197-204.

Holyoak, K. J., & Patterson, K. K. (1981). A positional discriminability model of linear-order judgements. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1283-1302.

Jamieson, D. G., & Petrusic, W. M. (1975). Relational judgments with remembered stimuli. *Perception and Psychophysics*, *18*, 373-378.

Johnson-Laird, P. N. (1983). *Mental Models.* Cambridge, UK: Harvard University Press.

Johnson-Laird, P. N. & Byrne, R. M. J. (1991). *Deduction.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Leth-Steensen, C., & Marley, A. A. J., (2000). A model of response time effects in symbolic comparison. *Psychological Review, 107* (1), 62-100.

Link, S. W. (1990). Modeling imageless thought: The relative judgment theory of numerical comparisons. *Journal of Mathematical Psychology*, *34*, 2-41.

Logie, R. H. (1995). *Visuo-spatial working memory.* Hove, UK: Lawrence Erlbaum Associates, Publishers Ltd.

Marschark, M., Azmitia, M., & Paivio, A. (1985). Associative priming in symbolic comparisons by adults and children. *Bulletin of the Psychonomic Society, 23* (6), 459-461

McConnell, J., & Quinn, J. G. (2004). Complexity factors in visuo-spatial working memory. *Memory, 12* (3), 338-350.

Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception and Psychophysics*, *13*, 180-184.

Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, *8*, 228-246.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1-18). New York: Plenum Press.

Petrusic, W. M. (1992). Semantic congruity effects and theories of the comparison process. *Journal of Experimental Psychology: Human Perception and Performance, 18* (4), 962-986.

Petrusic, W. M., & Baranski, J. V. (1989). Semantic congruity effects in perceptual comparisons. *Perception & Psychophysics, 45,* 439-452.

Poltrock, S. E. (1989). A random walk model of digit comparison. *Journal of Mathematical Psychology, 33,* 131-162.

Potts, G. R. (1972). Information processes used in the encoding of linear orderings. *Journal of Verbal Learning and Verbal Behaviour, 11,* 727-740.

Potts, G. R. (1974). Storing and retrieving information about ordered relationships. *Journal of Experimental Psychology, 103,* 431-439.

Potts, G. R., & Sholz, K. W. (1975). The internal representation of a three-term series problem. *Journal of Verbal Learning and Verbal Behaviour, 14,* 439-452.

Shaki, S., & Algom, D. (2002). The locus and nature of semantic congruity in symbolic comparison: Evidence from the stroop effect. *Memory & Cognition, 1,* 3-17.

Shoben, E. J., Čech, C. G., Schwanenflugel, P. J., & Sailor, K. M. (1989). Serial position effects in comparative judgments. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 273-286.

Sternberg, R. J. (1980). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General, 109,* 119-159.

Theberge, J. (2003). *Symbolic comparison and working memory.* Unpublished undergraduate Thesis, Carleton University, Ottawa, Ontario, Canada.

Te Linde, J., & Paivio, A. (1979). Symbolic comparison of color similarity. *Memory & Cognition, 7* (2), 141-148.

Toms, M., Morris, N., & Ward, D. (1993). Working memory and conditional reasoning. *Quarterly Journal of Experimental Psychology, 46A,* 679-699.

Towse, J. N. (1998). On random generation and the central executive of working memory. *British Journal of Psychology, 89* (1), 77-101.

Vandierendonck, A., & De Vooght, G. (1997). Working memory constraints on linear reasoning with spatial and temporal contents. *Quarterly Journal of Experimental Psychology, 50A,* 803-820.

Vandierendonck, A., Dierckx, V., & De Vooght, G. (2004). Mental model construction in linear reasoning: Evidence for the construction of initial annotated models. *The Quarterly Journal of Experimental Psychology, 57* (8), 1369-1391.

Vandierendonck, A., Kemps, E., Fastame, M.C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology, 95,* 57-79.

Woocher, F. D., Glass, A. L., & Holyoak, K. J. (1978). Positional discriminability in linear orderings. *Memory and Cognition, 6,* 165-173.

TABLES

Table 1

Summary of Distance Effect Results in Terms of Response Time

Condition	Effect of Condition	Distance Effect	Condition x Distance Effect
AS	Yes	Yes	No
VT	Yes[a]	Yes	No
RG	Yes	Yes	No

[a] But only marginally

Table 2

Summary of Distance Effect Results in Terms of Accuracy

Condition	Effect of Condition	Distance Effect	Condition x Distance Effect
AS	Yes	Yes	No
VT	Yes	Yes	No
RG	Yes	Yes	Yes[a]

[a] But only when Split 5 was omitted from the analysis.

Table 3

Summary of End and Semantic Congruity Effect Results in Terms of Response Time

Condition	Split	Condition Effect	End Effect	Condition x End Effect	SCE	Condition x SCE
AS	1	Yes	Yes	No	Yes[d]	No
	2	Yes	Yes	No	Yes	No
	3	Yes	Yes	No	Yes	No
	4	Yes[a]	No	No	Yes	No
VT	1	No	Yes	No	Yes	No
	2	No	Yes	No	Yes	No
	3	Yes	Yes[b]	No	No	No
	4	Yes[a]	Yes	No	Yes[a]	No
RG	1	Yes	Yes	No	No	No
	2	Yes	Yes	No	No	No
	3	Yes	Yes[c]	No	Yes	No
	4	Yes	No	No	No	No

[a] But only marginally

[b] But only marginally for the quadratic trend

[c] Not for the quadratic trend

[d] But only marginally for the non-trend Pair x Instruction effect

Table 4

Summary of End and Semantic Congruity Effect Results in Terms of Accuracy

Condition	Split	Condition Effect	End Effect	Condition x End Effect	SCE	Condition x SCE
AS	1	No	Yes[a]	No	No	No
	2	No	Yes[a]	No	No	No
	3	No	No	No	No	No
	4	Yes	No	No	Yes	No
VT	1	Yes	Yes[b]	No	No	No
	2	Yes	Yes	No	No	No
	3	Yes	No	No	No	No
	4	Yes	No	No	No	Yes[c]
RG	1	Yes	No	No	No	No
	2	No	No	No	No	Yes[c]
	3	Yes[c]	No	No	No	No
	4	No	No	No	No	No

[a] But only marginally for the non-trend pair effect

[b] But only for the quadratic trend

[c] But only marginally

FIGURES

Figure 1. Distance effect plots of mean response times (RT; Figure 1A) and accuracy (in terms of proportion correct [PC]; Figure 1B) for the control and secondary task conditions (error bars represent one standard error for the corresponding data point). Separate plots are given for each secondary task group.

Figure 2. Response time (RT) plots for the control (top plot) and articulatory suppression (bottom plot) groups under each of the "Shorter?" and "Taller?" instructions. Separate plots are assigned to each split.

Figure 3. Response time (RT) plots for the control (top plot) and visuo-spatial tapping (bottom plot) groups under each of the "Shorter?" and "Taller?" instructions. Separate plots are assigned to each split.

Figure 4. Response time plots (RT) for the control (top plot) and random letter generation (bottom plot) groups under each of the "Shorter?" and "Taller?" instructions. Separate plots are assigned to each split.

Figure 5. Accuracy plots (in terms of proportion correct [PC]) for the control (top plot) and articulatory suppression (bottom plot) groups under each of the "Shorter?" and "Taller?" instructions. Separate plots are assigned to each split.

Figure 6. Accuracy plots (in terms of proportion correct [PC]) for the control (top plot) and visuo-spatial tapping (bottom plot) groups under each of the "Shorter?" and "Taller?" instructions. Separate plots are assigned to each split.

Figure 7. Accuracy plots (in terms of proportion correct [PC]) for the control (top plot) and random letter generation (bottom plot) groups under each of the "Shorter?" and "Taller?" instructions. Separate plots are assigned to each split.

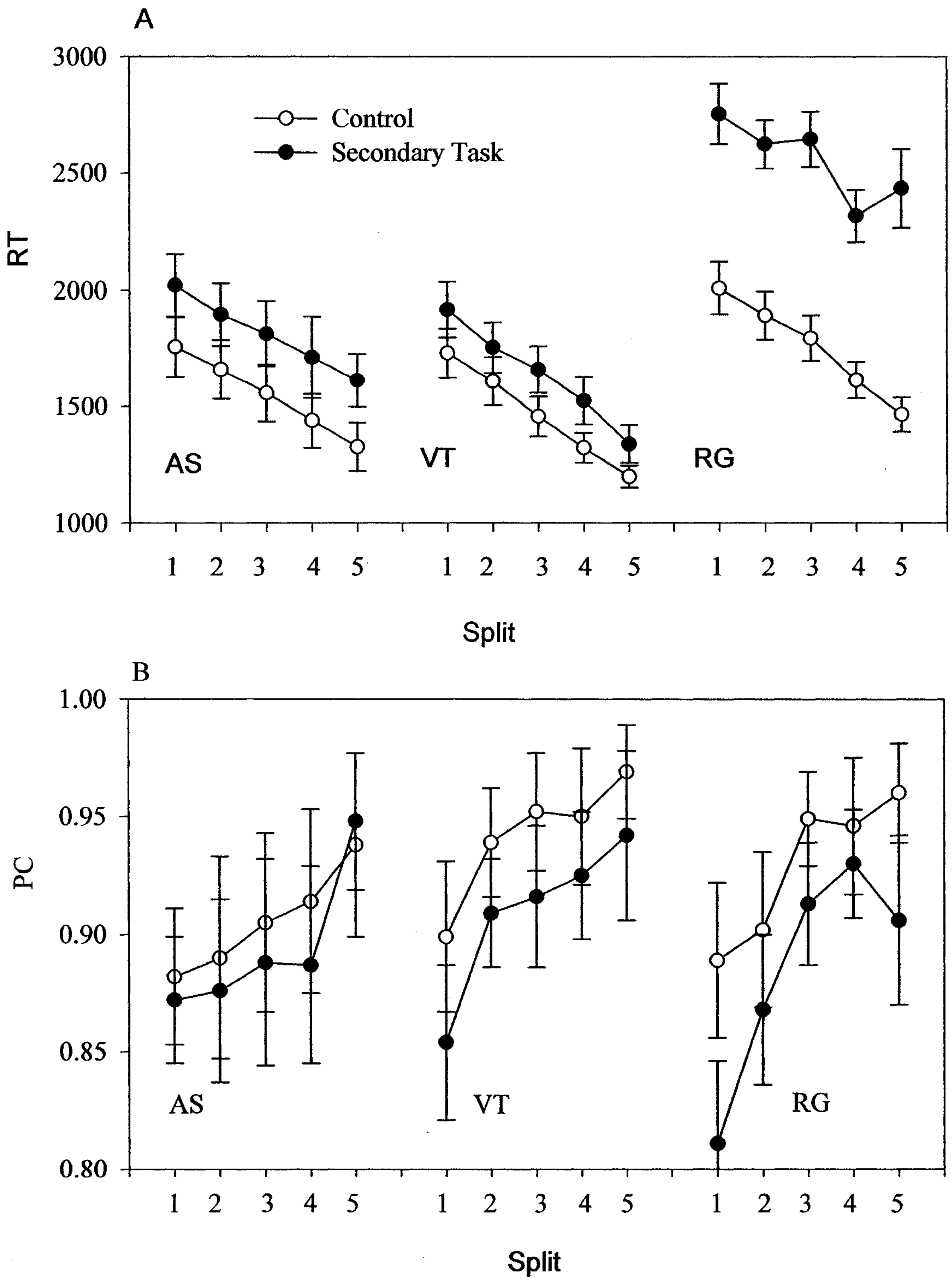

A
3000
2500
2000
1500
1000
RT
Control
Secondary Task
AS
VT
RG
1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
Split
B
1.00
0.95
0.90
0.85
0.80
PC
AS
VT
RG
1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
Split

Figure 1

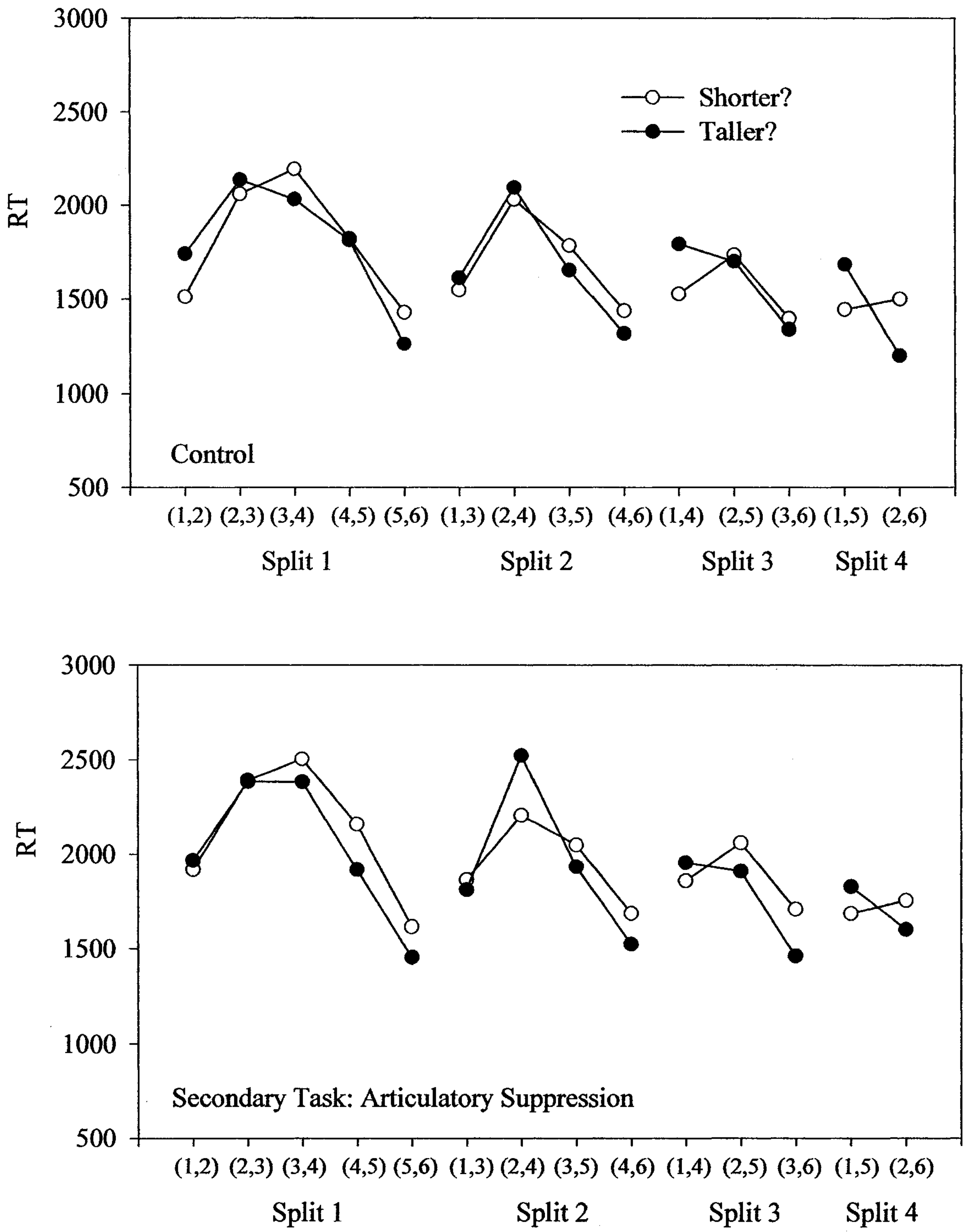
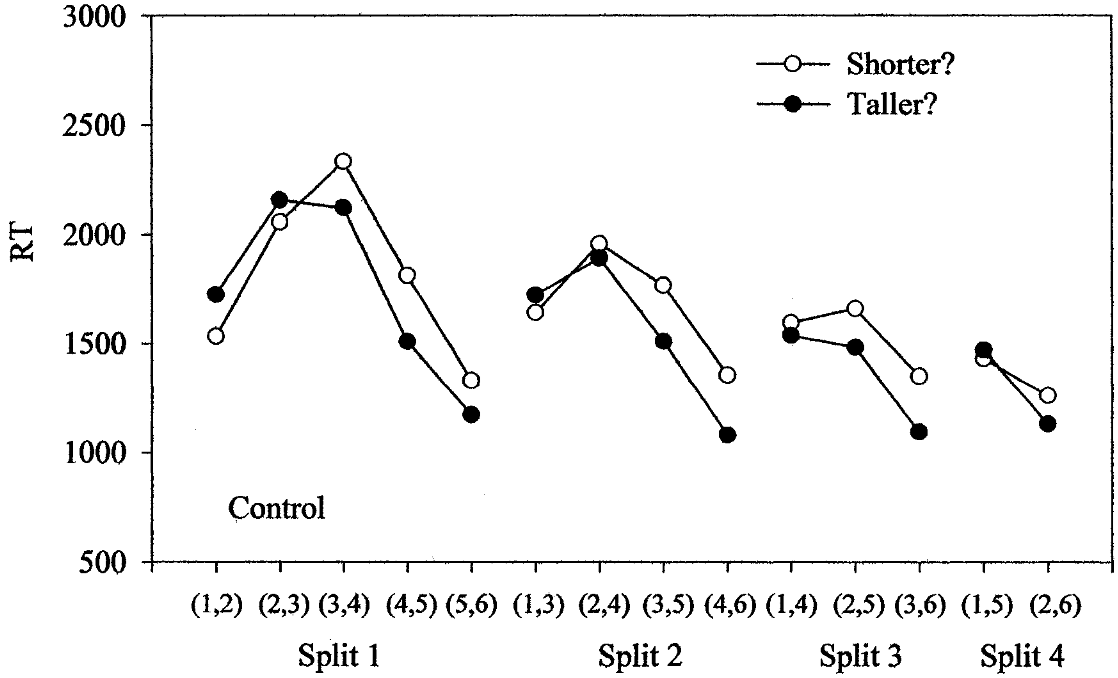
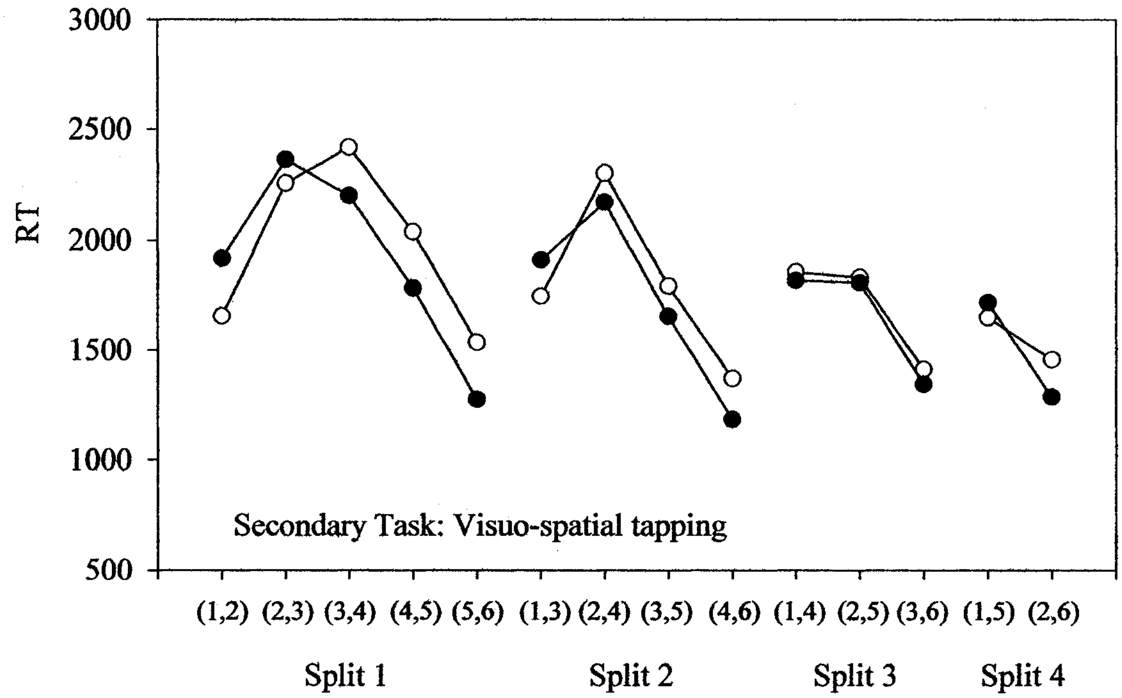

3000
2500
2000
1500
1000
500
RT
Shorter?
Taller?
Control
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4
3000
2500
2000
1500
1000
500
RT
Secondary Task: Articulatory Suppression
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4

Figure 2



3000
2500
2000
1500
1000
500
RT
Shorter?
Taller?
Control
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4



3000
2500
2000
1500
1000
500
RT
Secondary Task: Visuo-spatial tapping
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4

Figure 3

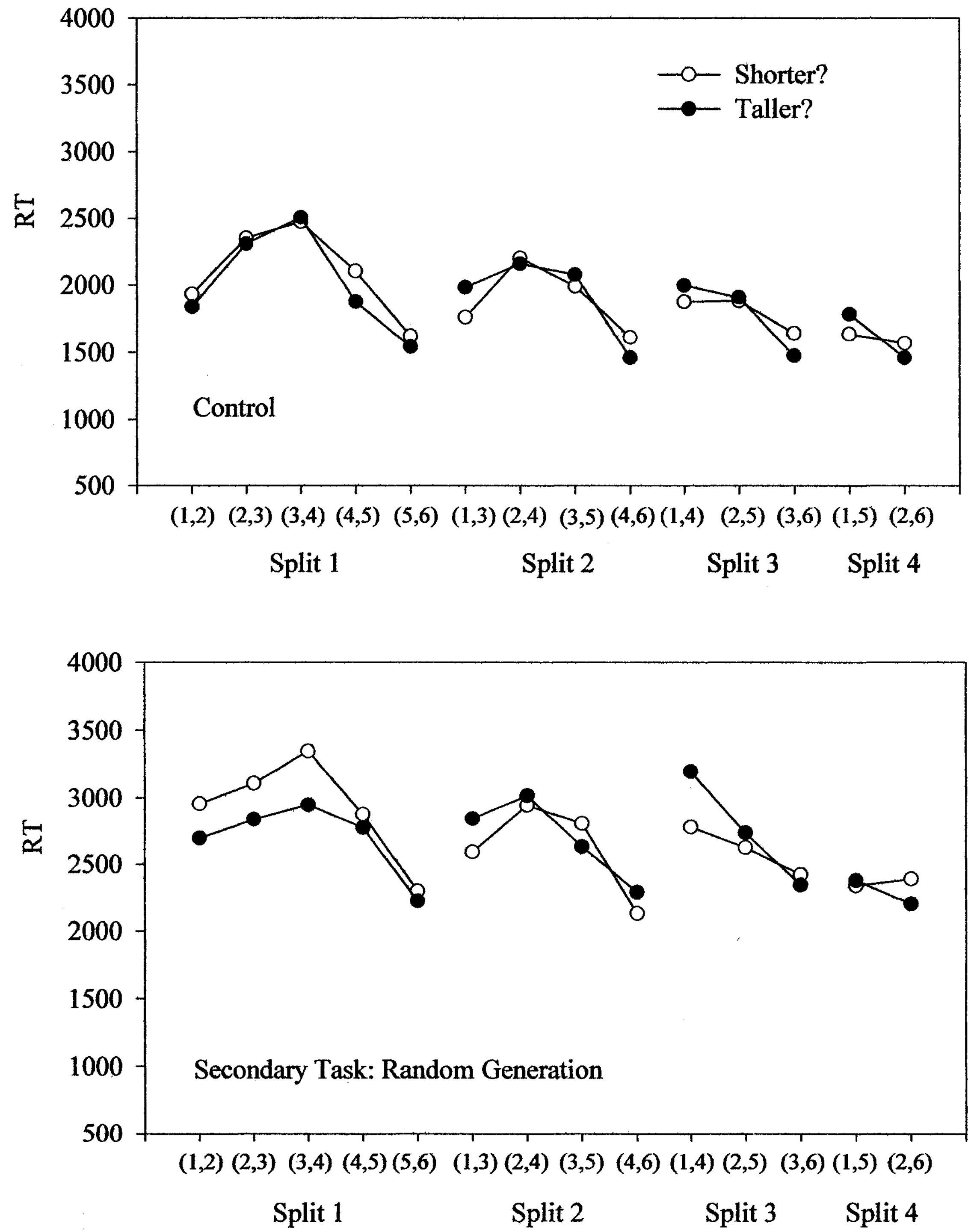
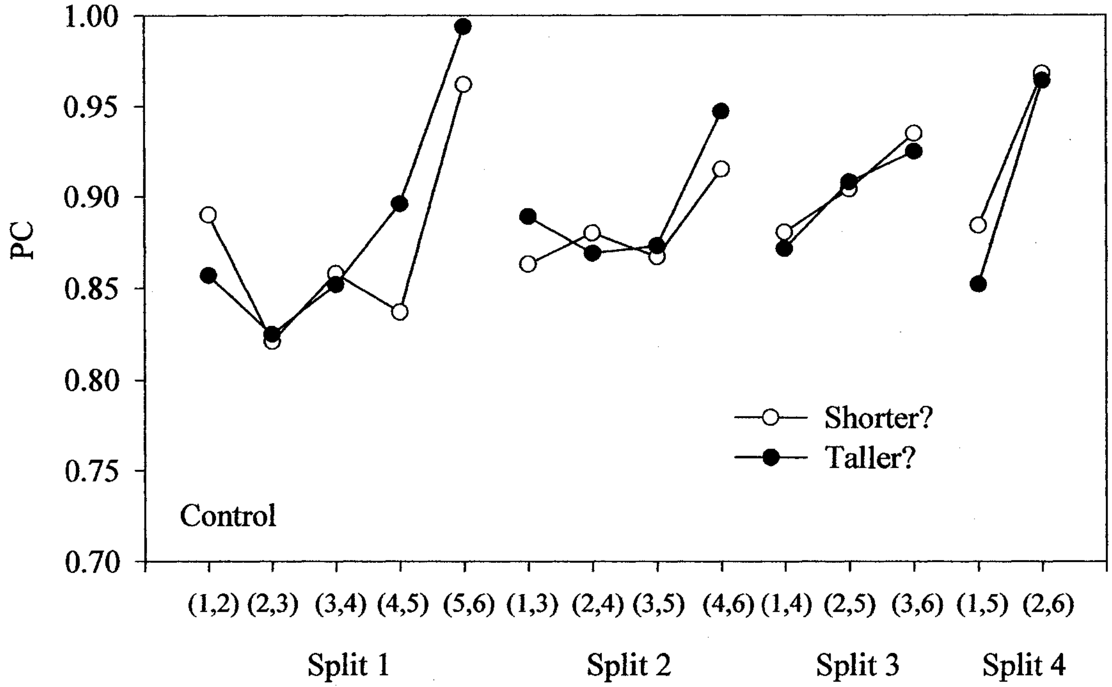

4000
3500
3000
2500
2000
1500
1000
500
RT
Shorter?
Taller?
Control
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4
4000
3500
3000
2500
2000
1500
1000
500
RT
Secondary Task: Random Generation
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4

Figure 4


1.00
0.95
0.90
0.85
0.80
0.75
0.70
PC
Shorter?
Taller?
Control
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4

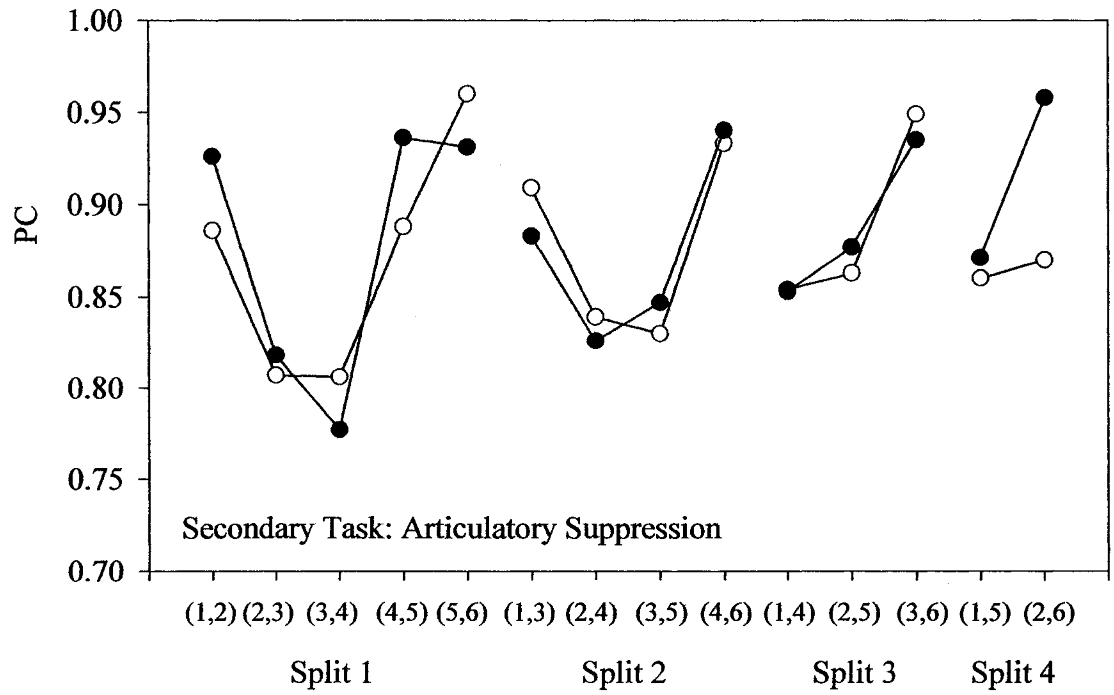
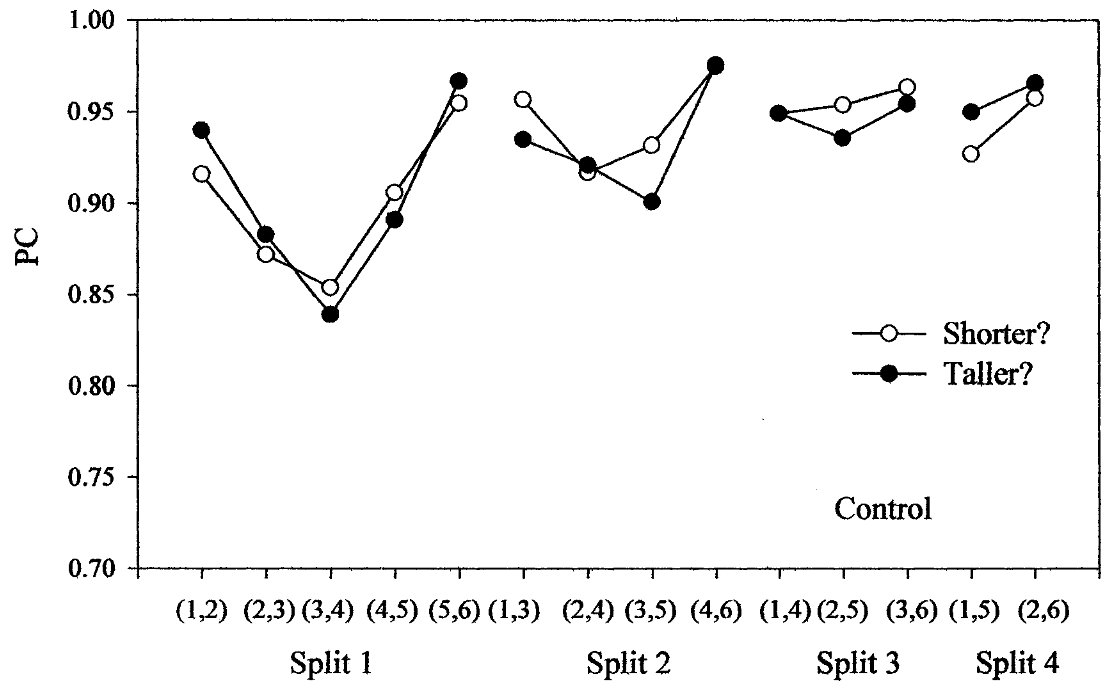

1.00
0.95
0.90
0.85
0.80
0.75
0.70
PC
Secondary Task: Articulatory Suppression
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4

Figure 5


1.00
0.95
0.90
0.85
0.80
0.75
0.70
PC
Shorter?
Taller?
Control
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4

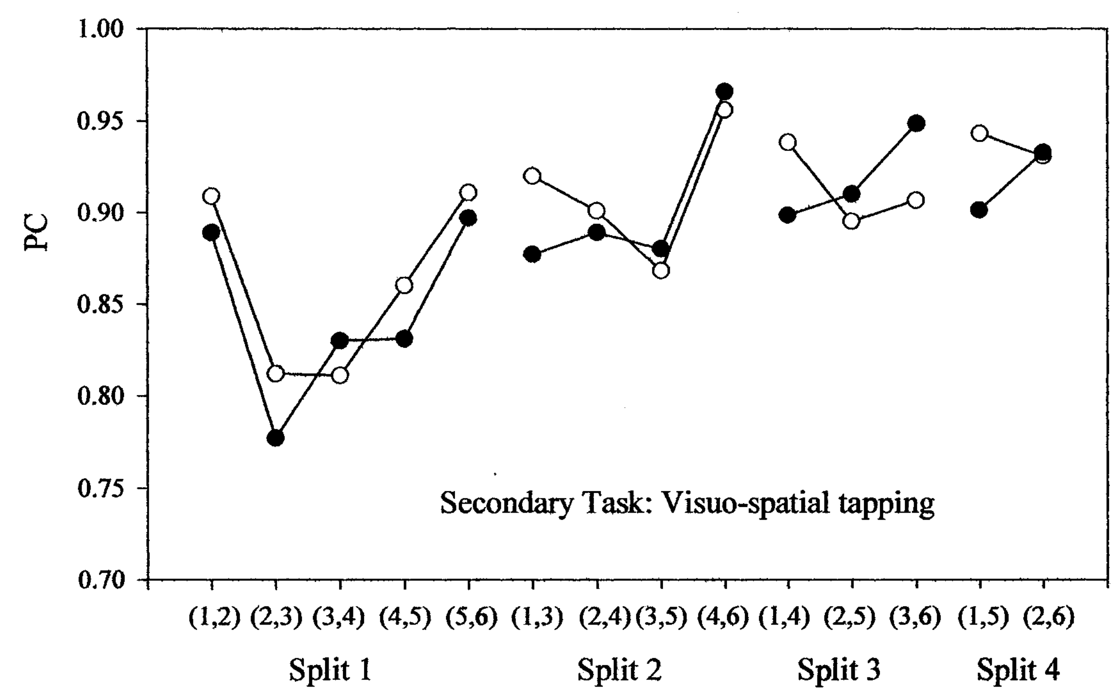

1.00
0.95
0.90
0.85
0.80
0.75
0.70
PC
Secondary Task: Visuo-spatial tapping
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4

Figure 6

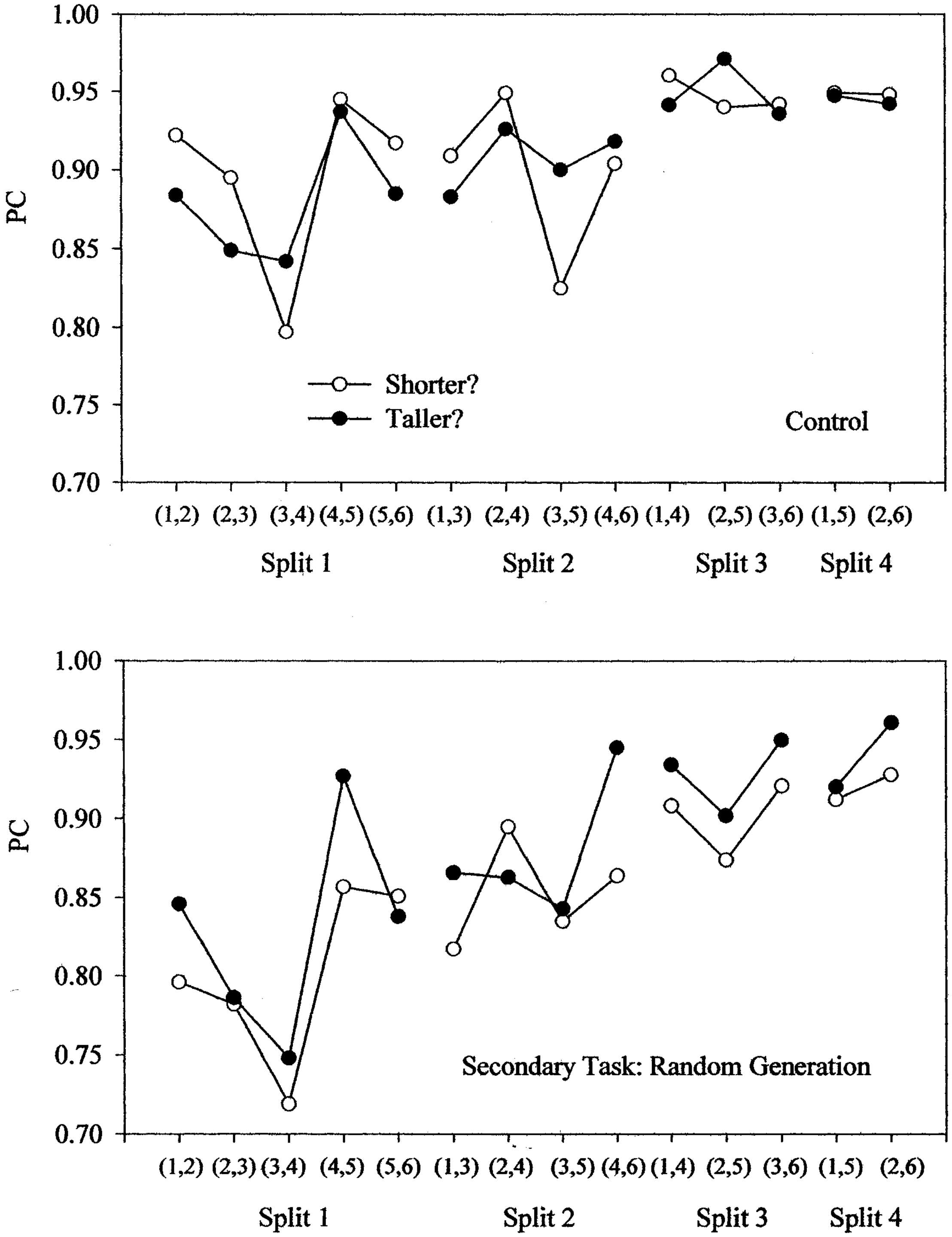

1.00
0.95
0.90
0.85
0.80
0.75
0.70
PC
Shorter?
Taller?
Control
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4
1.00
0.95
0.90
0.85
0.80
0.75
0.70
PC
Secondary Task: Random Generation
(1,2) (2,3) (3,4) (4,5) (5,6) (1,3) (2,4) (3,5) (4,6) (1,4) (2,5) (3,6) (1,5) (2,6)
Split 1
Split 2
Split 3
Split 4

Figure 7

APPENDIX

Phase 1

There will be two parts to this experiment.

In the first part you will be presented with **pairs** of 3-letter **names** and will be required to make a comparison between them. Each name is a label that stands for an imaginary "person". There are six "people" in this experiment, all of whom are assumed to differ in height. Hence, on each comparison trial you will be presented with two of these names (side by side in the center of the computer screen) and will have to decide which "person" is either the "**Taller?**" or the "**Shorter?**".

After each trial you will be informed as to whether you were correct or not and will be provided with the actual correct response. The comparison pairs will be shown repeatedly throughout this part of the experiment so you should use this **feedback** to help you to learn which person is the "taller" and which person is "shorter" in each pair. You will not be expected to know this information at the very start of the experiment (and hence you will initially be guessing) but as you go along you will gradually be able to learn the correct relations between the members in each pair.

On each trial an instruction will be presented. This instruction will tell you the kind of comparison that you will be making on that trial (i.e., choose the "Shorter?" or choose the "Taller?"). Shortly afterwards the "plus sign" in the center of the screen will be replaced by the comparison pair. To make your response simply press the **left** key marked in **yellow** on the keyboard in front of you if you feel that the "person" on the left is the correct response and press the **right** key marked in **yellow** if you feel that the "person" on the right is the correct response. Use your index and middle fingers on your left hand to make these responses (please try to keep these fingers on these keys as each trial starts). After each choice has been made a new instruction will appear on the screen indicating the start of the next trial.

This task takes about 20-30 minutes to perform and will be completed after you have reached a certain number of correct responses. The computer will indicate when the task has been completed. I will be here at the beginning to aid you. When you have finished this task please come and get me so I can set up the second part of the experiment. You may proceed when ready.

Phase 2A

The procedure for the second part of this experiment is quite similar to that in the first part except for a few differences. First, there will be **no feedback** provided and the comparison trials will be presented in a continuous fashion (one right after the other). Second, some of the comparison pairs will be **new** to you but you should have no trouble responding to those pairs given the information that you have learned in the first part of the experiment. Finally, during this part of the experiment, there will be four blocks of about 120 trials each. For two of these blocks you will be asked to do a secondary task at the same time that you make the comparison responses. Before these blocks, instructions will be provided as to the nature of the secondary task that you will be performing and a small set of practice trials will also be given in order to get you used to doing the two tasks at once. Try to be accurate in each decision **without** taking too much time to respond. You may **rest** between each block (you will be informed of this by the computer).

<u>Secondary Task</u> (1)

In these two blocks of comparison trials, you will have to simultaneously perform a letter vocalizing task. This task simply involves repeating out loud the letter sequence "F, J, N, W" throughout the whole course of the block. Please try to repeat these letters as quickly <u>as possible</u>. At the start of this block you will be given a 15 second period to commence this task (which you will simply continue to perform when the names appear and the comparison task begins). Please try to be consistent in this letter vocalizing task while performing the Shorter-Taller comparison task at the same time. There are no tricks involved here we simply want to see how well you can perform both tasks at the same time. However, the experimenter will be in the room to make sure that you are keeping "on track" with both tasks.

<u>Secondary Task</u> **(2)**

In these two blocks of comparison trials, you will have to simultaneously perform a spatial tapping task. This task simply involves using your right index finger to tap the four keys marked in green on the right hand side of the keyboard in order (see numerical labels) throughout the whole course of the block. Please try to tap out these keys as quickly <u>as possible</u>. At the start of this block you will be given a 15 second period to commence this task (which you will simply continue to perform when the names appear and the comparison task begins). Please try to be consistent in this tapping while performing the Shorter-Taller comparison task at the same time. There are no tricks involved here we simply want to see how well you can perform both tasks at the same time. However, the experimenter will be in the room to make sure that you are keeping "on track".

Secondary Task (3)

In these two blocks of comparison trials, you will have to simultaneously perform a random letter generation task. This task simply involves repeating out loud the letters "F, J, N, W" in a random order throughout the whole course of the block. Please try to generate these random letters as quickly as possible. Note that while random sequences could be regarded as fairly patternless, they do sometimes contain some repetition and/or symmetry. At the start of this block you will be given a 15 second period to commence this task (which you will simply continue to perform when the names appear and the comparison task begins). Please try to be consistent in this random generation task while performing the Shorter-Taller comparison task at the same time. There are no tricks involved here we simply want to see how well you can perform both tasks at the same time. However, the experimenter will be in the room to make sure that you are keeping "on track".