Since the seminal study of Nissen and Bullemer (1987), the serial reaction time (SRT) task has been one of the preferred paradigms for researchers interested in the study of implicit learning. In a typical study using the SRT task, participants are instructed to respond to the position of a stimulus (such as a dot or an asterisk), appearing at one of a number of fixed locations on a computer screen. Unbeknownst to the participants, the order in which the stimuli appear in the locations follows a fixed sequence. In these sequenced blocks, reaction times (RTs) decrease reliably as the training progresses. When the task is changed such that the position of the stimulus is determined randomly, RTs increase. The difference between average RT on such a random block and the preceding sequence block can be taken as a measure of learning of the sequence (Cleeremans, Destrebecqz & Boyer, 1998).

Nissen and Bullemer (1987) manipulated attentional load by comparing a single-task condition in which participants performed only the SRT task with a dual-task condition in which a tone-counting task was performed in addition to the SRT task. One of the key findings of Nissen and Bullemer was that when the tone-counting task was performed concurrently with the SRT task, learning of the repeating sequence did not seem to occur. They interpreted this as an indication that attention is necessary for the sequence to be learned. Cohen, Ivry and Keele (1990) also examined the effects of a secondary task on sequence learning and controlled for the structure of the sequence. Cohen et al. distinguished between stimulus locations that can be predicted by the immediate predecessor (unique or first-order associations) and stimulus locations that cannot be predicted only from one preceding element (ambiguous associations).\(^5\) For example, in the sequence 123243, 2 after 1 and 3 after 4 are unique associations because position 1 always predicts position 2 and position 4 always predicts position 3. The other associations are ambiguous because given the first element in any of them, there is not certainty as to what position will follow (2

\(^5\) In the case of ambiguous associations, when two preceding elements are needed to make a correct prediction we can talk of second-order associations; third-order associations if three elements are needed and so forth.
Sequence Learning

can be followed by either 3 or 4 while 3 can be followed by 2 or 1). In this last case, the participant cannot code the transition as a pairwise association and more than one previous element must be remembered to make a successful prediction (4-3 must be remembered to predict 1 for example). Cohen et al. constructed three types of sequences: unique, ambiguous and hybrid, composed respectively of unique, ambiguous or both types of associations. In a series of experiments, they showed that the concurrent performance of a secondary task affected the learning of unique and hybrid sequences (there was less learning in comparison with a single-task condition) but only eliminated the learning of the ambiguous sequence.

To explain these results, Cohen et al. (1990; see also Curran & Keele, 1993) hypothesized that two different mechanisms underlie learning in the SRT task: One that can operate under distraction and forms associations between adjacent items and another that requires attention and builds a hierarchic code that permits the learning of ambiguous sequences. The idea is that parts of the ambiguous sequence that have previously occurred in the same order are recognized within working memory and stored as subgroups. A secondary task interferes with learning because it disrupts working memory and prevents recognition of parts of the sequence and their grouping into higher hierarchies. From this follows that working memory is necessary for learning the sequence and for expressing what has been learned. Consistently, Curran and Keele found that participants learning under single-task conditions did not perform differently from participants that had learned under dual-task conditions if the testing conditions involved an attentional load.

The learning of hybrid sequences poses a challenge to the dual mechanism hypothesis because, although these sequences contain ambiguous associations, they can be learned under dual-task conditions. Curran and Keele (1993; see also Keele & Jennings, 1992) suggested that the hybrid sequences are learned by the associative system, which uses the unique associations as cues that define points to start parsing the sequence. Therefore, in the presence of a secondary task, working memory is

---

6 The idea of hierarchical coding comes from studies in motor learning that show that sequences of movements are commonly represented in chunks or subgroups. Evidence of such grouping has been found in the form of pauses between the subgroups, patterns of errors and transfer of performance (Gordon & Meyer, 1987; Povel & Collard, 1982). Povel and Collard give a simple but clarifying example: If a person has to dial the number 654456 we can expect different latencies in the response depending on whether the internal representation is (65)(44)(56) or (654)(456). Building a hierarchical code implies making chunks with the sequence and recoding every chunk at a higher level.

7 Cohen et al. (1990) talked of short-term memory (STM) whereas Curran and Keele (1993) used the more modern term of working memory (WM). This different terminology reflects the change in focus, from the passive construct of STM to the more dynamic concept of WM. In that sense and following Conway, Engle and colleagues (Conway et al., 2005), we favored the measure of WM span as indicative of WM capacity, understood as the maintenance of relevant information in the face of active, ongoing processing.
loaded and the hierarchical coding mechanism does not recognize previous parts of the sequence and cannot recode them in subsections. As a result, an ambiguous sequence cannot be learned under distraction. The hybrid sequence, on the other hand, can be learned because the associative, attention-independent system uses the unique associations as starting point to parse the sequence.

If constructing a hierarchic code requires the involvement of working memory, it can be predicted from the dual-mechanism hypothesis that working memory capacity (WMC) will correlate with learning only when the hierarchical-coding system is at work. Thus, a measure of WMC should correlate with learning under single- but not dual-task conditions, where the secondary task interferes with the hierarchical-coding system. The prediction that introducing a secondary task will decrease or eliminate any correlation between learning in the SRT task and WMC is also consistent with the general theory of implicit learning put forward by Reber (1993). Reber suggested that the range of phenomena related to the cognitive unconscious (implicit memory, implicit learning, subliminal perception, etc.) is based on a primitive system that developed before the appearance of conscious functioning. One of the predictions of this theory is that implicit learning is not influenced by individual differences factors like age or measures of intelligence, whereas explicit learning is. Whereas the correlation between explicit learning and factors such as intelligence is well established (Carroll, 1993), evidence for a correlation between IQ and implicit learning has been contradictory. For example, Reber, Wakenfield and Hernstadt (1991) showed that IQ correlated with performance in an explicit series-completion problem-solving task but did not correlate with learning of an artificial grammar, and Gebauer and Mackintosh (2007) found that measures of intelligence correlated with performance in implicit learning tasks only under explicit rule discovery instructions and not under standard implicit instructions. On the other hand, Fletcher, Maybery and Benett (2000) compared gifted and disabled children on explicit and implicit tasks and observed similar correlations between measures of learning and IQ in both types of task.

Working memory capacity has a clear relation with attention and explicit processing (Miyake & Shah, 1999; Baddeley, 2003). Most models of WM contain some mechanism that manages the allocation of attention like the central executive of Baddeley and Hitch (1974), the supervisory attention system of Norman and Shallice (1980) or the focus of attention of Cowan (1995). Independently of whether this “attention manager” allocates resources to peripheral stores (Baddeley) or constitutes one single system that activates a subset of long-term memory (Cowan), WMC is likely to reflect individual differences in the control of attention and will be related
Sequence Learning

with controlled, explicit processing but not with automatic, implicit processing (Barrett, Tugade & Engle, 2004; Conway & Engle, 1996).

Therefore, working memory capacity should correlate with explicit learning but not with implicit learning performance (see also Reber, 1993). Indeed, Unsworth and Engle (2005), in a comparison of people with high or low working memory span on a SRT task, found differences between the two groups when they were given intentional instructions but not when they followed standard, incidental instructions. In contrast, Frensch and Miner (1994) found significant correlations between learning and short-term memory capacity when the SRT task was combined with a tone-counting task but not when the SRT task was performed alone. Although Frensch and Miner’s results seem to contradict the prediction that adding a secondary task will interfere with a hierarchical-coding mechanism that involves working memory, it is important to notice that they measured STM capacity and not WM capacity. Although STM and WM are highly related constructs, STM capacity reflects domain-specific storage, whereas WM capacity reflects domain-general executive attention (Conway et al., 2005). It is this more general, domain-independent capacity to control and allocate attentional resources (captured in WM span tasks but not in STM span tasks) that we expect to correlate with learning in the SRT task when attention is not loaded with a secondary task.

The dual-mechanism hypothesis (Cohen et al, 1990; Curran & Keele, 1993) assumes that the attention-dependent system builds a hierarchical code that permits learning the ambiguous sequences. In the absence of attention, hybrid sequences may still be learned by the associative, attention-independent mechanism that parses the sequence using the unique associations as cues. It seems reasonable to expect that the associative system will first learn the unique associations because this is where the parsing of the sequence will start. Then, the ambiguous associations will also be learned. We would thus predict faster responses to unique than to ambiguous associations early in training under dual-task conditions. If unique associations are responded to faster than are ambiguous associations, learning, as indexed by the introduction of a random block, might be greater too. Under single-task conditions, unique associations might also be initially responded to faster than ambiguous associations: If the hierarchical coding system requires attention, we might expect this to be an effortful process that will be preceded by the attention-independent associative system, which will start learning the unique associations. However, when with time and practice the higher code is built and the sequence is integrated as a whole, the difference between types of associations should disappear.
To our knowledge, only one study (Frensch, Buchner & Lin, 1994) has directly tested the prediction that attentional load may interact differently with the learning of unique and ambiguous associations within a hybrid sequence. In this study, one of the hybrid sequences of Cohen et al. (1990) was practiced for eight 120-trial blocks under single or dual-task conditions, followed by transfer blocks in which one type of association (unique or ambiguous) within the sequence was modified. The longer RTs in transfer, as compared to sequence blocks, when either unique or ambiguous associations were changed, revealed that both types of associations were learned. Both task structure and attentional load influenced learning such that unique associations were learned better (i.e., showed a greater difference in RT between sequence and transfer blocks) than ambiguous associations and participants in the single-task condition showed more learning than participants in the dual-task condition did. Importantly, however, there was no interaction between task load and type of association, leading Frensch and colleagues to conclude that only a single, associative learning mechanism operates under both single- and dual-task conditions and on unique and ambiguous associations.

Although task load and type of association did not interact in Frensch et al.’s (1994) study, the difference in RT to the two types of association during training was greater in the single-task condition than in the dual-task condition (15 vs. 8 ms, Frensch et al., Experiment 1). This is inconsistent with the hypothesis that the introduction of a secondary task will increase the difference between unique and ambiguous associations by virtue of having a greater effect on ambiguous associations. The finding of Frensch et al. might be the result of the way the secondary task was implemented. Contrary to common practice (e.g., Cohen et al., 1990; Jimenez & Vazquez, 2005; Nissen & Bullemer, 1987; Stadler, 1995), Frensch et al. did not give participants in the dual-task condition feedback about their accuracy in the tone counting task. Normally, feedback is given and participants are encouraged to keep their accuracy over 90%. This lack of feedback during the tone-counting task might have been the reason that over 30% of the participants in the dual-task conditions did not meet the minimum accuracy criterion and were excluded from the analyses. The remaining 70% of participants may have differed from the excluded participants. Participants who, without being encouraged to do so, still manage to keep a high level of accuracy in the secondary task might have shared some characteristic such as higher motivation or a higher working memory capacity than

---

8 This is a high number of participants excluded. A short literature review of 12 experiments in 6 different studies showed that, on average, 9.4% of participants were excluded. Only two experiments showed 22% and 26% of participants excluded respectively (Shanks & Channon, 2002).
Sequence Learning

those excluded. That would imply that single- and dual-task conditions would have differed not only in the experimental manipulation, but also in the measure of this individual differences variable.

Another unusual finding in Frensch et al. (1994) is the relatively lower error rate of the single-task condition in the SRT task itself (2.7 % and 2.1 % in the single-task conditions vs. 7.0 % and 5.8 % in the dual-task conditions). Error rates in the SRT task are normally fairly low and the analyses are centered on the more relevant RT data. Most studies do not report whether there are differences between the conditions in the error rate, but when differences are reported, they consistently show a higher error rate for the single-task condition (Frensch, Lin & Buchner, 1998; Jimenez & Mendez, 1999; Shanks & Channon, 2002; Stadler, 1995). The tendency towards a higher error rate in the single-task condition can be attributed to a greater proportion of incorrect anticipation responses (Schvaneveldt & Gomez, 1998). The higher error rate of the dual-task conditions in Frensch et al. might have had a relation with the lack of feedback in the tone-counting task: The participants from the dual-task condition who performed with accuracy over 90% in the tone-counting task may have done so at a cost to the SRT task.

Given that the findings of Frensch et al. (1994) may have been influenced by selection error, we revisited the question of whether two learning processes operate concurrently in the learning of hybrid sequences. The study specifically examined the effect of a secondary task on the learning of unique and ambiguous associations within a sequence. Moreover, we included a test of working memory span to examine the relation of working memory capacity and sequence learning. It was hypothesized that there would be an interaction between attentional load and type of association, both during training and in terms of sequence learning, and that working memory capacity would correlate with learning under single-task but not under dual-task conditions.

**EXPERIMENT 1**

Experiment 1 was conducted to test the hypothesis that the introduction of a secondary task would have more effect on the learning of ambiguous than on unique associations. The experimental manipulation involved the comparison of single- and dual-task conditions on a SRT task. The single-task condition performed only the primary SRT task while the dual-task condition performed the SRT task and, in addition, a tone-counting task. Working memory span was measured to test the prediction that working memory capacity may correlate with learning under single-but not under dual-task conditions.
Method

Participants

Fifty-nine undergraduate students (39 female) from the University of Groningen participated in the study as part of a course requirement. The mean age was 21.6 years (range 18 to 52, \( sd = 5.2 \)). Participants were randomly assigned to either a single- or a dual-task condition. After eliminating six participants (19%) who made more than 10% of tone-counting errors and one participant who could not finish the experiment because of a software problem, there were 26 participants in each condition.

Stimuli and apparatus

The experiment was run on IBM-compatible personal computers with 51-cm color monitors (model iiyama MA203DT vision master pro 513) and standard keyboards. Stimuli were presented and responses were registered with a program written in E-prime (Schneider, Eschman & Zuccolotto, 2002).

The stimulus display consisted of four black unfilled rectangles 3 cm high by 2 cm wide on a white background, separated by 2 cm and arranged horizontally in the center of the computer screen. The imperative stimulus was a black dot with a diameter of 1 cm. On each trial, the dot appeared inside one of the rectangles and the participant’s task was to press a key on the keyboard corresponding spatially to its position. Responses were performed with the index and middle fingers of each hand resting on the “z”, “x”, “n” and “m” keys, for outer-left, inner-left, inner-right and outer-right positions, respectively.

The location of the stimulus on the screen was determined by one of six sequences, based on those used by Cohen, et al. (1990, Experiments 3 and 4): 123243, 123134, 142312, 132412, and 423213, where the numbers 1, 2, 3, and 4 correspond to the four locations from left to right. Each sequence is composed of 2 associations in which a number can be completely predicted by its immediate predecessor (unique associations) and 4 associations in which two predecessors have to be considered to make a successful prediction (ambiguous associations). Sequences were randomly assigned to participants.

The test used to measure working memory capacity was the automated version of the reading span task (Conway et al., 2005). This test, as other versions of the reading span task (Daneman & Carpenter, 1980) requires participants to recall sets of stimuli of varied length (3 to 7 elements) presented sequentially on the screen while performing a concurrent task. Each set size was presented three times, for a total of 75 trials. The stimuli were letters and the concurrent task to read sentences in English and
judge whether they made sense or not. Thus, a normal trial of the automated reading span task started with the presentation of the sentence for the sense judgment task. After the participant gave a response, a letter was presented for 800 ms followed by the next sentence (with no interval between the two tasks). When a complete set of letters had been presented, the participant was prompted to recall them in order. The method used to calculate the score was the sum of all perfectly recalled sets, as recommended by Conway et al.

**Procedure**

The experiment started with two practice blocks of 32 trials each followed by 12 training blocks of 120 trials each. Every block started with a randomly selected stimulus location. Subsequent stimuli followed the assigned sequence except in blocks 9-10 and during the practice trials in which all stimulus positions were chosen randomly with the constraint that successive presentations of stimuli at the same location were not allowed.

Participants were instructed to respond as quickly and accurately as possible. Stimuli remained in view for 3 s or until a response key was pressed. If the allowed time was exceeded a box appeared for 750 ms centered on the screen with the text “Too slow” (presented in Dutch). If an error was committed, the word “Error” (in Dutch) was displayed in a box centered on the screen for 750 ms. The next stimulus appeared 250 ms after a response or after the offset of the feedback.

In the dual-task condition, a 2000- or 1000-Hz tone was presented for 250 ms after a randomly selected interval of 40, 80 or 120 ms following the response on correct trials. Participants were instructed to count the high tones presented during the block and to report this number at the end of each block. Participants entered their responses using the keyboard and were informed of their accuracy via a feedback display that encouraged them to keep their accuracy above 90%.

Following the SRT task, a generation task was performed under inclusion and exclusion conditions. In this task, participants were informed that the stimuli for the...
SRT task had followed a fixed short sequence that repeated for most of the experiment. They were then asked to generate a series of keypresses that resembled the trained sequence (inclusion condition). During this task, the screen showed the four reference rectangles and when any given key was pressed, the dot appeared inside the corresponding location and stayed there until another key was pressed. There was no time limit for the task and participants were instructed to avoid repetitions of the same keypress. After the generation task was performed under inclusion instructions, the same task was performed again under exclusion instructions. In this case, participants were instructed to generate a series of trials while avoiding the reproduction of the trained sequence. That is, on any given trial they were to select a location where the next stimulus would not have been presented during the training. In both generation conditions, participants were told to guess if they could not remember exactly what the sequence was. Each block of the generation task consisted of 120 trials. The session ended with the test of working memory capacity.

Results

SRT task

Accuracy in the SRT task was above 95% under single- and dual-task conditions (see Figure 3.1). A mixed ANOVA on the proportion of correct responses with condition (single- or dual-task) as a between-subjects factor and block (1 to 12) as a within-subjects factor showed a main effect of condition ($F(1, 50) = 5.44, p = 0.024, \eta^2_p = 0.098$) reflecting lower accuracy under single- than under dual-task conditions (96.78 % and 97.73 % respectively). The effect of block was also significant ($F(11, 550) = 10.34, p < 0.001, \eta^2_p = 0.171$) but there was no Condition x Block interaction ($F(11, 550) = 1.16, p = 0.326, \eta^2_p = 0.023$).
Figure 3.1. Accuracy in the SRT task as a function of condition (single- or dual-task) and block (1 to 12).

The responses to the predicted element of the association were used for the analysis of RT as a function of type of association. That is, in a sequence like 123243, the RT to unique associations is calculated from the RTs to elements 2 after 1 and 3 after 4, while the RT to ambiguous associations comes from the RTs to all the other positions. Error trials, responses exceeding 3 s and the first five trials of each block were excluded in the computation of the mean RT per block (6.7 % of trials were filtered out). Figure 3.2 shows the average RT as a function of block, condition and type of association.
Mean RT data were submitted to a mixed ANOVA with condition (single- or dual-task) as a between-subjects factor, and type of association (unique or ambiguous) and block (pre-transfer sequence blocks 1-8) as within-subjects factors. The results showed significant effects of condition \((F(1, 50) = 22.29, p < 0.001, \eta^2 = 0.308)\), such that RTs were faster under single- than under dual-task conditions (304 ms vs. 403 ms), type of association \((F(1, 50) = 10.74, p = 0.002, \eta^2 = 0.177)\), such that responses to unique associations were faster than those to ambiguous associations (345 ms vs. 361 ms); and block \((F(7, 350) = 52.38, p < 0.001, \eta^2 = 0.512)\) which reflects the general decrease in RT across blocks (from 439 ms in Block 1 to 297 ms in Block 8). There was also a Type of Association x Condition interaction \((F(1, 50) = 6.01, p = 0.018, \eta^2 = 0.107)\), which indicates that the average difference between unique and ambiguous associations within the dual-task condition (28 ms) was greater than that within the single-task condition (3.9 ms). The Block x Type of Association \((F(7, 350) = 1.97, p = 0.079, \eta^2 = 0.038)\), the Block x Condition \((F(7, 350) = 0.98, p = 0.421, \eta^2 = 0.019)\) and the Condition x Block x Type of Association interactions \((F(7, 350) = 1.43, p = 0.209, \eta^2 = 0.028)\) were not significant.

Extra support for the involvement of two types of learning in the SRT task is evident in the comparison of the learning rates between conditions and within association. For this purpose, power-law functions were fit per participant to the mean
Sequence Learning

RT of Blocks 1 to 8 separately per condition and type of association. Data were log-log transformed and fit to a linear model per subject to extract the individual slopes, which can be considered a measure of the steepness of the learning curve (i.e., rate of learning). The variance explained by the model (averaged across participants) was 50%, 46%, 62% and 62% for unique associations in the single-task condition, ambiguous associations in the single-task condition, unique associations in the dual-task condition and ambiguous associations in the dual-task condition respectively. The resultant slopes were submitted to a mixed ANOVA with condition (single-task or dual-task) as between-subjects factor and type of association (unique or ambiguous) as a within-subjects factor. The only significant effect was that of condition ($F(1, 50) = 5.52, p = 0.023, \eta^2 = 0.099$) which showed that participants in the single-task condition learned at a faster rate than did participants in the dual-task condition (average slopes were -0.255 and -0.164 for single-task and dual-task conditions respectively). There was no significant effect of type of association ($F(1, 50) = 1.49, p = 0.229, \eta^2 = 0.029$) nor was there a Condition x Type of Association interaction ($F(1, 50) = 2.22, p = 0.143, \eta^2 = 0.042$), although planned post-hoc comparisons showed that the learning rates differed between associations in the dual- (unique = -0.182, ambiguous = -0.144; $t(25) = -4.38, p < 0.001$) but not in the single-task condition (unique = -0.253, ambiguous = -0.256; $t(25) = 0.142, p = 0.888$).

Regarding learning of the sequence (defined as the difference in RT between pre-transfer block and transfer block), Figure 3.2 shows that transfer to the random sequence seemed to produce a greater deterioration of performance under single- than under dual-task conditions. In order to test the degree of learning in the SRT task as a function of type of association and condition, the difference in RT between blocks 9 and 8 was calculated per participant and submitted to a mixed ANOVA with condition (single-task or dual-task) as a between-subjects factor and type of association (unique or ambiguous) as a within-subjects factor. The results showed a significant effect of condition ($F(1, 50) = 22.01, p < 0.001, \eta^2 = 0.306$), reflecting that participants in the single-task group learned more than participants in the dual-task group did (209 ms and 101 ms respectively). There was no main effect of type of association ($F(1, 50) = 1.58, p = 0.215, \eta^2 = 0.031$). Finally, there was a Condition x Type of Association interaction ($F(1, 50) = 8.29, p = 0.006, \eta^2 = 0.142$) which shows that the secondary task produced a greater interference in the learning of ambiguous than unique associations.

**Generation task**
Generation task data were examined for evidence of differential implicit and explicit knowledge of the sequence in the single- and dual-task conditions. Scores were computed as the proportion of generated pairs (chunks of two elements) from the trained sequence out of the total number of pairs produced. Because the last element in any given association is the first element in the next, the total number of possible pairs was \(120 - 1 = 119\). Figure 3.3 shows the proportion of pairs from the trained sequence generated during the generation task as a function of condition and type of test.

![Figure 3.3](image)

*Figure 3.3.* Average proportion of pairs of elements from the trained sequence as a function of condition and generation test. Error bars depict the standard errors of the means. The dotted horizontal line represents chance level.

The total number of possible combinations of two (avoiding repetitions) from a set of four elements is 12. From these 12 pairs, six were present in the training sequence, thus the probability of producing a pair from the trained sequence by chance is 0.5. Generation scores were compared to chance performance. Inclusion scores were significantly above chance level for single-task \((t(25) = 9.90, p < 0.001)\) and dual-task conditions \((t(25) = 4.27, p < 0.001)\). Exclusion scores were significantly below chance for single-task condition \((t(25) = -2.47, p = 0.021)\) but not for dual-task condition \((t(25) = -1.39, p = 0.178)\).
Figure 3.4. Average proportion of pairs of elements from the trained sequence as a function of condition, generation test and type of associations. Error bars depict the standard errors of the means. The dotted horizontal lines represent chance level.

Generation scores were calculated separately for unique and ambiguous associations (see Figure 3.4). Because there were twice as many ambiguous as unique associations in the sequences, the probabilities of producing by chance unique and ambiguous pairs are 0.17 and 0.33 respectively. Comparisons of the proportion of generated pairs to chance performance by means of t-tests showed that participants in both the single- and dual-task conditions produced more unique (single: $t(25) = 5.27, p < 0.001$, dual: $t(25) = 2.98, p = 0.006$) and ambiguous associations (single: $t(25) = 11.83, p < 0.001$, dual: $t(25) = 3.56, p = 0.002$) than expected by chance during the inclusion test. In the exclusion test, only unique associations in the single-task condition were produced below chance ($t(25) = -2.45, p = 0.022$). T-tests comparing the conditions in the production of pairs in the inclusion test were significant for both unique ($t(50) = 2.39, p = 0.021$) and ambiguous associations ($t(50) = 5.94, p < 0.001$). No differences between the conditions were found in the exclusion test in any type of stimuli.

**Working memory test**

The score of the working memory task was calculated as the sum of all perfectly recalled sets. Because working memory is conceived as a system responsible for maintenance of information in the face of ongoing processing (Conway et al., 2005), it is important to ensure accuracy in the processing component of the test, represented
by the judgment task. Thus, Conway et al. recommend excluding all participants with less than 85% accuracy in the concurrent sentence judgment task. Our participants found especially difficult the task and 20 of them failed to perform at the required level of accuracy. The reason of this high percentage (38.5%) of failure in the relatively simple judgment task might be that the great majority of participants in our sample were Dutch nationals. The use of English in the concurrent task probably added an element of extra difficulty thus, to compensate for the effects of using a foreign language, the minimum accuracy level was lowered to 80%. With this level of accuracy, 18 participants (69%) remained in the analyses from the single-task condition and 21 (81%) from the dual-task condition. We calculated the correlation between working memory capacity and the measure of learning (difference between blocks 8 and 9) in single- ($r = -0.006, p = 0.981$) and dual-task ($r = -0.178, p = 0.439$) conditions. Neither of the correlations reached significance.

**Discussion**

As predicted, performing a secondary task during the SRT task interfered more with the learning of ambiguous than unique associations. This is consistent with the dual-mechanism hypothesis (Cohen et al., 1990; Curran & Keele, 1993). If the secondary task disrupts a hierarchical coding mechanism and all the learning is carried out by the associative system, then the unique associations will be learned first because they serve as cues to start parsing the sequence. The differential impact of the secondary task on unique and ambiguous associations was evident during training in the greater difference between associations under dual- than under single-task conditions. Learning, as measured by the difference in RT between sequence and random blocks, was also differentially disrupted by the secondary task, with the learning of ambiguous associations being reduced more than the learning of unique associations. Although the difference between the learning rates of unique and ambiguous associations was markedly greater in the dual- than in the single-task condition, the analysis of the learning rates did not replicate the interaction effect between type of association and condition found in the analysis of the training RTs.

The analysis of the data from the generation task showed that both conditions resulted in the acquisition of a similar amount of implicit knowledge and that the general differences in performance were due to more explicit knowledge (especially regarding ambiguous associations) in the single-task condition. It seems that the introduction of a secondary task mainly impairs the development of explicit knowledge of ambiguous associations.
Finally, the prediction that working memory capacity would correlate with learning under single-task conditions was not confirmed. As explained, this could have been caused by the small number of observations left because participants had difficulties performing the test in English. To avoid this problem, in the next experiment the working memory test was changed to a non-verbal task. The training phase was also extended and several transfer blocks introduced to test learning at different moments. The expectation was that the main results of Experiment 1 would be replicated, such that the effect of a secondary task would be greater on ambiguous than on unique associations, that participants in both single- and dual-task conditions would develop implicit and explicit knowledge and that the advantage in learning for the ST condition would be explained by differences in explicit knowledge. Moreover, extending the training phase allowed further testing of the dual-mechanism hypothesis (Cohen et al., 1990; Curran & Keele, 1993). For example, with extended training the difference between types of association under dual-task conditions might remain. This could be interpreted as proof that the associative system is incapable of learning perfectly the ambiguous associations without the help of a hierarchical coding system. However, with extended training the difference between associations could disappear. This second outcome could indicate that the associative system can learn perfectly the ambiguous associations with extended practice, but it could also be due to an automatization of the secondary task. Experiment 2 was designed to address these questions.

EXPERIMENT 2

Experiment 1 showed that the introduction of a dual-task load disrupts the learning of ambiguous associations more than that of unique associations, which are learned similarly regardless of the attentional manipulation. This finding seems to support the involvement of two systems in sequence learning during the SRT task: an associative system that requires minimal attention and learns unique associations, and an attention-dependent system that builds a hierarchical code to help learning ambiguous associations. Experiment 2 was designed to test whether this difference in learning between unique and ambiguous associations might disappear with extended training. If the difference between the learning of unique and ambiguous associations under dual-task conditions disappeared, this could be indicative of something about the way the associative system can or cannot deal with ambiguous associations; but it could also be due to an automatization of the secondary task (Shanks, Rowland & Ranger, 2005). That is, if the secondary task became automatic, then it would no longer divert attentional resources and the attention-demanding hierarchical-coding system would
be available, as under single-task conditions. To check for this possibility, the training phase was extended in Experiment 2 and the secondary task was changed to a symbol-counting task, (see, e.g., Jimenez & Mendez, 1999) in which the imperative stimulus for the SRT task was one of four symbols, and one of these symbols was to be counted for the counting-task. The new format of the secondary task allowed the use of the same stimulus for the SRT task and the counting task and, moreover, the introduction of two dual-task conditions: One in which the mapping between stimulus of the SRT task and the target of the secondary task was consistent (dual-task/consistent) and should become automated, and one in which this mapping was varied (dual-task/varied) to avoid automatization of the secondary-task (Schneider & Shiffrin, 1977).

Finally, to test whether automatization of the secondary-task indeed took place, a search task was introduced at regular intervals during the training. Based on Schneider and Shiffrin (1977), it is reasonable to expect that the change of target stimulus from block to block should interfere with the automatization of the secondary-task and that this difference in automaticity should be evident in a search-task. The prediction is that, as training progresses, the number of distractors in the search display will influence performance to a lesser extent in the case of the dual-task/consistent condition because of automatization.

**Method**

**Participants**

Participants were 101 undergraduate students (76 female) enrolled in a first-year psychology course at the University of Groningen. They participated in the study for course credit. The mean age was 20.69 years (range 17 to 52, \( sd = 2.89 \)). Participants were randomly assigned to one of the three conditions: single-task, dual-task/consistent and dual-task/varied. Participants with overall symbol-counting accuracy below 90% or accuracy below 70% on two or more blocks of the SRT task were excluded from further analyses. With this criterion, four participants from each dual-task condition were excluded for low accuracy in the secondary task (11.1 % and 11.4 %, in dual-task/consistent and dual-task/varied conditions respectively). The final distribution was 30 participants in the single-task condition, 32 in the dual-task/consistent condition and 31 in the dual-task/varied condition.

**Stimuli and apparatus**
Sequence Learning

The apparatus and sequences were as in Experiment 1. However, the random presentations in the transfer blocks were substituted by transfer sequences. Every participant was assigned randomly to use one sequence during the training blocks and a different sequence during transfer.

The stimulus display consisted of four rectangles of 3 cm high and 2 cm wide separated by 2 cm and arranged horizontally in the center of the computer screen, as in Experiment 1. However, the imperative stimulus was any of the four stimuli resulting from the combination of the symbols “□” and “△”, and the colors green and magenta. The symbols were 1 cm high and 1.2 cm wide and appeared centered inside the squares. The symbols were sampled randomly; each stimulus appeared a roughly equal number of times inside each of the four rectangles. Thus, although the stimuli positions followed the sequence the stimuli themselves did not. Participants were instructed to respond to the position of each symbol by pressing the spatially corresponding key.

The working memory task was changed to avoid the language problems encountered during Experiment 1. The test used was the symmetry span task (Conway et al., 2005). In this task, participants are to remember sequences of colored squares on a grid (see Figure 3.5). As in the reading span task, there are 5 different set sizes that range from 3 to 7 elements. Every set size is presented a total of three times, thus the test consists of 75 trials, and participants must perform a concurrent task after each presentation which consists of deciding whether a geometric figure is symmetric respect to its vertical axis or not (see Figure 3.5). A trial of the symmetry span task started with the presentation of the figure for the symmetry judgment task. Once the participant had responded, a grid was presented for 800 ms with one of the squares colored in red. This was followed by the next figure for the symmetry judgment task. When a complete set of squares had been presented, the participant was prompted to recall them in order.

11 Concerns have been expressed regarding the use of random presentations during the transfer blocks to measure sequence learning (Reed & Johnson, 1994). For this reason we decided to change the presentation of the stimuli during the transfer blocks from random to an alternative sequence.
Figure 3.5. Screenshots of the computer on a given trial during presentation of the stimulus for the main memorization task (left panel) and the figure for the symmetry task (right panel)

Procedure

The experiment began with one block of practice of the search task (12 trials), followed by the first block of the search task. Then participants began the SRT task with two practice blocks of 30 trials each followed by 24 training blocks. During the training blocks with the SRT task, participants followed the assigned sequence except in Blocks 6, 14 and 22 where the transfer sequence was presented. The search task was presented after training Blocks 8, 16 and 24. Each SRT and search task block consisted of 120 trials. After the last block of search, participants performed a generation task as in Experiment 1. The experimental session ended with a test to measure working memory span, as in Experiment 2.

In the practice blocks of the SRT task, the stimuli appeared in random locations with the restriction that repetitions were avoided. Each block of the SRT task started with the presentation of a stimulus in a randomly selected location. The subsequent stimuli followed the assigned sequence. Participants were instructed to respond as quickly and accurately as possible and a maximum of 3 s was allowed for responses. When the time was exceeded a message with the text “Too slow” was displayed in a box centered on the screen for 750 ms. Incorrect responses were signaled with a 150-Hz tone for 250 ms. The following stimulus appeared 250 ms after a response or after the offset of the feedback.

In the single-task condition, participants were instructed simply to respond to the location of the stimulus by pressing the corresponding key and to ignore the type of symbol. In the dual-task/consistent condition, participants performed the SRT task
as in the single-task condition and, in addition, were to count the number of times that one of the four possible symbols appeared. The target symbol for the counting task was the same in all blocks of the SRT task for participants in the dual-task/consistent condition. In the case of the dual-task/varied condition, participants also performed the SRT task and had to count symbols as in the dual-task/consistent condition, but with the difference that the target symbol for the counting task changed on every block. The target symbol for the counting task was indicated at the beginning of every block and, at the end of the block, a message box prompted the participant for the response. After the response was entered, participants received feedback about the counting task and about the SRT task in two consecutive message boxes that stayed centered on the screen for 2 s each. The feedback about the counting task was given in terms of percentage correct and a message encouraged the participant to maintain accuracy above 90%. The SRT-task feedback gave the number of errors committed during the block and either (1) congratulated the participant if errors had been made in less than 5% of the trials, (2) encouraged the participant to avoid making errors in the next block, if errors had been made in more than 5% of the trials or (3) encouraged the participant to try to respond faster, if no errors had been made. Averaged across blocks and participants, the target symbol was presented 29.98 times per block (sd = 0.91).

The search task consisted in searching for one of the stimuli used in the SRT task on a display of distractors (the other stimuli from the SRT task). The target was indicated at the beginning of the block and participants were instructed to press the key “1” if the target was present or the key “2” if it was not. Each trial started with the presentation of a fixation point centered in the screen for 250 ms, followed by the search display. Correct responses ended the trial and incorrect responses were signaled with a 250-ms tone. The target was the same in all four blocks of the search task and, in the case of the dual-task/consistent condition, coincided with the target of the counting task. The displays contained 8, 16 or 32 distractors and were presented with a program written in E-prime (Schneider et al., 2002) that applied the following constraints: The three display sizes were presented an equal number of times (40 times each), the target was present in half of the trials of each display size, half of the distractors shared color with the target and the other half shared shape, the distractors were equally distributed along the four quadrants of the screen, and no distractor was placed within 1.5 cm of the border of the screen or within 3 mm of another distractor.

**Results**

*SRT task*
The single-task condition consistently had worst accuracy across all blocks than the dual-task conditions did (see Figure 3.6). The mean proportion of correct responses averaged across all blocks of training was 97%, 98% and 98% for ST, dual-task/consistent and dual-task/varied respectively. A mixed ANOVA with condition (single-task, dual-task/consistent or dual-task/varied) as a between-subjects factor and block (blocks 1 to 24 of the SRT task) as a within-subjects factor showed main effects of condition ($F(2, 90) = 12.95, p < 0.001, \eta^2 = 0.223$) and block ($F(23, 2070) = 21.7, p < 0.001, \eta^2 = 0.194$) but no Condition x Block interaction ($F(46, 2070) = 1.39, p = 0.088, \eta^2 = 0.030$). Tukey’s post hoc tests showed that the single-task group had lower accuracy than the dual-consistent group ($p < 0.001$) and the dual-varied group ($p < 0.001$), while these last two did not differ among them ($p = 0.912$).

![Figure 3.6. Accuracy in the SRT task as a function of condition (single-task, dual-task/consistent or dual-task/varied) and block (Blocks 1 to 24). The vertical lines mark the transfer blocks (Blocks 6, 14 and 22).](image)

Reaction times above three standard deviations over the mean (computed per condition) were excluded from the analyses, as were error trials and trials following an error and the first five trials of each block. These filters led to the exclusion of 3.8%, 3.3% and 3.2% of data in the single-task, dual-task/consistent and dual-task/varied conditions, respectively. In the analysis of the RTs we used, as in the previous experiment, the response to the predicted element within each association. Figure 3.7 shows the mean RT per condition, block and type of association during the SRT task.
Figure 3.7. Mean RT in the SRT task as a function of condition (single-task, dual-task/consistent or dual-task/varied), block (Blocks 1 to 24) and type of association (unique associations or ambiguous associations). The vertical lines mark the transfer blocks (Blocks 6, 14 and 22).

As in the previous experiment, learning was defined as the difference in RT between the transfer blocks and the immediately preceding sequence block. Because there were three blocks of transfer, learning was measured at three points. Separate mixed ANOVAs on learning with condition (single-task, dual-task/consistent or dual-task/varied) as a between-subjects factor and type of association (unique associations or ambiguous associations) as a within-subjects factor were carried out. Only the analysis at the first learning point showed a significant effect of condition ($F(2, 90) = 6.49, p < 0.002, \eta^2 = 0.126$). The effects of type of association ($F(1, 90) = 1.24, p = 0.269, \eta^2 = 0.014$) and the Type of Association x Condition interaction ($F(2, 90) = 1.57, p = 0.215, \eta^2 = 0.034$) were non-significant. Tukey’s post-hoc comparisons showed that participants in the single-task condition acquired more sequence knowledge than those in the dual-task/consistent (103 ms vs. 52 ms, $p = 0.007$) and dual-task/varied (103 ms vs. 50 ms, $p = 0.006$) conditions. No differences were found between the two dual-task conditions ($p = 0.994$). The lack of interaction shows that the amount of learning did not vary between conditions as a function of type of association.

The next analysis was designed to check whether the secondary task had a different impact during training for unique than ambiguous associations. A mixed ANOVA was carried out on average RT with condition as a between-subjects factor and type of association (unique or ambiguous) and block (the blocks of the SRT task
excluding transfer and block after transfer, N = 17) as within-subjects factors. There were significant main effects of condition ($F(2, 90) = 35.81, p < 0.001, \eta^2 = 0.443$), which reflects that the single-task condition produced on average the fastest responses (179 ms, 313 ms and 331 ms for the single-task, dual-task/consistent and dual-task/varied conditions respectively); block ($F(16, 1440) = 415.99, p < 0.001, \eta^2 = 0.822$) due to a progressive decrease in RT along training; and type of association ($F(1, 90) = 111.14, p < 0.001, \eta^2 = 0.553$) reflecting that the responses to unique associations were on average faster than the responses to ambiguous associations (263 ms vs. 286 ms). Finally, there were also interactions between block and condition ($F(32, 1440) = 1.822, p = 0.053, \eta^2 = 0.039$), block and type of association ($F(16, 1440) = 2.07, p = 0.029, \eta^2 = 0.022$) and block, type of association and condition ($F(32, 1440) = 2.19, p = 0.003, \eta^2 = 0.046$). The Block x Type of Association x Condition interaction shows that, as in Experiment 1, there was a greater impact of the secondary task on ambiguous than on unique associations.

In order to investigate further the Block x Type of Association x Condition interaction, we calculated the mean difference between RT to unique and ambiguous associations per training block, and averaged across groups of blocks (averages of blocks 1-5, 8-12 or 16-20). ANOVAs with condition (single-task, dual-task/consistent or dual-task/varied) as a between-subjects factor revealed that the difference in RT between responses to unique and ambiguous associations did not vary between conditions during Blocks 1 to 5 ($F(2, 92) = 1.26, p = 0.287$), and during Blocks 8 to 12 ($F(2, 92) = 2.46, p = 0.092$). The conditions differed though during Blocks 16 to 20 ($F(2, 92) = 3.27, p = 0.043$) were Tukey’s post-hoc tests showed that the difference between RT to unique and ambiguous associations was marginally smaller in the single- than in the dual-task/consistent (11.8 ms vs. 25.9 ms, $p = 0.064$) and the dual-task/varied (11.8 ms vs. 25.3 ms, $p = 0.082$) conditions, whereas participants in the two dual-task conditions did not differ among them (25.9 ms vs. 25.3 ms, $p = 0.995$).

As in Experiment 1, power-law functions were fitted per subject to the data shown in Figure 3.7. The data was log-log transformed and fitted to a linear model per subject to extract the individual slopes. These individual learning rates (see Table 3.1 for the slopes) per type of association were submitted to a mixed ANOVA with condition (single-task, dual-task/consistent and dual-task/varied) as a between-subjects factor and type of association (unique or ambiguous) as a within-subjects factor. Results showed a main effect of condition ($F(2, 90) = 27.99, p < 0.001, \eta^2 = 0.384$) and a main effect of type of association ($F(1, 90) = 15.97, p < 0.001, \eta^2 = 0.151$) but no interaction ($F(2, 90) = 0.59, p < 0.552, \eta^2 = 0.013$). Tukey’s post-hoc tests showed that the main effect of condition was due to significant differences
Sequence Learning

between the means of single-task and dual-task/consistent conditions (-0.573 vs. -0.307, \(p < 0.001\)) and the means of single-task and dual-task/varied conditions (-0.573 vs. -0.349, \(p < 0.001\)), but not between the two dual-task conditions (\(p = 0.513\)). The main effect of type of association shows that, on average, the learning rate of unique associations was faster than the learning rate of ambiguous associations (-0.424 vs. -0.395). The analysis of the learning rates shows that the secondary task slowed equally the learning of both types of association.

Last, we also fitted power-law functions across three different sections of training (five blocks each) to check for possible changes in the learning rate and the influence of introducing testing blocks. The resultant slopes are shown in Table 3.1.

Table 3.1. Rates of learning as a function of condition, type of association and section of training.

<table>
<thead>
<tr>
<th></th>
<th>17 Blocks</th>
<th>Blocks 1-5</th>
<th>Blocks 8-12</th>
<th>Blocks 16-20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-task unique</td>
<td>-0.582</td>
<td>-0.229</td>
<td>-0.297</td>
<td>-0.198</td>
</tr>
<tr>
<td>Single-task ambiguous</td>
<td>-0.562</td>
<td>-0.221</td>
<td>-0.281</td>
<td>-0.138</td>
</tr>
<tr>
<td>Dual-task/consistent unique</td>
<td>-0.321</td>
<td>-0.102</td>
<td>-0.167</td>
<td>-0.136</td>
</tr>
<tr>
<td>Dual-task/consistent ambiguous</td>
<td>-0.293</td>
<td>-0.067</td>
<td>-0.157</td>
<td>-0.131</td>
</tr>
<tr>
<td>Dual-task/varied unique</td>
<td>-0.368</td>
<td>-0.077</td>
<td>-0.183</td>
<td>-0.144</td>
</tr>
<tr>
<td>Dual-task/varied ambiguous</td>
<td>-0.329</td>
<td>-0.053</td>
<td>-0.149</td>
<td>-0.104</td>
</tr>
</tbody>
</table>

The learning rates were compared through an ANOVA with condition as a between-subjects factor and section of training (blocks 1-5, blocks 8-12 or blocks 16-20) and type of association as within-subjects factors. The results, as with the comparison of the global slopes, showed main effects of condition (\(F(2, 90) = 12.80, p < 0.001, \eta^2 = 0.222\)) and type of association (\(F(1, 90) = 19.07, p < 0.001, \eta^2 = 0.175\)) that will not be investigated further. More interesting is the main effect of section of training (\(F(2, 180) = 7.00, p = 0.001, \eta^2 = 0.072\)) which was caused by differences in the learning rate between first and second sections (-0.125 vs. -0.206, \(p = 0.001\)) and differences between second and third sections (-0.206 vs. -0.142, \(p = 0.010\)). No differences in learning rates were found between first and third sections of training (-0.125 vs. -0.142, \(p > 0.05\)). None of the interactions was significant. The comparison of the learning rates shows that the highest amount of learning took place immediately after the first transfer block.

**Generation task**
As in Experiment 1, the dependent variable in the generation task was the proportion of generated chunks of two elements from the trained sequence. Figure 3.8 presents these proportions across condition, generation test and type of association.

The proportion of generated pairs (unique and ambiguous associations) from the trained sequence was compared to chance performance. The probability of producing a unique pair by chance is 0.17 and the probability of producing an ambiguous pair by chance is 0.33. When compared with chance performance, all conditions produced more pairs (unique and ambiguous) than expected by chance during the inclusion test, and none of them differed from chance during the exclusion test. It seems that the differences between the conditions were related mainly with the scores during the inclusion test.

The scores of the inclusion test were submitted to a mixed ANOVA with condition (single-task, dual-task/consistent and dual-task/varied) as between-subjects factor and type of association (unique or ambiguous) as within-subjects factor. The results revealed main effects of condition \( (F (2, 90) = 7.82, p = 0.001, \eta^2 = 0.148) \) and type of association \( (F (1, 90) = 518.42, p < 0.001, \eta^2 = 0.852) \) and a Condition x Type of Association interaction \( (F (2, 90) = 11.68, p < 0.001, \eta^2 = 0.206) \). Follow-up
ANOVAs showed differences between the conditions in the production of ambiguous associations ($F (2, 90) = 12.48, p < 0.001$), but not unique ($F (2, 90) = 1.35, p = 0.264$). Tukey’s post-hoc tests showed that the single-task condition produced more ambiguous pairs during the inclusion test than dual-task consistent (59.2 % vs. 46.2 %, $p = 0.001$) and dual-task/varied (59.2 % vs. 43.1 %, $p < 0.001$). No differences were found between the two dual-task conditions ($p = 0.615$). The results of the generation task are consistent with the results from Experiment 1 and can be interpreted as evidence of an advantage in explicit knowledge of ambiguous associations in the case of the ST condition respect to both dual-task conditions.

**Search task**

Figures 9a and 9b present the results of the search task regarding mean reaction time and mean accuracy.

![Graphs showing search task results](image)

*Figure 3.9a. Mean proportion accuracy as a function of condition, block of search and set size.*
As it was previously mentioned, the search task was introduced as a means to check the progressive automatization of the secondary task. Figure 3.9b shows that the lines became increasingly flatter across blocks, which means that, as training progressed, the number of distractors influenced less the time to detect the target. The prediction was that, if the secondary task were indeed becoming automatic, this flattening of the lines would be more severe in the dual-task consistent condition. A mixed ANOVA on the average RT with search block (search blocks 1 to 4) and set size (8, 16 and 32 distractors) as within-subjects factors and condition (single-task, dual-task consistent and dual-task/varied) as a between-subjects factor showed significant effects of block \((F(3, 270) = 219.16, p < 0.001, \eta^2 = 0.709)\), set size \((F(2, 180) = 689.12, p < 0.001, \eta^2 = 0.884)\) and a Block x Set size interaction \((F(6, 540) = 63.88, p < 0.001, \eta^2 = 0.415)\). No differences between conditions \((F(1, 90) = 0.86, p = 0.428, \eta^2 = 0.019)\) were found, nor were any of the interactions involving the factor condition significant. The main effect of block shows that RTs decreased significantly across the four blocks (average RTs were 1592.23, 1382.45, 1293.27 and 1285.24 in blocks 1 to 4 respectively). Pairwise comparisons showed that the decrease in RT was
significant between blocks 1-2 (p < 0.001) and blocks 2-3 (p < 0.001) but not between blocks 3-4 (p = 0.397). The main effect of set size indicates that the time of search increased as it did the number of distractors (1018 ms, 1346 ms and 1799 ms for 8, 16 and 32 distractors respectively). The interaction shows that the practice effect, consequence of repeating the task in four occasions was more acute as the set size increased (average difference in RT between blocks 1 and 4 was 165, 288 and 465 ms for set size 8, 16 and 32 respectively). The lack of influence of the experimental manipulation in the results of the search task reveals that the dual-task did not become less automatic as a result of changing the mapping. Analysis with the data regarding accuracy in the search task showed a similar pattern of results.

**Working memory test**

As in the previous experiment, we calculated the correlation between the measure of learning (at the first transfer block) and working memory span. The correlation was marginally significant in the case of the single-task condition ($r(30) = 0.34, p = 0.066$). No significant correlations were found in the dual-task consistent ($r(30) = -0.29, p = 0.116$) nor dual-task/varied ($r(73) = 0.29, p = 0.104$) conditions. Because correlation tests are highly dependent on the number of observations, we aggregated the data from both experiments. The correlation between working memory capacity and learning was significant in the single-task condition ($r(48) = 0.39, p = 0.006$) but not significant in the dual-task condition ($r(51) = 0.05, p = 0.744$). Figure 3.10 shows the relation between the two variables for single-task and dual-task conditions.

---

12 In order to calculate the correlation between working memory span and sequence learning in the dual-task group, the data from the dual-task/varied condition in Experiment 2 were excluded. The reason is that the change of target for the secondary task could have made participants in this condition process information in a different way than participants in the other two dual-task conditions did. Nevertheless, if their data is aggregated, the results do not change significantly and show a non-significant correlation between working memory capacity and learning ($r(82) = 0.13, p = 0.247$).
Figure 3.10. Relation between working memory span and learning. The left panel presents the aggregated data from the single-task conditions and the right panel the aggregated data from the dual-task conditions.

The correlation between working memory span and learning was also calculated separately for unique and ambiguous associations. The correlations in the single-task condition were significant in the case of unique ($r(48) = 0.29$, $p = 0.040$) and ambiguous associations ($r(48) = 0.42$, $p = 0.003$). No correlation reached significance in the dual-task condition (unique: ($r(51) = 0.07$, $p = 0.622$), ambiguous: ($r(51) = 0.03$, $p = 0.830$)).

Discussion

The second experiment showed, as did Experiment 1 that learning of both unique and ambiguous sequence elements can occur under attentional load. The effect of an attentional load was a general increase in RT, and this increase seemed to be more pronounced in the case of ambiguous associations. Although the difference in RTs during the training phase between the two types of associations was greater in the case of the dual-task conditions than the single-task condition, ambiguous associations were also responded to more slowly than were unique associations in the single-task condition. This fact is compatible with the dual-mechanism hypothesis: Assuming that both mechanisms work together in the case of single-task conditions, we might expect the associative system to start working earlier because it does not require effortful processing whereas the hierarchical-coding system requires identification and storage in working memory of preceding elements in the sequence. However, with extended
training, the hierarchical system slowly helps to represent the sequence as a whole (i.e., in a *verbatim representation*, to use the term of Stadler, 1992) and the difference in RT between the two types of associations tends to disappear (as occurred in the single-task condition). In the case of the dual-task conditions, the lack of attentional resources impairs the use of the hierarchical system and the difference in RT to unique and ambiguous associations remains constant even after 24 blocks of training.

The pattern of results in the generation task was similar to that of Experiment 1. All subjects produced more pairs from the original sequence when asked to do so (inclusion conditions) than when asked the opposite (exclusion), which shows some degree of control over the knowledge acquired. Furthermore, participants in the single-task condition produced more ambiguous pairs than those in the two dual-task conditions but only in the inclusion test. This can be interpreted as an advantage in explicit knowledge regarding ambiguous associations (Destrebecqz & Cleeremans, 2003) However, it is not clear why the superior explicit knowledge of the single-task condition with respect to the two dual-task conditions only led to a difference in learning at the time of the first transfer block.

An intuitive explanation for this lack of difference in learning between the conditions in the second and third transfer blocks would be that the secondary-task became automatic, such that it no longer made demands on the cognitive system and freed participants in the dual-task conditions to devote attention to the SRT task. However, the fact that there were no differences between the dual-task/varied and dual-task consistent conditions during the search task seems to counter this explanation. There was some degree of automaticity developed during training, as revealed by the decreasing effect of set size as a function of block of search, but this did not differ between conditions. An alternative explanation is that the first transfer block acted as a cue that pushed participants into a more explicit, hypothesis-testing mode of processing, and that event eliminated the differences between single-task and dual-task conditions. This prediction can be derived from theories in skill acquisition (Schmidt & Bjork, 1992) that hypothesize an increase in explicit processing as a consequence of an increase in the difficulty of the task. It would also be supported by the analysis of the learning rates that shows that all participants learned at their fastest rate during the period following the first transfer block. However, it cannot be shown from the data that the reason for the superior learning after the first transfer block was an increase in explicit processing.

---

13 Considering that the consistent result of introducing the transfer blocks is an increase in the error rate and an increase in the average RT for all participants, it can be safely assumed that the task becomes more difficult during transfer relative to the sequence blocks.
Overall, the results in Experiment 2 partially replicate those of Experiment 1. To some extent, the introduction of a secondary task affected differentially the responses to unique and ambiguous associations, dual-task participants learned the training sequence and showed that part of that knowledge was conscious, and participants in the single-task condition seemed to develop more explicit knowledge over ambiguous associations.

**GENERAL DISCUSSION**

In two experiments with the SRT task, we found that the influence of a load on attention interacted with the structure of the sequence. The secondary task seemed to affect especially the learning of ambiguous associations and, more specifically, the explicit knowledge developed in the task regarding these associations. A measure of working memory capacity correlated with learning only in the case of the single-task condition and the correlation was higher with ambiguous than with unique associations. Together, the results might be taken as evidence for the existence of two separate systems that process information in the SRT task.

It seems clear that it is possible to learn unique and ambiguous associations under conditions of attentional load. This was shown by Cohen et al. (1990) who hypothesized that the associative, non-attentional system might be able to learn hybrid sequences because the unique associations serve as a cue to start parsing the sequence. Although the secondary task affected both types of associations, ambiguous associations were disrupted the most, as revealed by the significant interactions between type of association and condition during the training blocks in Experiments 1 and 2. It was also in the explicit knowledge related with ambiguous associations were more differences between the conditions were found in the generation task. Introducing a secondary task disrupted learning, slowed down the RT during training and deteriorated the explicit knowledge developed during the SRT task. Moreover, in all these effects, the ambiguous associations were affected the most. This differential reaction of the two types of association to the attentional load can be interpreted as evidence for the existence of two systems that vary in their attentional requirements.

The correlations between working memory capacity and learning can give extra information about the nature of the mechanisms involved in sequence learning. Working memory capacity has been related with attention and control (Baddeley, 1993; Cowan, 2000; Kane, Conway, Hambrick & Engle, 2007) while implicit learning has been hypothesized to be independent of individual differences (Gebauer & Mackintosh, 2007; Reber, 1993; Reber & Allen, 2000). In our experiments, people learning under dual-task conditions showed tighter distributions in their learning
scores than people learning with full availability of attention. Furthermore, the measure of working memory capacity only correlated with learning in the case of the single-task condition, and this correlation was greater with ambiguous associations. It seems that lack of attention rendered working memory ineffective and that when working memory was available, it was in the learning of ambiguous associations were it exerted more influence.

The disruption of a hierarchical coding mechanism hypothesis, as described by Cohen et al. (1990) and Curran and Keele (1993) posits the existence of two mechanisms of implicit learning, one that depends on attention and is related with working memory and another that does not. Under dual-task conditions, only the associative, attention-independent system can work and it learns ambiguous associations because it uses the unique associations as cues to start parsing the sequence. However, it could equally be hypothesized that the attentional, hierarchical-coding system is not completely disabled by the attentional manipulation and some residual activity is responsible for the learning of the ambiguous associations. This would give way to a less strict version of the dual-mechanism hypothesis in which the effects of the attentional manipulation are graded, and affect both systems (Heuer & Schmidtke, 1996). A graded version of the hypothesis implies that the secondary task might interfere with both systems. This is supported by the finding that the secondary task also interferes with the learning of unique associations, as shown in the measure of learning and the learning rates. This interpretation is consistent with the view expressed by Stadler (1997) in which there is only one single implicit learning mechanism, which is non-hierarchical. Explicit learning also takes place during the SRT task, and this is the second mechanism that constructs a hierarchical representation, remembers the context for a given element, requires attention and relates to working memory. The product of the explicit learning mechanism is some amount of knowledge that can be consciously controlled or explicit knowledge. This conclusion is consistent with the main findings in the SRT task and explains why both conditions developed some control about their knowledge and why the differences in learning were mainly related with differences in explicit knowledge. However, it leaves unanswered the question of the relation between type of association and type of knowledge.

Implicit and explicit knowledge were operationalized in the present study in terms of conscious control. Knowledge that cannot be controlled is implicit (Berry & Dienes, 1993; Stadler, 1997; Squire, 2004); therefore, all the pairs from the trained sequence produced under exclusion conditions represent unconscious knowledge. If the conditions do not differ in their scores during the exclusion test, then the assumption is that they acquired a similar amount of implicit, unconscious
knowledge. We can then hypothesize that: 1) if knowledge in both types of association can be processed implicitly and explicitly, and 2) the main differences between conditions relate to explicit knowledge, especially of ambiguous associations; then a reason for this difference might be that there is more explicit knowledge involved with ambiguous than with unique associations. That is, unique and ambiguous associations might be thought of as differing in the amount of information necessary to predict their last element. The idea can be better understood if a unique association is described in the following terms: “IF 4 occurs, THEN get ready for 3”. In this rule, the conditional part constitutes a description that can be conceived as a statement, that is, declarative knowledge. In that sense, an ambiguous association would require more declarative knowledge because more details would have to be added to the conditional. For example, a typical ambiguous association would be: IF 2 (preceded by 1) occurs, THEN get ready for 3. The two examples can be described in more schematic terms: 4 \rightarrow 3 (unique) and 1, 2 \rightarrow 3 (ambiguous). We can see that the conditional part in the case of the ambiguous association contains more information than the unique does and that this information relates not just to the number of elements to be considered but also to their temporal relation.

A last interesting finding is the speed-up in the learning rate after the first transfer block in Experiment 2. The literature of learning shows that the biggest gains during the process of mastering any type of skill occur early and the learning curve typically takes the form defined by a power-law function. It is counterintuitive in that sense that the learning rate becomes faster when the training is well under way. The fact that the speed up occurred after the first transfer block might indicate a relation between both events. The literature in skill acquisition (Bjork, 1994; Schmidt & Bjork, 1992) suggests that the introduction of difficulties in training might push the learner into an explicit mode of processing information. If we assume that: 1) the transfer blocks involve an increase in the difficulty of the task (as measured by RT and error rate) and 2) explicit processing takes place to some extent during the SRT task as consistently shown in both experiments by the results of the generation task, it can then be hypothesized that the transfer blocks produced an increase in explicit processing. This hypothesis would explain both, the speed up after transfer and the lack of difference in learning between the conditions in the second and third transfer blocks.

Although the assumptions of the process dissociation procedure (Jacoby, 1998) applied for our generation task are not free of critics, we reckon that those critics affect the quantitative estimations of the conscious and unconscious components. Therefore, we did not calculate the parameters and based instead our conclusions in the comparison of inclusion vs. exclusion performance only (see Destrebecqz & Cleeremans, 2001, 2003; Neal & Hesketh, 1997; Shanks et al., 2005). Based on that comparison, it can be concluded that the conditions did not differ in their amount of implicit knowledge but they did differ in the amount of general, and thus, explicit knowledge.
blocks. The data from the present experiments however does not convey the necessary evidence and further research will be needed to support this claim.

In conclusion, the present study extends our knowledge about sequence learning and the type of knowledge developed during the SRT task in several ways. First, we show that explicit knowledge can be acquired under attentional load conditions. Second, we show that the advantage in learning for the single-task condition might be related with a superior explicit knowledge related with ambiguous associations. Third, we show that the impact of the secondary task can be found not just in the learning but also in the progression through training, a finding that counters previous results by Frensch, et al. (1994). Fourth, we show that working memory capacity correlates with learning under single but not under dual-task conditions. Last, our data suggests that the introduction of transfer blocks to measure learning might have an effect in learning per se. This is not something new, as tests have long been considered powerful learning events (Bjork, 1994). The relevance here is that if the learning promoted by the introduction of transfer blocks involves an increase in explicit processing, then any explicit measure of learning taken after the end of the SRT task (e.g., generation tasks, explicit questions, etc.) could be biased and show more explicit learning than what was actually acquired before transfer took place. Further research might be designed to address the impact of the transfer blocks in learning and, if extra knowledge is acquired, analyze its nature.