Spatial Tapping Interferes With the Processing of Linguistic Spatial Relations

Matthijs L. Noordzij, Rob H. J. van der Lubbe, Sebastiaan F. W. Neggers, and Albert Postma
Psychological Laboratory, Helmholtz Institute, Utrecht University

Abstract Simple spatial relations may be represented either in a propositional format that is dependent on verbal rehearsal or in a picture-like format that is maintained by visual-spatial rehearsal. In sentence-picture and picture-picture verification tasks, we examined the effect of an articulatory suppression and a spatial tapping dual task on the encoding of simple spatial relations (e.g., triangle left of circle). Articulatory suppression did not interfere, while spatial tapping lowered performance in both tasks. Apparently, both linguistic and perceptual inputs of simple spatial relations engaged the visual-spatial working memory. In the sentence-picture verification experiments, spatial tapping only hampered performance of participants who were classified on the basis of their RT patterns as having used a visual-spatial strategy, while it had no effect for those who were classified as having applied a verbal strategy. Therefore, this study provides converging evidence, using a dual-task methodology, that both separate verbal and visual-spatial strategies exist for the processing of simple spatial sentences.

A picture of a spatial configuration of two objects can be easily described in a sentence. This description will usually mention the objects and the categorical spatial relation between them: "The knife is to the right of the plate." Conversely, it is also easy to construct a generic or specific image of the aforementioned spatial description. An interesting question is to what extent representations generated by verbal or pictorial information are similar. In a well-known spatial description task (Clark & Chase, 1972; Just & Carpenter, 1975), which enables an answer to this question, participants read a sentence describing the categorical spatial relation between two objects (e.g., star is above plus), followed by a picture that does (e.g., a star above a plus) or does not (e.g., a plus above a star) exemplify the description. The participant has to decide whether the picture is a true or false depiction of the spatial description. Several authors (e.g., MacLeod, Hunt, & Mathews, 1978; Reichle, Carpenter, & Just, 2000) suggested that this sentence-picture verification task can be achieved in two distinct ways: using a verbal or a visual-spatial strategy. The verbal strategy involves the formation of a propositional representation dependent on verbal working memory, whereas the visual-spatial strategy involves the formation of a pictorial representation dependent on visual-spatial working memory. The current study used a dual task methodology to improve our understanding of the type of representation that is employed in the verification of sentences and pictures that specify *left* and *right* relations between two objects.

In a sentence-picture verification task participants compare linguistic information in a sentence with visual-spatial information in a picture. Some transformations seem necessary to enable a comparison between the two types of information. One possibility is that the representation of a spatial sentence has a proposition-based format that is maintained through active verbal rehearsal. Subsequently, the picture is encoded into the same propositional format as the sentence, after which a comparison is possible. Evidence for the employment of verbal strategies comes from studies in which the linguistic complexity of the spatial sentence was varied. The underlying assumption is that the representations (of the sentence and of the picture) are compared component by component, until all components have been resolved. Adding more linguistic components, such as a negation, would lengthen the verification time. Hence, the sentence "star is above plus" should be compared faster to a subsequent picture than the sentence "star is not above plus." In line with this prediction, Clark and Chase (1972) found that participants became slower when the sentence became more complex (e.g., the star is not above the plus).

Another possibility is that the initial sentence is transformed immediately into a representation with a pictorial format, which can be maintained by active visual-spatial rehearsal until the picture appears.

Subsequently, the pictorial representation of the sentence can be directly compared to the picture. For a visual-spatial strategy the linguistic complexity of the sentence should not play a role in the comparison process because the sentence is converted to a pictorial format that does not retain the additional complexities (e.g., negation). Therefore, the sentence "star is above plus" should be compared equally fast to a subsequent picture as the sentence "star is not above plus." Evidence for the employment of a visual-spatial strategy comes from studies in which linguistic complexity had little or no effect on picture verification time (e.g., Seymour, 1974).

Apparently, evidence is found for the presence of verbal strategies as well as visual-spatial strategies in sentence-picture verification tasks. Interestingly, it has been suggested that both strategies might apply (i.e. there may be intersubject differences in which strategy is preferred). Gluschko and Cooper (1978) showed that the effects of linguistic complexity could be diminished by explicitly instructing participants to use imagery to solve the task. MacLeod et al. (1978) argued that they could divide their participant group on the basis of RTs into one group that employed the verbal strategy and a group that was supposed to employ the visual-spatial strategy. Furthermore, participants who scored high on psychometric measures of spatial ability were more likely to have chosen a visual-spatial strategy.

Eley (1981) compared RT-patterns of participants, who were classified as having followed a visual-spatial or verbal strategy, in sentence-picture and picture-picture verification tasks. For one group this classification was based upon self-reports after the experiment, whereas another group was instructed to follow a specific strategy. Participants who were classified as having used a visual-spatial strategy were equally fast on a sentence-picture verification task as on a picture-picture verification task, while participants who were classified as having used a verbal strategy were slower on a sentence-picture verification task than on a picture-picture verification task. Eley argued that participants who follow a visual-spatial strategy transform the sentence representation into a pictorial representation before the picture appears. Consequently, the comparison between the pictorial representations of the sentence and the picture should be very similar to the comparison process in a picture-picture verification task. In contrast, participants who follow a verbal strategy have to transform the picture into a propositional representation in the sentence-picture verification task. This additional transformation, which is absent in picture-picture verification, may be responsible for longer RTs in the sentence-picture task than in the picture-picture task. Further support for the division of participants into a verbal and visual-spatial group comes from neuroimaging research. In a recent fMRI study (Reichle et al., 2000), the sentence-picture verification paradigm was used to investigate the cortical systems that are involved with verbal and visual-spatial processing. Participants who were instructed to follow a verbal strategy produced more activation in traditional language areas such as Broca's area, whereas participants who were instructed to follow a visual-spatial strategy showed more activation in traditional visual-spatial areas such as the parietal cortex.

The sentence-picture verification task has proven to be a paradigm that is useful for contrasting different types of processes, namely verbal and visual-spatial, the application of which can differ between participant groups as a whole as well as between individual participants. Nevertheless, there could be some problems with classifying a type of processing by looking at RT-patterns (or possibly error scores) in a sentence-picture verification task. As noted by Marquer and Pereira (1990), two identical individual verification RT-patterns can originate from two different strategies. Hence, Marquer and Perreira suggested that a more valid approach would be to use self-reports to classify participants into strategy groups and then compare this classification against RT-patterns. However, a commonplace finding in cognitive psychology is that introspective reports provide inaccurate information about the processes that participants actually employ in a task (e.g., Evans, 1989). In a review of the sentence-picture verification task, Roberts, Wood, and Gilmore (1994) argued that all the main strategy classification systems based upon verification RT-patterns, and also sources of converging evidence based on psychometric test scores and introspective reports, seem to be problematic. The lack of clarity that seems to be associated with the methods that have been used for strategy classification in sentence-picture verification tasks could be avoided by introducing additional tasks that interfere unambiguously with either verbal or visual-spatial processes. It was exactly the goal of the present study to employ dual tasks, which selectively interfere with the verbal or spatial components of working memory, in order to shed light on which type of representation, that is, propositional or pictorial, is employed in the verification of spatial relationships.

In the domain of working memory there are many studies that support the view that verbal and visual-spatial information are processed by dissociable subcomponents. Therefore, the examination of the subcomponents of working memory that are engaged during a sentence-picture verification task might provide an insight into which of the two strategies are employed. In Baddeley's (1986) model of working memory, the storage and processing of task-relevant information is achieved by three modules. Verbal information is encoded and maintained by the so-called phonological loop, whereas visual-spatial information is processed by another specialized system – the so-called visual-spatial sketchpad. Both systems are coordinated and controlled by a central executive of limited capacity. The verbal strategy in a sentence-picture verification task can be expected to depend on the phonological loop; the propositional representation of the first sentence needs to be constructed and maintained in verbal working memory. In contrast, the visual-spatial strategy can be expected to depend on the visual-spatial working memory; the transformation of the first sentence into a representation with a pictorial format and the subsequent rehearsal of this visual-spatial representation would require the use of visual-spatial working memory.

Dual-task procedures have been used extensively in studies aiming to disrupt the phonological loop or the visual-spatial working memory. A secondary task that is thought to disrupt the phonological loop is articulatory suppression (e.g., Baddeley & Andrade, 2000; Chincotta & Underwood, 1997; Milner, Jeeves, Ratcliff, & Cunnison, 1982), which requires participants to repeat aloud a single word as "the" or a predictable sequence such as the digits 1, 2, 3, and 4. The working memory model accounts for the effects of articulatory suppression by suggesting that it places a load on the articulatory rehearsal process, thereby undermining an important facility for retention. A secondary task that is thought to disrupt the visual-spatial working memory is spatial tapping (e.g., Beech, 1984; Logie, 1995; Sussman, 1982), which requires participants to tap a predetermined spatial array, like the four corner points of a square following a counterclockwise direction. The working memory model accounts for these effects by suggesting that it places a load on an active (spatial) rehearsal process, thereby impairing retention possibilities.

In the present study we employed sentence-picture and picture-picture verification tasks, in combination with the aforementioned verbal and spatial dual tasks. Our main interest was to further examine what type of representation is employed when participants encode categorical spatial relations. If participants follow a verbal strategy and construct a propositional representation then their performance should be more impaired by articulatory suppression than by spatial tapping. If participants follow a visual-spatial strategy and construct a pictorial representation then their performance should be more impaired by spatial tapping than by articulatory suppression. In addition, the possibility that different participants employed different strategies in the sentence-picture verification tasks was examined. Following Eley (1981), participants who showed similar RTs in a sentence-picture and a picture-picture task were classified as having used a visual-spatial strategy, whereas participants who showed slower RTs, in a sentence-picture task than in a picture-picture task were classified as having used a verbal strategy. To verify the validity of the classification on basis of RTs, we examined whether "verbal" participants were indeed most interfered by articulatory suppression, and "visual-spatial" participants were more hampered by spatial tapping.

Experiment 1

Method

Participants. Eighteen right-handed participants (9 men, 9 women, all undergraduate students), with normal or corrected-to-normal vision, gave informed consent. They were naive with respect to the hypotheses and were paid €7 per hour for participating.

Materials. Stimuli were presented on a 19" Dell monitor with E-Prime software running on a Pentium III computer. Visual stimuli were composed of three shapes (circle, triangle, and square). The stimuli were the same colour (black), the same size, and at the same distance from one another (although this distance was twice as far for the first and second picture). Pictorial objects subtended 4.8° x 4.8° of visual angle. The text that participants read (e.g., Triangle left of Circle) was 18-point Times New Roman type, and was written in Dutch. Individual letters subtended 1.0° x 0.6° of visual angle. In a sentence-picture verification task participants read a sentence and in the picture-picture verification task they saw a picture; both were placed in the centre of the screen and provided information concerning a simple categorical relation (only "to the left of" and "to the right of") between two objects (circle, square, and triangle). Subsequently, a picture was presented and participants had to decide as fast and accurately as possible whether this second picture exemplified the sentence or first picture correctly. The second picture was placed randomly in one of four positions (top-left, top-right, bottom-left, bottom-right). The second picture depicted the two objects that were presented in the first sentence in a horizontal (left-right) relation.

A serial response box was used to collect key-press responses from the participants. A quartz metronome was used to train the participants to tap and count at a speed of 184 beats per minute. The spatial tapping task consisted of four wooden plates, each 70 mm square, arranged in a square on a horizontal board with 25 mm between each plate. The three-dimensional position of the tapping finger was sampled at a rate of 100 Hz by means of a miniBIRD 800 motion tracking system from Ascension Technology Corporation.

Design and procedure. Before the start of the experiment participants were trained on the interference tasks. For the articulatory suppression task, participants were asked to count aloud repeatedly from one to four throughout presentation of the first stimulus. For the spatial tapping task, participants were asked to tap repeatedly with their left (nondominant) hand throughout presentation of the first stimulus. The tapping pattern consisted of touching with the index finger in a counterclockwise direction each of the plates (which were arranged in a square) in turn. The wooden plates were placed in a closed box with a half-open front for the tapping hand to make sure that participants did not look at their hands while executing the spatial tapping task. The experiment started when participants were able to perform the interference tasks without mistakes and at the correct speed. The presentation of stimuli and interference tasks was blocked and counterbalanced over participants. Each condition consisted of 4 practice trials and 36 experimental trials.

At the beginning of each trial a fixation cross was presented in the centre of the screen and after 1,000 ms a beep of 1,000 Hz sounded (500 ms), indicating that participants had to start the interference task (counting or tapping). After 1,500 ms a sentence or a picture appeared for 3,000 ms. When the stimulus disappeared another 1,000Hz beep sounded to signal participants to stop counting or tapping. After a delay of 750 ms with a fixation cross, a picture was presented and participants had to press the correct button with their right hand for the same/different decision. Participants responded by pressing the left or right button of the response box. A left button press equalled "same" for one-half of the participants and a right button press equalled "same" for the other half of the participants. The beeps were also presented in the conditions without an interference task (baseline conditions), but participants were told to ignore them.

Data analysis. The data of the practice trials were discarded, as were trials on which the RT was either 2.5 standard deviations above or below the mean of the condition. As a result of this criterion, .5% of the trials were considered to be an outlier. The between-subjects variable Group was constructed by using a K-means clustering algorithm that maximized a t-test for two groups based on the differences between RTs in the sentence-picture and picture-picture baseline task (see Mathews, Hunt, & MacLeod, 1980). This resulted in a visual-spatial group (n = 11) who showed a difference in RTs (M = 118 ms, minimum = -72 ms, maximum = 244 ms, SD = 92 ms) between the sentence-picture and picture-picture baseline tasks that was significantly smaller than the difference in RTs (M = 409 ms, minimum = 290 ms, maximum = 542 ms, SD = 107 ms) for the verbal group (n = 7), $t(16) = 5.9$, $p < .001$. Mean RTs, computed over correct trials, and mean percentage error scores were analyzed using separate 2 x 3 x 2 ANOVAs with Format (Sentence or Picture) and Interference (Baseline, Articulatory Suppression or Spatial Tapping) as within-subjects variables and Group (Verbal or Visual-Spatial) as between-subjects variable. In addition, planned comparisons were carried out to test specifically whether the verbal and visual-spatial groups were differently impaired by the two dual tasks.

Previous studies showed that sentences with "to the right of" were compared faster to a subsequent picture than sentences with "to the left of" (e.g., Just & Carpenter, 1975). However, Just and Carpenter also found that if a picture stored in long-term memory had to be compared to a sentence then the asymmetry reversed (i.e., "left of" faster than "right of"). Therefore, an analysis of the different sentences was carried out to establish whether there was a difference between sentences containing "to the left of" and sentences containing "to the right of." Mean RTs, computed over correct trials, and mean percentage error scores were analyzed using separate 3 x 2 x 2 ANOVAs, with Interference (Baseline, Articulatory Suppression or Spatial Tapping) and Locative Preposition (Left or Right) as within-subjects variables and Group (Verbal or Visual-Spatial) as between-subjects variable. Greenhouse-Geisser correction was applied in all tests involving variables with more than two levels to correct for possible violations of sphericity assumptions (e.g., Maxwell & Delaney, 1990). An alpha level of .05 was used for all statistical tests.

Results

Reaction times. The main effect of Format was significant, $F(1, 16) = 51.8$, $p < .001$, MSE = 47,339.1, showing that RTs were faster when the format of the first stimulus was a picture (M = 754 ms) than when it was a sentence (M = 1,063; see Figure 1). The main effect of Interference was significant, $F(2, 32) = 5.3$, $p = .02$, $\varepsilon = .72$, MSE = 54,251.7. Analysis of this effect indicated that RTs were slower in the Spatial Tapping condition (M = 996 ms) than in the Baseline (M = 884 ms) and the Articulatory Suppression condition (M = 845 ms), $t(17) = 3.1$, $p = .006$ and $t(17) = 2.6$, $p = .02$, whereas RTs in the Articulatory Suppression condition did not differ from the RTs in the Baseline condition, $t(17) < 1$.

The main effect of Group was not significant, F(1,

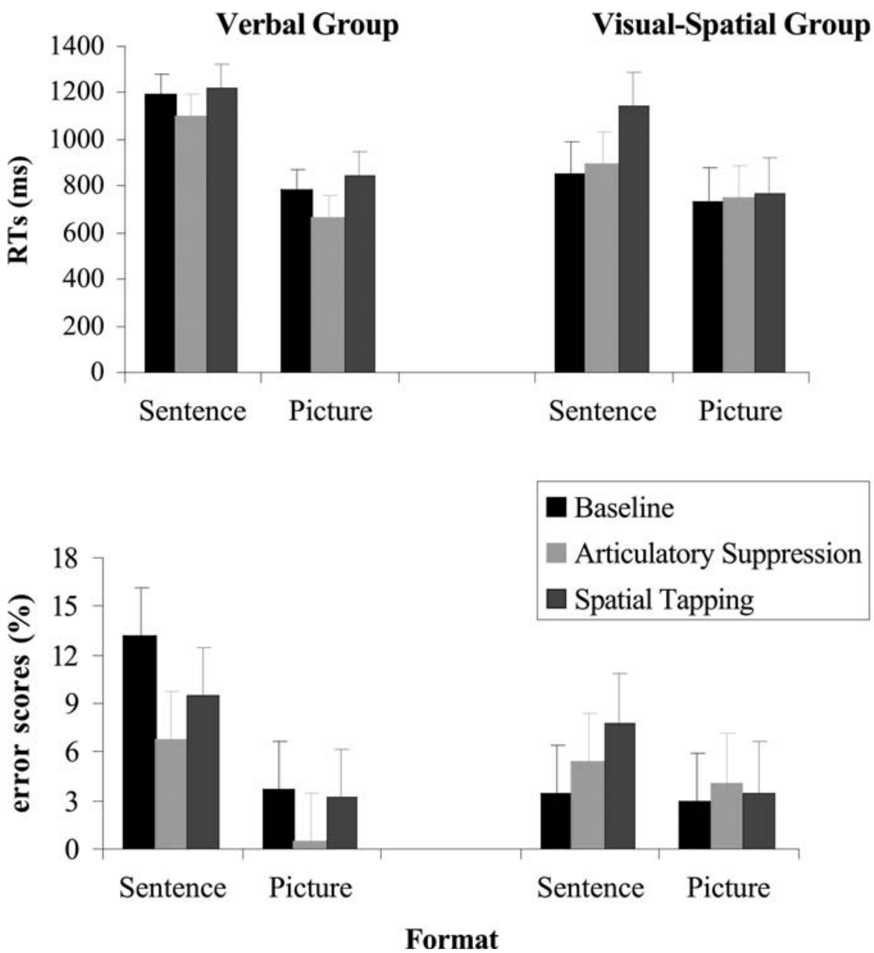

Figure 1. Mean RTs (ms) and error scores (%) for sentence-picture and picture-picture verification tasks in baseline and interference conditions in Experiment 1. Computation of within-subjects standard errors based on Loftus and Masson (1994).

16) < 1. The interaction between Format and Group was significant, $F(1, 16) = 5.2$, $p = .04$, $MSE = 47{,}339.1$, showing that the participants classified as the visual-spatial group ($M = 959$ ms) had faster RTs than the participants classified as the verbal group ($M = 1{,}167$ ms) when the first stimulus was a sentence, $t(16) = 2.6$, $p = .02$. In contrast, both groups had similar RTs when the first stimulus was a picture, $t(16) < 1$. No other interactions among Format, Interference, and Group were significant, all Fs < 1.3. Planned comparisons for the sentence-picture verification trials showed that the verbal group was neither impaired by Articulatory Suppression or by Spatial Tapping, both ts < 1, whereas the visual-spatial group was impaired by Spatial Tapping, but not by Articulatory Suppression, $t(10) = 2.6$, $p = .03$ and $t < 1$. The analyses with the variable Locative Preposition (Left or Right) showed no significant effects, all Fs < 1.6 (see Table 1).

Error scores. The main effect of Format was significant, $F(1, 16) = 28.7$, $p < .001$, $MSE = 19.9$, showing that participants made more errors when the first stimulus was a sentence ($M = 7.6\%$) than a picture ($M = 2.9\%$). The main effects of Interference and Group were not significant, $F(2, 32) = 2.0$, $p = .16$, and $F(1, 16) = 1.7$, $p = .22$. The interaction between Format and Group was significant, $F(1, 16) = 9.3$, $p = .007$, $MSE = 19.9$. Analysis of this interaction showed that participants who were classified as the verbal group made more errors when the first stimulus was a sentence ($M = 9.8\%$) than a picture ($M = 2.4\%$), $t(6) = 6.8$, $p < .001$. Participants who were classified as the visual-spatial group made the same amount of errors with sentences and pictures, $t(10) < 1.7$, $p = .13$. The interaction between Type of Interference and Group was significant, $F(2, 32) = 5.3$, $p = .01$, $\varepsilon = .98$, $MSE = 18.1$, showing that the verbal group ($M = 8.3\%$) made more errors than the visual-spatial group ($M = 3.0\%$) in the baseline condition,

TABLE 1
Mean RTs (ms) and Error Scores (%) for Sentence-Picture Tasks in Experiment 1

	RTs (ms)				Error Scores (%)			
Group	Verbal		Visual-Spatial		Verbal		Visual-Spatial	
Locative Preposition	"Left of"	"Right of"	"Left of"	"Right of"	"Left of"	"Right of"	"Left of"	"Right of"
Type of Interference								
Baseline	1,151	1,234	857	835	9.5	16.7	4.0	2.5
Articulatory Suppression	1,127	1,059	904	870	4.8	8.7	5.6	5.1
Spatial Tapping	1,220	1,229	1,196	1,109	7.1	11.9	8.1	7.6

Note: Computation of within-subjects standard errors amounted to 100 ms and 3.0% for the Verbal, and 147 ms and 3.1% for the Visual-Spatial group (based on Loftus and Masson, 1994).

$t(16) = 4.8$, $p < .001$, whereas both groups made the same amount of errors in the Articulatory Suppression and Spatial Tapping conditions, both ts < 1. No other interaction between Format, Interference, and Group was significant, all Fs < 2.0. Planned comparisons for the sentence-picture verification trials showed that the verbal group was not impaired by Spatial Tapping and, surprisingly, facilitated by Articulatory Suppression, $t <$ 1.1 and $t(6) = 2.8$, $p = .03$, whereas the visual-spatial group was marginally impaired by Spatial Tapping and not by Articulatory Suppression, $t(10)= 2.1$, $p = .06$ and $t < 1$.The analyses with the variable Locative Preposition (Left or Right) showed no significant effects, all Fs< 2.1.

Discussion

The main focus of this first experiment was to examine which representational format (pictorial or verbal) is involved in the processing of linguistic and perceptual categorical spatial relations. Participants were slower verifying the second picture when they performed the spatial tapping task than when they performed the articulatory suppression task or when they performed no secondary task. This pattern of interference suggests that in the present task conditions participants primarily relied on a visual-spatial strategy rather than on a verbal strategy in processing simple spatial relations. Participants were faster and made fewer errors when the first stimulus was a picture than when it was a sentence. Apparently, the pictorial representation based on a picture was more effectively compared to a second picture than the visual-spatial representation based on a sentence. The pictorial representation of the sentence could have been less effective for a subsequent comparison to a picture because participants made more errors reading and interpreting the sentence than viewing the picture. Although previous research showed that there might be differences in processes related to reading and understanding sentences with opposite locative prepositions such as "left of" and "right of," we found no differences in this experiment.

The division of the participants into two groups on the basis of the difference between their sentence-picture and picture-picture verification RTs, resulted in a visual-spatial group of eleven participants and a verbal group of seven participants. The visual-spatial group was faster and better than the verbal group, but only when the first stimulus was a sentence. Of course, this is consistent with the above-mentioned criterion that was used to divide the participants in groups. We found that spatial tapping significantly impaired the visual-spatial, and not the verbal group. Articulatory suppression did not impair the visual-spatial and verbal group, and surprisingly, the verbal group made fewer errors in the articulatory suppression condition than in the baseline condition. Thus, we did find that the visual-spatial group was more impaired by spatial tapping, but we did not find that the verbal group was more impaired by articulatory suppression. Possibly, the present verbal dual task interferes mostly with the phonological surface structure of the spatial sentence and not with the underlying propositional content. Therefore, a different verbal dual task might be needed to yield verbal interference effects for participants who employ a verbal strategy in sentence-picture verification paradigms. Previous research has consistently shown that articulatory suppression and spatial tapping are approximately equally demanding in their own domain (i.e. verbal or visual-spatial), and that both tasks put a negligible load on the central executive (e.g., Brooks, 1967, Quinn & McConnell, 1996; Smyth & Pelky, 1992). Hence, we want to point out here that it is not simply the case of making the verbal dual task more demanding, as we found that an equally demanding visual-spatial dual task did yield clear interference effects (indicative of a visual-spatial strategy). In addition, simply increasing the complexity of the verbal dual task would unavoidably turn it into a central executive interference task as well, with all sorts of undesired, nonmodality-specific, side effects. The notion that it was not simply differences in dual task difficulty that accounted for selective interference effects was further supported by the fact that most articulatory suppression condition scores were not even inbetween baseline and spatial tapping scores, but either completely the same or even slightly better than the baseline scores.

An alternative explanation for the interference of spatial tapping, which is hard to refute, is that it originates from similarity between this task and the response mode (pressing the left or right button of a response box). Spatial tapping could hamper a participant because the motor activity related to the interference task slows down the subsequent motor preparation and carrying out of the key-press. This is not the case for articulatory suppression because activity from this task stems from a completely different modality. To avoid the possible confound between spatial tapping and the response mode, we conducted a second experiment in which we changed the response mode from the spatial-motor domain (same for spatial tapping) to the verbal-vocal domain (different for spatial tapping). For Experiment 2 we used a voice key as response mode and we only examined the effect of spatial tapping relative to baseline in a sentence-picture and picture-picture verification task.

Experiment 2

Method

Participants. Twenty-four right-handed participants (6 men, 18 women, all undergraduate students) cooperated, with normal or corrected-to-normal vision. They all gave informed consent and were naive with respect to the hypotheses, and were paid €7 per hour for participating.

Design and procedure. The interference task used was a visual-spatial (spatial tapping) task; the articulatory suppression task was not included in Experiment 2. All other aspects of the design and procedure of Experiment 2 were identical to those used in Experiment 1, except for the response mode (voice key); when S2 was presented participants had to say "goed" or "fout" ("true" and "false" in Dutch) as fast as possible without making any mistakes.

Data analysis. The data of the practice trials were discarded, as were trials on which the RT was either 2.5 standard deviations above or below the mean of the condition. As a result of this criterion, .8% of the trials were considered to be an outlier. The cluster analysis resulted in a visual-spatial group (n = 16) who showed a difference in RTs (M = 17 ms, minimum = -94 ms, maximum = 83 ms, SD = 48.2) between the sentence-picture and picture-picture baseline tasks that was significantly smaller than the difference in RTs (M = 174 ms, minimum = 110 ms, maximum = 377 ms, SD = 85.7) for the verbal group (n = 8), $t(22) = 5.8$, $p < .001$. Mean RTs, computed over correct trials, and mean percentage error scores were analyzed using separate 2 x 2 x 2 ANOVAs with Format (Sentence or Picture) and Interference (Baseline or Spatial Tapping) as within-subjects variables and Group (Verbal or Visual-Spatial) as between-subjects variable.

In addition, an analysis of the different sentences was carried out to establish whether there was a difference between sentences containing "to the left of" and sentences containing "to the right of." Mean RTs, computed over correct trials, and mean percentage error scores were analyzed using separate 2 x 2 x 2 ANOVAs, with Interference (Baseline or Spatial Tapping) and Locative Preposition (Left or Right) as within-subjects variables and Group (Verbal or Visual-Spatial) as between-subjects variable.

Results

Reaction times. The main effect of Format was significant, $F(1, 22) = 52.5$, $p < .001$, MSE = 5,899.2, showing that RTs were faster when the first stimulus was a picture (M = 699 ms) than a sentence (M = 820 ms; see Figure 2). The main effect of Interference was not significant, $F(1, 22) < 1$. The main effect of Group was not significant, $F(1, 22) < 1$.

The interaction between Format and Group was not significant, $F(1, 22) = 3.0$, $p = .1$, MSE = 5,899.2. The interaction between Format, Interference, and Group was significant, $F(1, 22) = 7.1$, $p = .01$, MSE = 7,332.1. Analysis of this interaction revealed that the visual-spatial group showed a significant interaction between Format and Interference, $F(1, 16) = 13.0$, $p = .003$, MSE = 6,807.8, while the verbal group did not, $F(1, 6) < 1$. Further analysis indicated that the visual-spatial group was impaired by spatial tapping when the first stimulus was a sentence, $t(15) = 2.8$, $p = .013$, while the verbal group was not impaired by spatial tapping when the first stimulus was a sentence, $t(7) < 1.1$. No other interactions among Format, Interference, and Group were significant, all Fs < 2.2. The analyses with the variable Locative Preposition (Left or Right) showed no significant effects, all Fs < 1 (see Table 2).

Error scores. The main effect of Format was significant, $F(1, 22) = 15.6$, $p = .001$, MSE = 20.4, indicating that fewer errors were made when the first stimulus was a picture (M = 4.7 %) than a sentence (M = 8.6%). The main effect of Interference was not significant, $F(1, 22) < 1$. The main effect of Group was not significant, $F(1, 22) = 1.4$, $p = .24$, MSE = 56.5.

The interaction between Format and Group was sig-

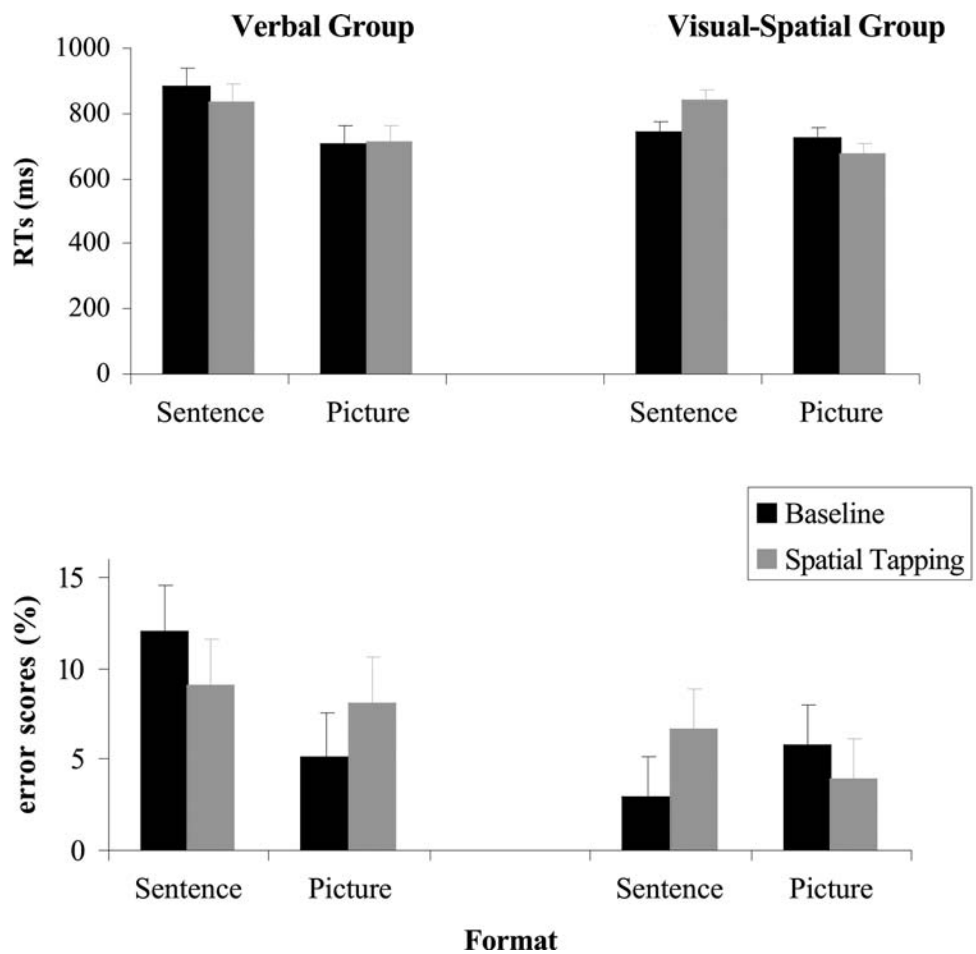

Figure 2. Mean RTs (ms) and error scores (%) for sentence-picture and picture-picture verification tasks in baseline and spatial tapping conditions in Experiment 2. Computation of within-subjects standard errors based on Loftus and Masson (1994).

nificant, $F(1, 22) = 4.3$, $p = .049$, $MSE = 20.4$. Analysis of this interaction showed that the verbal group (M = 10.6%) made more errors than the visual-spatial group (M = 6.6%) when the first stimulus was a sentence, but not when it was a picture, $t(22) = 2.2$, $p = .04$ and $t(22) < 1$. The interaction between Format, Type of Interference, and Group was significant, $F(1, 22) = 12.7$, $p = .002$, $MSE = 14.2$. Analysis of the component interactions revealed that both the visual-spatial group and the verbal group showed a significant interaction between Format and Interference, $F(1, 15) = 5.7$, $p = .03$, $MSE = 15.4$, and $F(1, 7) = 8.3$, $p = .02$, $MSE = 11.6$. Further analysis indicated that the visual-spatial group was impaired by spatial tapping when the first stimulus was a sentence and facilitated by spatial tapping when it was a picture, $t(15) = 2.4$, $p = .03$. In contrast, the verbal group was facilitated by spatial tapping when the first stimulus was a sentence and impaired by spatial tapping when it was a picture, $t(7) = 2.9$, $p = .02$. No other interactions between Format, Interference, and Group were significant, all Fs < 1.

The analyses with the variable Locative Preposition (Left or Right) showed a significant interaction between Type of Interference and Locative Preposition, $F(1, 22) = 6.1$, $p = .02$, $MSE = 26.9$. Further analysis showed that Spatial Tapping (M = 10.0%; Baseline M = 6.3%) lowered performance for sentences containing “left of,” $t(23) = 2.2$, $p = .04$, whereas there was no difference between Baseline (M = 8.8%) and Spatial Tapping (M = 6.7%) conditions if the sentence contained “right of,” $t(23) < 1.1$. No other effects concerning Locative Preposition were significant, all Fs < 1.

TABLE 2
Mean RTs (ms) and Error Scores (%) for Sentence-Picture Tasks in Experiment 2

	RTs (ms)				Error Scores (%)			
Group	Verbal		Visual-Spatial		Verbal		Visual-Spatial	
Locative Preposition	"Left of"	"Right of"	"Left of"	"Right of"	"Left of"	"Right of"	"Left of"	"Right of"
Type of Interference								
Baseline	877	874	738	735	11.1	13.2	3.8	6.6
Spatial Tapping	819	843	838	832	10.4	7.6	9.7	6.3

Note: Computation of within-subjects standard errors amounted to 60 ms and 2.6% for the Verbal, and 39 ms and 2.3% for the Visual-Spatial group (based on Loftus and Masson, 1994).

Discussion

The main focus of Experiment 2 was to determine whether the effect of spatial tapping found in Experiment 1 originated from the specific response mode used in that experiment. In Experiment 2 we also observed diminished performances (i.e., slower RTs and more errors) in comparison with the baseline condition. Hence, the spatial tapping interference obtained in Experiment 1 did not originate from the similarity between the response mode (a left or right key-press) and the dual task (tapping a simple spatial pattern), because the interfering effect of spatial tapping was also present in this second experiment that used a voice key as response mode. Similar to Experiment 1, participants were faster and made fewer errors when the first stimulus was a picture than when it was a sentence. Therefore, the pictorial representation based on a picture seemed to be compared more effectively to a second picture than the pictorial representation based on a sentence.

The division of the participants into two groups in Experiment 2 resulted in a visual-spatial group of sixteen participants and a verbal group of eight participants. Only the visual-spatial group showed a diminished performance in the spatial tapping condition in comparison to the baseline condition. For the visual-spatial group, as became apparent from the second order interaction between Format, Type of Interference, and Group, the interference caused by spatial tapping was limited to the condition in which the first stimulus was a sentence. In contrast, spatial tapping facilitated the performance of the verbal group when the first stimulus was a sentence, but impaired the performance when the first stimulus was a picture. Interestingly, in this experiment the detrimental effect of spatial tapping on the accuracy of participants in the visual-spatial group was confined to trials in which the spatial sentence contained the preposition "left of." We will further elaborate on the issues of the classification criterion, the resulting groups, and the difference between "left of" and "right of" in the general discussion.

General Discussion

This study started from the premise that a linguistic spatial relation may be represented with either a propositional format that is dependent on rehearsal in the phonological loop or a pictorial format that is dependent on rehearsal in the visual-spatial working memory. A spatial and a verbal dual task were employed, which were assumed to load on the visual-spatial working memory and the phonological loop, respectively. We found that the spatial dual task interfered in sentence-picture verification tasks, whereas the verbal dual task showed no effect. Apparently, most participants followed a visual-spatial strategy and not a verbal strategy in the sentence-picture verification task. Experiment 2 ruled out a possible confounding between the spatial dual task and the response mode. The fact that spatial tapping interfered with a sentence-picture verification task suggests that most participants formed a pictorial representation of the first sentence. In both experiments participants performed better when they had to verify a picture after viewing an initial picture than after reading a sentence. This could indicate that the pictorial representation based on the picture was more effectively compared to a following picture than the pictorial representation based on a sentence. The difference in effectiveness might arise because participants made more errors reading and interpreting the sentence than viewing the picture. Taken together, the two experiments provide evidence for the pictorial nature of the representation of linguistic categorical spatial relations in a sentence-picture verification task.

If participants formed a pictorial representation of the spatial sentence then this representation had to be generated and subsequently maintained in working memory. Whether the spatial tapping interfered mostly with the generation or the maintenance process cannot be inferred from the present findings. However, in either case, the relevant processes were clearly spatial. This finding does not fit a model for processing of simple spatial sentences that assumes that the representational format of a spatial sentence is always propositional (e.g., Logan & Sadler, 1996).

Previous research has used RT patterns to demonstrate that individual participants might choose different strategies within a sentence-picture verification experiment (e.g., Kroll & Corrigan, 1981; MacLeod et al., 1978; Russell & Taggert, 1995). To examine the possibility of the employment of different strategies in our experiments, we used a classification criterion based upon the difference between sentence-picture and picture-picture verification RTs. Participants who were about equally fast in both tasks were thought to be using a visual-spatial strategy and participants who were much faster in the picture-picture task than in the sentence-picture task were classified as using a verbal strategy. Participants were divided into two groups with the above-mentioned criteria (using a cluster analysis), and we did indeed find that spatial tapping impaired participants who were classified as using a visual-spatial strategy, while the participants who were classified as using a verbal strategy were not impaired by spatial tapping. The fact that the verbal group was not hampered by articulatory suppression might indicate that this task mostly interfered with the phonological surface structure of the spatial sentence and not with the underlying propositional content. Importantly, spatial tapping did not hamper (and in some cases facilitated) the performance of the verbal group. Therefore, the present study, using a dual-task methodology, provides converging evidence that both separate verbal and visual-spatial strategies exist for the processing of simple spatial sentences (e.g., Glushko & Cooper, 1978; MacLeod et al.; Reichle et al., 2000).

In Experiment 2 there was a significant difference in the effect of spatial tapping on the picture-picture verification task: Spatial tapping impaired the verbal group, but not the visual-spatial group. It may be that the picture-picture task was a passive "image maintenance" task, and that the sentence-picture task was an active "image generation" task. Spatial tapping impaired the verbal group on the image maintenance task, while this dual task impaired the visual-spatial group only on the more demanding image generation task. This could indicate that the verbal and visual-spatial group were not only employing different strategies, but also were composed of participants with different spatial abilities, that is, low (verbal group) and high spatial abilities (visual-spatial group).

Previous studies (e.g., Just & Carpenter, 1975) found that there are differences in the way sentences with different locative prepositions (e.g., "above" vs. "below" and "left of" vs. "right of") are encoded. Usually, sentences with "above" and "right of" were found to be easier to compare to subsequent pictures than sentences with "below" and "left of." In the present study we found no differences between sentences with "left of" and "right of," except for the result that in Experiment 2 the detrimental effect of spatial tapping on accuracy was only found for sentences containing "left of" and not for sentences containing "right of." First, the fact that we found no systematic differences between sentences with different locative prepositions is probably because of the specific types of locative prepositions in our experiment. Just and Carpenter already noted that the above-below asymmetry seems to be more general than the right-left asymmetry. Second, it could be that the spatial tapping task only yielded its deteriorating effects because participants always code "left of" and "right of" in terms of their hands. Hence, the tapping task, which was always executed with the left hand, might have interfered more with the understanding of "to the left of" than with "to the right of" (as we found in Experiment 2). Yet, we did not find this in Experiment 1, where participants also always tapped with their left hand. Importantly, throughout all the analyses no clear effect of "left of" versus "right of" was observed for the present study, and therefore it seems unlikely that tapping works by hampering body-centred prepositional coding.

How does the present conclusion that most participants followed a visual-spatial strategy bear upon previous studies that examined the presence of different strategies in the sentence-picture verification task? These studies (e.g., MacLeod, Hunt, & Mathews, 1978) typically seemed to find that the majority of the participants followed a verbal strategy, while a small subset of participants was classified as having followed a visual-spatial strategy. How can we account for this discrepancy? It could be that the current classification method based on RT differences on sentence-picture and picture-picture tasks might not be reliable, while the method of previous studies (i.e., measuring the effect of linguistic complexity) could be a better way to assign participants to a strategy group. Therefore, it could be that this latter, more reliable, classification method, although it has also been criticized (cf. Roberts et al., 1994) would have resulted in assigning more participants to the verbal strategy group. However, it would still not change the fact that the visual-spatial strategy (as indicated by the effect of spatial tapping) and not the verbal strategy was dominant in the present set-up.

Finally, in order to explain our evidence concerning the dominance of a visual-spatial strategy, we turn to the specific characteristic of our sentence-picture verification task. There are several arguments supporting the view that the current procedure might be ideal for participants to choose a visual-spatial strategy. First, Tversky (1975) found that the effect of a negation in a spatial sentence on picture verification times was dependent on the moment the picture was presented in relation to the sentence: If a sentence and picture were presented simultaneously, a negation lengthened the verification RTs, while if a picture was presented after the sentence disappeared, a negation had no effect. Apparently, participants maintained a verbal strategy for simultaneous comparison, but in case of a delay they changed to a visual-spatial strategy. Given these findings, it is not surprising that we found participants to predominantly choose a visual strategy in the present study, because pictures followed the spatial sentences after a delay. Second, in a recent study (Noordzij, Van der Lubbe, & Postma, in press), we found support that participants choose a visual-spatial strategy when a picture was most likely to follow (80% of the trials a picture, 20% of the trials a sentence) a spatial sentence, while participants did not form a pictorial representation when a sentence was most likely to occur. Therefore, the expected stimulus-modality (context) seems to plays a major role in the availability of a visual-spatial strategy. In the current study a picture always followed a spatial sentence, which would constitute a situation in which participants seem very likely to adopt a visual-spatial strategy. Hence, our findings have relevance for a setting in which linguistic spatial information is compared to visual information. An interesting next step would be to examine the influence of verbal and spatial dual tasks on the processing of a sentence such as “X is to the left of Y” when the ensuing task is not visual but verbal: verifying this spatial relation in a subsequent sentence. If the context in which a spatial sentence is read is indeed a crucial factor with respect to processing strategies, then a spatial dual task should not interfere with the encoding of the sentence in a verbal context, indicating that people predominantly choose a verbal strategy.

In conclusion, the present study demonstrated that most participants consistently formed a pictorial representation of a spatial sentence in a sentence-picture verification task. Furthermore, we provided converging evidence, using a dual-task methodology, that both separate verbal and visual-spatial strategies exist for the processing of simple spatial sentences. Finally, our findings have relevance for a common situation in which people know that they have to use the information in a spatial sentence to verify information in a visual scene. However, further research is required to establish whether different types of processing occur when a spatial sentence is read in a purely verbal context.

We wish to thank Erik Reichle and one anonymous referee for many valuable suggestions on the manuscript. This study was supported by a grant from the Netherlands Organization for Fundamental Research (NWO, No. 440-20-00).

Requests for reprints should be addressed to Matthijs Noordzij, Psychological Laboratory, Utrecht University, Heidelberglaan 2, NL-3584 CS Utrecht, The Netherlands (E-mail: m.noordzij@fss.uu.nl).

References

Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Clarendon Press/Oxford University Press.

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General, 129(1),* 126-145.

Beech, J. R. (1984). The effects of visual and spatial interference on spatial working memory. *Journal of General Psychology, 110(2),* 141-149.

Brooks, L. R. (1967). The suppression of visualisation by reading. *Quarterly Journal of Experimental Psychology, 19,* 289-299.

Chincotta, D., & Underwood, G. (1997). Digit span and articulatory suppression: A cross-linguistic comparison. *European Journal of Cognitive Psychology, 9(1),* 89-96.

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

Eley, M. G. (1981). Distinguishing imagery from propositional recoding processes in sentence-picture verification tasks. *Canadian Journal of Psychology, 35(3),* 254- 269.

Evans, J. (1989). *Bias in human reasoning: Causes and consequences*. Hove & London: Lawrence Erlbaum Associates Ltd.

Glushko, R. J., & Cooper, L. A. (1978). Spatial comprehension and comparison processes in verification tasks. *Cognitive Psychology, 10(4),* 391-421.

Just, M. A., & Carpenter, P. A. (1975). The semantics of locative information in pictures and mental images. *British Journal of Psychology, 66(4),* 427-441.

Kroll, J. F., & Corrigan, A. (1981). Strategies in sentence-picture verification: The effect of an unexpected picture. *Journal of Verbal Learning & Verbal Behavior, 20(5),* 515-531.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476-490.

Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garret (Eds.), *Language and space* (pp. 493-529). Cambridge, MA: MIT Press.

Logie, R. H. (1995). *Visual-spatial working memory*. Hove, UK: Lawrence Erlbaum Associates, Inc.

MacLeod, C. M., Hunt, E. B., & Mathews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning & Verbal Behavior, 17(5),* 493-507.

Marquer, J., & Pereira, M. (1990). Reaction times in the study of strategies in sentence-picture verification: A reconsideration. *Quarterly Journal of Experimental Psychology, 42A(1),* 147-168.

Mathews, N. N., Hunt, E. B., & MacLeod, C. M. (1980). Strategy choice and strategy training in sentence-picture verification. *Journal of Verbal Learning & Verbal Behavior, 19*, 531-548.

Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective*. Pacific Grove, CA: Brooks/Cole Publishing Company.

Milner, A. D., Jeeves, M. A., Ratcliff, P. J., & Cunnison, J. (1982). Interference effects of verbal and spatial tasks on simple visual reaction time. *Neuropsychologia, 20(5),* 591- 595.

Noordzij, M. L., Van der Lubbe, R. H. J., & Postma, A. (in press). Strategic and automatic components in the processing of linguistic spatial relations. *Acta Psychologica*.

Oakhill, J. V., & Johnson-Laird, P. N. (1984). Representation of spatial descriptions in working memory. *Current Psychological Research & Reviews, 3(1),* 52-62.

Quinn, J. G., & McConnel, J. (1996). Irrelevant pictures in visual working memory. *Quarterly Journal of Experimental Psychology, 49A,* 200-215.

Reichle, E. D., Carpenter, P. A., & Just, M. A. (2000). The neural bases of strategy and skill in sentence-picture verification. *Cognitive Psychology, 40(4),* 261-295.

Roberts, M., Wood, D. J., & Gilmore, D. J. (1994). The sentence-picture verification task: Methodological and theoretical difficulties. *British Journal of Psychology, 85,* 413-432.

Russell, J., & Taggart, A. (1995). One model or two? Sentence verification in children and adults. *Educational Psychology, 15(4),* 379-404.

Seymour, P. H. (1974). Generation of a pictorial code. *Memory & Cognition, 2(2),* 224-232.

Smyth, M. M., & Pelky, P. L. (1992). Short-term retention of spatial information. *British Journal of Psychology, 83(3),* 359-374.

Sussman, H. M. (1982). Contrastive patterns of intrahemispheric interference to verbal and spatial concurrent tasks in right-handed, left-handed and stuttering populations. *Neuropsychologia, 20(6),* 675-684.

Tversky, B. (1975). Pictorial encoding of sentences in sentence-picture comparison. *Quarterly Journal of Experimental Psychology, 27*, 405-410.

Sommaire

Les relations spatiales simples peuvent être représentées soit sous forme de propositions dépendantes de la répétition verbale soit en tant qu'images entretenues par répétition visuelle-spatiale. La présente étude visait à utiliser des tâches doubles, qui entravent de manière sélective les composantes verbales ou spatiales de la mémoire de travail, afin d'établir le type de représentation, propositionnelle ou picturale, employée à la vérification des relations spatiales. Nous nous sommes reportés à des tâches de vérification faisant appel aux paires phrase-image et image-image afin d'examiner l'effet d'une tâche double de suppression articulatoire et de tapotement spatial sur l'encodage de relations spatiales simples désignées par les expressions « à gauche de » et « à droite de » (par ex., le triangle à gauche du cercle). De plus, nous avons tenu compte de la possibilité que différents participants utilisent diverses stratégies pour effectuer la tâche de vérification phrase-image. Selon les critères établis par Eley (1981), les participants dont les temps de réaction étaient comparables face à des tâches de types phrase-image et image-image étaient classés parmi ceux qui pratiquaient une stratégie visuelle-spatiale, tandis que ceux dont le temps de réaction face à une tâche de catégorie phrase-image étaient plus longs que celui que leur demandait une tâche image-image entraient dans le groupe de ceux qui possédaient une stratégie verbale.

Dans la première expérience, la suppression articulatoire ne constituait pas une entrave, tandis que le tapotement spatial se traduisait par une performance moindre aux deux tâches. Il semblerait que les intrants tant linguistiques que perceptuels des relations spatiales simples mobilisaient la mémoire de travail visuelle-spatiale. Lors des expériences de vérification du jumelage phrase-image, le tapotement spatial nuisait seulement au rendement des participants dont la tendance des temps de réaction les classait parmi ceux qui utilisaient une stratégie visuelle-spatiale, tandis qu'il n'avait aucun effet sur la performance de ceux dont on disait qu'ils appliquaient une stratégie verbale. De plus, les délais d'exécution étaient plus courts et les erreurs moins nombreuses lorsque le premier stimulus était une image plutôt qu'une phrase. Vraisemblablement, les participants réussissaient mieux à comparer à une image une représentation picturale

fondée sur une image plutôt qu'une représentation visuelle-spatiale fondée sur une phrase. La représentation picturale de la phrase aurait pu donner de moins bons résultats lors d'une comparaison ultérieure à une image du fait que les participants commettaient plus d'erreurs en lisant et en interprétant la phrase qu'ils ne le faisaient en regardant une image.

Afin d'éviter une éventuelle confusion entre le tapotement spatial et le mode de réaction, nous nous sommes livrés à une deuxième expérience pour laquelle nous avons substitué un mode de réaction du domaine verbal-vocal (qui diffère pour le tapotement spatial) à celui du domaine spatial-moteur (identique pour le tapotement spatial). Lors de la deuxième expérience, les performances relatives à la condition du tapotement spatial étaient moindres (c.-à-d. temps de réaction plus longs et plus grand nombre d'erreurs) en comparaison de la condition de référence. Par conséquent, l'entrave découlant du tapotement spatial observée lors de la première expérience n'était pas le fait de la similitude entre le mode de réaction (appuyer sur une touche à gauche ou à droite) et la tâche double (tapoter un motif spatial simple), car l'effet d'entrave du tapotement spatial a également été constaté lors de la seconde expérience dont le mode de réaction était une clé vocale. Il suit que l'étude offre des preuves convergentes, sur la base d'une méthode à doubles tâches, de l'existence de stratégies distinctes verbales et visuelles-spatiales appliquées au traitement de phrases spatiales simples.