Errors, feedback and attentional load
Guzmán-Muñoz, F.J.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2009

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
A number of manipulations, such as type of feedback and scheduling of training, have been shown to facilitate transfer to new contexts and stimuli. Bjork and colleagues (Bjork, 1994; Schmidt & Bjork, 1992) argued that a common effect of such manipulations is to introduce extra difficulties in the learning process. The consequence of encountering extra difficulties is that the learner pays more attention and engages in extra activities like feedback processing, error detection and encoding of multiple elements of the learning situation (Schmidt & Bjork). Extra information processing activities during learning might lead to elaborated memory representations that encode extra information of the learning situation. More elaborated memory representations should provide a greater number of retrieval routes, thus facilitating transfer. Because elaborated, transferable representations are characteristic of declarative memory (Eichenbaum & Cohen, 2001; Tulving, 1985), it can be argued that manipulations that increase the difficulty of learning also promote the involvement of declarative memory in the learning process.

Because declarative memory encodes relations between elements of the learning situation (Eichenbaum & Cohen, 2001), the hypothesis that introducing difficulties during learning promotes processing through the declarative memory system can be supported if we show that manipulations that increase the difficulty of training produce better memory for relations. Extra support for this hypothesis can be found if manipulations that interfere with declarative memory are shown to be especially detrimental for the encoding of relational representations. One paradigm that has been extensively used to study the learning of relations is implicit learning of sequences. Although implicit learning does not, by definition, depend on declarative memory, declarative memory has been shown to be needed for learning some type of associations. I first review research in implicit learning and, especially, studies that have manipulated attentional availability as a mean to interfere with explicit processing and declarative memory. The chapter continues with a review of research in support of the view that introducing difficulties during learning promotes elaborated processing. Finally, individual differences variables that might relate to
different types of LTM are proposed. The chapter ends with the description of the research questions that will be addressed in the empirical studies.

**Learning Through the Non-Declarative System**

Declarative memory is also called explicit memory or conscious memory because we are normally aware of the contents of this memory system, as opposed to the contents of implicit, non-declarative memory. Attention is involved in the stages of encoding, storing and retrieving declarative memory representations (Squire & Kandel, 2000). It has been argued that the more attention we pay to an item at the encoding stage, the greater the chance that that item will be successfully retrieved (Craik & Lockhart, 1972). Non-declarative memory seems to be less dependent on attention and many non-declarative memories, as those resulting from conditioning, perceptual-motor skills or habits can be acquired automatically (Schacter & Tulving, 1994).

A common way to study knowledge that can be acquired independently of declarative memory is through implicit learning studies. Implicit learning is said to have occurred when a person evidences having learned about the structure of the stimuli without intending to do so and without being able to verbalize the knowledge acquired (Dienes & Berry, 1997). Because of the incidental character of implicit learning, it has been proposed that it does not require attention and occurs as a side effect of any information-processing activity (Berry & Dienes, 1993; Cleeremans & Jimenez, 2002). In an implicit learning experiment, participants interact with a relatively complex, rule-governed environment without being informed of the existence of such rules. After a period of training, learning is assessed through indirect measures (e.g., time to respond) and direct measures (such as asking for the rule that governs the stimuli). When asked, participants typically fail to describe the rules that govern their behavior, even when learning is evident from their reaction times. Such **dissociations** between direct and indirect measures have been traditionally used as evidence of implicit learning. For example, in a typical sequence learning experiment using the serial reaction time (SRT) task, the participant is required to respond to a target stimulus that appears on each trial at one of several locations on the screen. Conditions in which target locations follow a sequence are compared with conditions in which the target is presented randomly. The difference in average reaction time between participants that respond to a sequence and participants that respond to random locations is used as indirect measure of learning. A dissociation is found (i.e., Nissen & Bullemer, 1987) when participants following a sequence show greater decrease in reaction times than those reacting to random locations but perform at chance level in direct measures such as free recall. Dissociations between direct and
indirect measures have also been found in artificial grammar learning (Reber, 1967) and control of complex systems (Berry & Broadbent, 1984).

Some concerns have been expressed regarding the use of dissociations as evidence of implicit learning. Perruchet and Amorin (1992) have argued that the reason dissociations are found is due to a lack of sensitivity of the direct measures. Using direct measures that targeted fragments of the sequence trained during the experiment, Perruchet and Amorin found a higher degree of explicit knowledge than previously reported. Such findings have been replicated (Perruchet & Pacteau, 1990; Vokey & Brooks, 1992) and the current view is that explicit processing and declarative memory have a greater involvement in the SRT task than previously hypothesized (Cleeremans, Destrebecqz & Boyer, 1998; Shanks, 2005). Shanks and St John (1994) also criticized the reliance on dissociations as evidence of implicit learning and argued that demonstrations of unconscious learning should satisfy two criteria: the information criterion and the sensitivity criterion. The information criterion states that both explicit and implicit tests should target the same pool of knowledge, and that this can be accomplished by keeping the testing conditions as equivalent as possible. The sensitivity criterion is more difficult to fulfill and assumes that the direct test should be sensitive to all the conscious knowledge developed by the participants.

Reingold and Merikle (1988) proposed a comparison of direct and indirect measures of the same knowledge as a way to fulfill the information criterion. An illustration of how this can be implemented is provided by Destrebecqz and Cleeremans (2001, 2003), who applied the process dissociation procedure (PDP, Jacoby, 1991, 1998) to the SRT task. In the studies of Destrebecqz and Cleeremans, participants first performed a typical SRT task. Afterwards, their knowledge of the sequence was measured through a generation task. During the generation task, the screen showed the same locations as in the SRT task (four reference rectangles horizontally aligned on the screen) and, when any given key was pressed, a dot appeared inside the corresponding location and stayed there until another key was pressed. Destrebecqz and Cleeremans asked their participants to generate a series of keypresses that resembled the trained sequence (“inclusion condition”) and, in a separate block, to make responses such that the trained sequence was avoided (“exclusion condition”). Because only the instructions were different, the two tests are identical in terms of task requirements and stimuli, thus satisfying the information criterion. Destrebecqz and Cleeremans used as a measure of learning the proportion of triplets from the trained sequence that participants produced during the generation task. Following the logic of the PDP, performance in the inclusion condition is supported by conscious recollection (declarative) and unconscious recognition.
Contributions to Learning

(procedural) processes. If participants produce triplets from the training sequence when they are explicitly instructed to avoid doing so (as in the exclusion condition), then those triplets can be regarded as the consequence of knowledge that cannot be controlled and is, therefore, unconscious.\(^1\) Applying the described adaptation of the PDP, Destrebecqz and Cleeremans found that the manipulation of the response to stimulus interval (RSI) could selectively decrease the explicit knowledge gained in the SRT task, but leave overall learning unaffected.

As explained previously, implicit learning is commonly assumed to rely on procedural memory and to occur independently of attentional availability. The role of attention in implicit learning has been studied through dual-task procedures. For example, in sequence-learning experiments, participants performing the SRT task only (single-task condition) have been compared with participants performing the SRT task and a concurrent secondary task (dual-task condition). After a series of blocks of training, participants are presented with testing blocks in which the sequence changes, with the difference in RT between training and testing blocks typically considered the measure of learning. Introducing a secondary-task has been shown to disrupt learning in sequence learning (Nissen & Bullemer, 1987), system-control (Hayes & Broadbent, 1988) and artificial grammar tasks (Dienes, Broadbent & Berry, 1991). These results suggest that attention is necessary for implicit learning and that the underlying process cannot be considered automatic. As will be discussed in detail in Chapter 5, learning under dual-task conditions is possible with some sequences (Cohen, Ivry & Keele, 1990; Curran & Keele, 1993), leading different authors to propose a variety of hypotheses to explain the effect of a dual-task load in implicit learning in the SRT task.

Stadler (1995; see also Frensch & Miner, 1994) tried to explain the interference of a secondary task in implicit learning by arguing that the stimulus for the dual-task disrupts the temporal organization of the sequence used in the SRT. However, Jimenez and Mendez (1999) introduced a secondary task that avoided this problem (by using the stimulus of the SRT task as target for the dual-task), and found no disruptive effect of the secondary task in learning. Frensch, Lin and Buchner (1998) proposed that a secondary task interferes with the expression rather than with the acquisition of knowledge. That is, participants in the dual-task condition may show worse performance because they are both trained and tested under an attentional load.

The PDP allows the calculation of the quantitative estimates of conscious and unconscious knowledge based on a series of assumptions (Jacoby, 1991, 1998). However, these assumptions have been disputed (Ratcliff, McKoon & Van Zandt, 1995; Tunney & Shanks, 2003). For this reason, Destrebecqz and Cleeremans (2001, 2003) based their adaptation of the PDP in the comparison between inclusion and exclusion scores only (see also Neal & Hesketh, 1997; Shanks, Rowland & Ranger, 2005).
load. Frensch and colleagues had their participants follow either most of their training (86% of trials) or only a small part (28% of trials) under dual-task conditions, and then tested all participants under a secondary-task load. The results supported their hypothesis showing no differences in learning between single- and dual-task training conditions. However, a recent study by Shanks and Channon (2002) has challenged their results by showing differences between participants trained under single-task conditions and participants trained under dual-task conditions, regardless of whether they were tested with or without a secondary task.

An alternative explanation of the disruptive effect of a secondary task in implicit learning has been proposed by Cleeremans and Jimenez (1998). Cleeremans and Jimenez argued that a secondary task might interfere with the development of explicit knowledge of the sequence during the SRT task. This hypothesis is consistent with the view of two general memory systems that work in parallel and with the assumption that no task can invoke exclusively one type of processing (Squire, 2004). Furthermore, application of the PDP has consistently shown that participants acquire some degree of explicit knowledge during the SRT task (Destrebecqz & Cleeremans, 2001, 2003; Shanks et al., 2005), so it is reasonable to assume that the acquisition of this knowledge might be disrupted by an attentional load. The application of the PDP described earlier seems like an appropriate way to answer the question of what type of knowledge is more disrupted by the introduction of a secondary task.

The SRT task seems an appropriate paradigm for studying learning when the use of declarative memory is interfered with. The incidental nature of the learning situation makes unlikely that the participant will engage in explicit strategies to test and develop hypothesis. Dual-task paradigms can be used to load attention and interfere with the working of declarative memory. Furthermore, the simple stimuli used can be easily manipulated to produce small increases, for example, in complexity, allowing thus extra experimental control. The interest of the paradigm for the purposes of the present thesis is to test the hypothesis that interfering with declarative memory especially disrupts the acquisition of complex, relational memory representations.

Variables that Promote the Use of Declarative Memory

A number of variables such as type of feedback, schedule of practice, making errors and testing knowledge have been shown to have a detrimental effect on performance during training but to, nevertheless, serve the long-term goal of promoting retention and transfer to different contexts (Goodman, 1998; Hesketh, 1997; Schmidt & Bjork, 1992). Bjork (1994) has argued that a common effect of these manipulations is the
introduction of difficulties in learning, which may lead to extra information processing. The consequence is the development of more elaborated representations encoding relationships between elements of the task and the context (Bjork; Schmidt & Bjork) and improved metacognitive skills (Hesketh). As relational representations are characteristic of declarative memory (Eichenbaum & Cohen, 2001) it can be argued that variables such as the provision of feedback, making errors, introducing testing events and scheduling of practice affect explicit processing and development of declarative memory representations.

**Effects of Feedback about Errors on Learning**

The two main manipulations of feedback involve the temporal contingency and specificity of the feedback with regard to the response (Schmidt & Bjork, 1992). Anderson and the ACT-R group (Anderson, 1993; Schooler & Anderson, 1990) have carried out numerous studies using a variety of tasks designed to determine when feedback should be given. For example, Lewis and Anderson (1985) manipulated the delay of feedback in a task of navigation through a maze and found that, although participants in an immediate-feedback condition performed more accurately and learned the correct moves faster, participants in a delayed-feedback condition learned to recognize “dead ends” better. Schooler and Anderson compared the performance of participants learning the LISP programming language and found that participants in an immediate-feedback group worked through the training material in less time than those in a delayed-feedback group. However, participants in the immediate-feedback group made more errors and were slower when they had to solve test problems, whereas participants in the delayed-feedback group self-corrected more errors. Schooler and Anderson hypothesized that immediate feedback competes for resources in working memory and decreases the probability that task-relevant information, necessary for the development of secondary skills such as error detection and self-correction, will be present and active during the time spent on task. Anderson, Corbett, Koedinger and Pelletier (1995) argued that immediate feedback might be beneficial in cutting down the time spent in error states, whereas delayed feedback might be more effective in promoting other abilities that can facilitate transfer.

The idea that immediate feedback after a response might interfere with the development of secondary skills such as error detection and correction is expressed in the guidance hypothesis of Schmidt, Young, Swinnen and Shapiro (1989). Schmidt et al. compared the performance of people learning a task involving a complex arm movement and receiving feedback about results after every trial or as a summary after a number of trials. The results showed that giving feedback as a summary had a
negative effect on performance in training but improved performance in a delayed test. They interpreted this as evidence that feedback after every trial provides guidance that benefits performance in early stages of acquisition. However, the learner might become too dependent on the guidance and this dependence might prevent him from processing other aspects of information related with the task (environmental cues, proprioceptive feedback), that might prove fundamental for performance in transfer.

Specificity is another characteristic of feedback that can be manipulated and that may result in improved retention performance. Specificity of feedback refers to the level (detail) of the information presented in feedback messages. Kluger and DeNisi (1996) point out that an increase in feedback specificity has traditionally been considered beneficial for performance, but this is not always the case. For example, Anderson, Conrad and Corbett (1989) manipulated the specificity of the feedback given by a LISP tutor. Students were assigned to a condition in which errors were signaled immediately after their commission, or to a condition in which errors were signaled and an explanation was provided. Students who received explanations made fewer errors and progressed through the training with more ease, however, no differences were found in final tests of performance.

Goodman and Wood (2004) pointed out that highly specific feedback during training decreases the chances that learners will experience non-optimal scenarios and thus provides fewer opportunities to learn from errors. Consequently, learners who receive specific feedback might acquire the knowledge needed to behave under optimal conditions but will not know how to behave when the conditions are less than optimal. In order to test this hypothesis, Goodman and Wood had participants perform a complex, personnel management task under conditions that differed in the specificity of the feedback (low, medium or high specificity). In the task, participants had to make decisions (e.g., to assign goals, rewards, etc) to manage a group of workers. During the practice period, starting conditions were set to optimal and participants in the high-feedback condition used the information from the feedback to keep their performance high, whereas the performance of participants in the low-feedback condition quickly deteriorated. In the testing phase, starting conditions were randomly assigned and, as predicted, participants in the high-specificity feedback condition showed less capacity to make correct decisions under non-optimal circumstances than participants in the low-feedback condition.

Goodman and Wood (2004) argued that increased specificity of feedback leads to better performance in general and, therefore, fewer opportunities to learn the set of responses needed to confront more difficult situations. When the specificity of
feedback increases, the feedback focuses more on particular behaviors and gives more information about the locus of errors, thus guiding learners in identifying correct and incorrect actions. This has a beneficial effect in short-term performance. However, the guidance leads learners to work less actively to find the link between their actions and the outcomes. This seems to impair information-processing related activities such as error diagnosis, encoding and retrieval. On the other hand, low levels of specificity of feedback force learners to experiment with different actions to discover which response is correct (Goodman & Wood; Kluger & DeNisi, 1996).

Overall, feedback contingency and specificity seem to have similar effects. Feedback given immediately or with a great degree of specificity seems to promote performance of the learners during training by way of guiding behavior and reducing the difficulty of the task. This has, however, the effect that the training situation loses some of its resemblance to the real-world situation, where often the feedback is vague or non-existent. A consequence is that the learner might not develop certain skills (e.g., error detection) that will be very important when the guidance is not present.

**Effects of Making Errors in Learning**

Although the opportunity to make errors may be associated with improved performance in the long run, in some type of situations it might be advantageous to structure learning in order to avoid errors entirely. Training in which errors are prevented has been shown to result in improved performance in clinical settings. In such errorless training (Terrace, 1963), participants are prevented from making errors by giving them the response in advance (Baddeley & Wilson, 1994; Kessels & De Haan, 2003). As an example, Baddeley and Wilson compared the performance of amnesic and control participants learning lists of words under errorful and errorless conditions. The task was a stem completion (for example, the experimenter would say; “I am thinking of a five-letter word beginning with QU”) and participants were either given the correct response in advance (errorless condition: “And the word is QUOTE. Please write it down”) or encouraged to try until they succeeded (errorful condition: “Can you guess what the word might be?”). The results showed that all participants performed better under errorless conditions and that this advantage was especially significant in the case of the amnesic group. Baddeley and Wilson hypothesized that declarative memory serves to filter out incorrect responses and because amnesics lack this type of memory, errorful training is especially harmful in their case. That is, contextual information provides the necessary information to filter wrong from correct associations but as amnesics cannot code this extra information, it is especially beneficial in their case to avoid errors and learn only correct associations.
The errorless approach has been successfully applied to participants with memory decrements due to dementia of the Alzheimer type (Clare et al., 2000) and advanced age (Kessels & De Haan, 2003). Tulving, Hayman and MacDonald (1991) used a study-only procedure to introduce a considerable amount of factual knowledge to K.M., an amnesic patient. The study-only procedure consists of numerous presentations of the items for study without testing, avoiding thus the possibility of committing errors. In that sense, it is a methodology in which errors are avoided and constitutes an example of successful application of error-free training to a profoundly amnesic patient. Finally, in a series of experiments comparing errorless and trial-and-error learning techniques, Evans et al. (2000) found that errorless learning methods were more beneficial in tasks and situations that could be completed using implicit memory. Tasks that required the explicit recall of novel associations did not benefit from errorless learning.

The errorless training approach has also been applied in healthy populations. For example, Maxwell, Masters, Kerr and Weedon (2001) trained participants in a task of golf putting under errorless and errorful conditions. Errorful learners showed more deterioration of performance under dual-task conditions than errorless learners did and seemed to engage more often in hypothesis-testing behavior. Maxwell et al. hypothesized that committing errors led to the development of an explicit strategy in learning motor skills, which in turn motivated the development of explicit hypotheses and rules. Explicit revision of those rules was disrupted by the secondary task and this resulted in a reduction in the level of performance under dual-task conditions of participants trained under errorful conditions with respect to those trained under errorless conditions.

As mentioned earlier, in fields such as problem solving, errors have traditionally been considered to have some positive effects. The reason may be that failing to solve a problem (impasse) can be an event that forces the learner to stop, step back and try to reconsider and re-analyze all the available information. This re-conceptualization is what leads to the discovery of not-yet-perceived relations between the stimuli, to a change of mind-set and to a new approach to solving the problem (insight) that, in the last instance, might determine reaching the solution. According to the classical theory of problem solving (Newell & Simon, 1972; VanLehn, 1989), an impasse occurs because the procedure mandates an impossible action or makes a false claim about the current state. The behavioral correlates are periods where no problem-solving activity takes place or where the same problem-solving activity is repeated. The solution to an impasse is called repair, and it might consist of the rejection of the current schema and the selection of another (VanLehn). Gick and McGarry (1992) gave an example of positive relation between failures in
attempts to solve a problem and spontaneous transfer when solving an analogous problem. In their experiments, participants trained with a source problem and tried afterwards to solve a target problem. Gick and McGarry found that transfer was promoted through a manipulation of the surface similarity of the problems that affected the number of erroneous solutions during training. When the training problem produced solution failures similar to those characteristic of the target problem, a successful solution of the target problem was more likely to occur. When solution failures of the training problem were prevented, transfer did not take place.

Another line of research deals with the emotional and meta-cognitive effects of errors during learning. In the so-called error-management technique developed by Frese and collaborators (Dorman & Frese, 1994; Frese et al., 1991; Heimbeck, Frese, Sonnentag & Keith, 2003) learners are allowed to make errors during training and are encouraged to learn from them. Errors might be provoked by, for example, giving the learners problems that exceed their level of expertise (Frese et al., 1991) or providing minimal instructions and encouraging the learners to learn through exploration of the system (Heimbeck et al., 2003).

The main objective of the error-management technique is to minimize the negative effects of errors while maximizing their positive effects. The emotional negative effects are minimized by way of instructions; the strategic negative effects (e.g., damage to the system, incompletion of the task) are prevented by providing a safe environment for the errors to appear (Heimbeck, et al., 2003). Among the positive consequences of making errors, Dormann and Frese (1994) proposed that errors instigate and require exploration, and by way of exploring, the learner gets to know more of the system and a richer mental model is developed. Frese and Sabini (1985) also argued that errors give a chance to reintellectualize strategies that have been automatized. Similarly, Ohlsson (1996) proposed that knowledge of skills is encoded in the form of production rules and errors push the learner to modify old, faulty rules and derive from them new, more specific rules, which can be indeed seen as an enrichment of the mental model. A second positive effect of making errors is that the person learns to deal with the negative emotions associated, as well as with the whole error situation per se (Dorman & Frese). Finally, allowing the trainee to make errors makes the training situation more similar with a real-life scenario, a

\[ Ohlsson (1996) \] proposed a model in which errors are caused by overly general knowledge structures. These structures have the form of production rules \((G, S \rightarrow A)\), meaning “IF the goal \(G\) is pursued and the situation \(S\) is encountered, THEN carry out the action \(A\)”. Error correction is made through refinement of the declarative, conditional part of the rule (adding to \(S\) the details of the situation in which the error was produced, thus creating a new rule more situation-specific).
characteristic that is a well-known facilitator of transfer (Thorndike & Woodworth, 1901).

A study by Frese et al. (1991) can serve as illustration of the error-management approach. Frese et al. compared the performance of participants in error-avoidant and error-training groups in a task that involved learning how to use a word processor. Errors were avoided in the error-avoidant group by giving detailed and precise instructions of how to complete each of the problems. If the subject misspelled a word, the experimenter quickly intervened to avoid the mistake. Exploration was discouraged and, when the task had been finished, the subject was asked to repeat it to fill the training time. In the error-training group, the participants only received the commands necessary to finish the tasks but no instructions on how to do it. They were encouraged to explore the system and to solve the problems by themselves. The results showed superiority for the error-training group in solving difficult tasks under non-timed conditions but no significant differences under speeded conditions.³

Dormann and Frese (1994) also found that errors promoted exploration in a task of learning to use the SPSS data analysis environment and exploration activity correlated positively with performance in a posttest. Finally, some relation has been found between error-management training and individual differences. Gully, Payne, Koles and Whiteman (2002) measured cognitive ability and openness to experience in their participants and trained them in a decision-making simulation task under conditions that avoided or encouraged making errors. Their results showed that individuals with higher cognitive ability or higher scores in openness to experience were more likely to benefit from conditions that encourage making errors than individuals with lower cognitive ability or lower openness to experience scores.

Effects of Testing in Learning

Testing, as long as it involves the retrieval of information, constitutes in itself a powerful learning event. The enhanced retention due to the introduction of testing blocks (or trials) during a study phase has been called the testing effect (Bjork, 1994; ³ This is a very interesting result that points to a superior development of explicit knowledge in the error-training group. If the error-training group had an advantage in the development of explicit knowledge, it would also need time to apply this type of knowledge. Without time, the advantage over the error-avoidant group might just vanish. This can be expected from the very nature of explicit memory and explicit processing which require the conscious, time-consuming application of rules as opposed to implicit processes, which work in an automatic fashion. Studies in implicit learning also show that manipulations of attention affect not only the acquisition of knowledge but also its expression (Curran & Keele, 1993; Frensch et al., 1998). Finally, R. Ellis (2005) proposes that imposing a time limit (speeded test) is the difference that changes the nature of a recognition test from explicit into implicit.
Roediger & Karpicke, 2006a). In a typical study of the testing effect, participants are initially given a list of words to memorize. After this study phase in which the words are simply presented, one group continues with a number of study blocks (study condition) and another group follows an equivalent number of blocks in which study is combined with free recall (test condition). The typical finding is that, if retrieval is successful, testing can promote long-term recall more efficiently than an equivalent number of study opportunities, even when the tests are given without feedback (Carrier & Pashler, 1992; Hogan & Kintsch, 1971; Kuo & Hirshman, 1996).

An important feature of the testing effect is that the advantage in performance for the testing groups is especially visible in delayed tests. For example, Wheeler, Ewers and Buonanno (2003) had subjects learn a list of 40 words either by studying the list five consecutive times (study condition) or by studying the list once and recalling it, without feedback, four times (test condition). When the participants were tested, participants in the study group outperformed those in the test group in an immediate free recall test. However, a week later this effect reversed and participants in the test group recalled twice as many words as those in the study group did. Roediger and Karpicke (2006b) found a similar effect using passages of scientific texts as study material. Their participants studied the texts four times (study-only condition), studied three times and took a recall test (single-test condition), or studied once and took three recall tests (multiple-tests condition). They then performed an immediate-recall test after five minutes or a delayed-recall test after a week. Participants in the study-only condition showed better recall than those in the other two conditions in the immediate test but were outperformed in the delayed test, with participants in the multiple-tests condition showing the best performance. Thus, introducing tests during the learning stage reveals itself as a manipulation with a similar effect to that of introducing errors or delaying the feedback during training: deterioration of immediate performance but promotion of long-term retention (Bjork, 1994).

The beneficial effect of testing on long-term recall when the testing is delayed with respect to the initial study opportunity has also been called the spacing effect (Dempster, 1988). In one of the earliest examples of the phenomenon, Jacoby (1978) had people study word pairs and then re-study the pair (study condition) or take a test in which they had to generate the right-member of the pair after being prompted with the left-member. This test was given, either immediately after studying the pair

---

4 The study of Jacoby (1978) has been cited as example of the testing effect (Roediger & Karpicke, 2006a), the spacing effect (Dempster, 1988) and the generation effect (Burns, 1990), which gives an index of the relevance of the study and the close relation between the three effects.
(immediate-test condition) or after an interval filled with studying other pairs (delayed-test condition). All participants received at the end of the experiment a cued-recall test in which they were presented with the left member of the pair and they had to produce the right member. The results of this final test showed an advantage for testing versus studying (a testing effect) and an advantage within the testing conditions for the delayed-test group (a spacing effect). In a recent study, Pashler, Zarow & Triplett (2003) had their participants learn pairs of Eskimo-English words and tested them twice with different intervals between the two tests. Long intervals produced a greater deterioration of performance in the second test than short intervals but improved performance on a retention test the following day. However, the participants in the study of Pashler et al. received feedback after every test on the first session, thus the effect of delayed re-testing might be confounded with the effects of delayed re-study. Nevertheless, Pashler et al.’s study represents another good example of a manipulation that caused many errors in an immediate test but proved to be positive in terms of long-term retention.

The positive effects in long-term retention derived from introducing tests during learning have been shown in studies with free-recall tests of word lists (Hogan & Kintsch, 1971; Wheeler, Ewers & Buonanno, 2003), cued recall and paired-associate learning (Carrier & Pashler, 1992; Kuo & Hirshman, 1996), and free-recall and multiple-choice tests of educational texts (Agarwal, Karpicke, Kang, Roediger & Mcdermott, 2007; Nungester & Duchastel, 1982; Roediger & Karpicke, 2006b). Frequent testing has also been shown to promote recall and retention of the learned material in studies outside the laboratory using students attending college courses (Bangert-Drowns, Kulik & Kulik, 1991; Leeming, 2002). The testing effect has been explained as the result of an increased elaboration of the memory trace, which facilitates multiple retrieval routes (McDaniel & Mason, 1985). Roediger and Karpicke (2006a) interpret the testing effect as yet another example of manipulations that slow initial performance but promote long-term retention by way of introducing desirable difficulties (Bjork, 1994).

Other Variables that Influence Learning

Bjork (1994) describes a number of variables that can be manipulated to introduce difficulties during training and promote long-term retention and transfer, thus improving learning. A first symptom of an increase in the difficulty of the learning situation would of course be an increase in the number of errors (and/or an increase in the average RT). One manipulation that can influence learning, in addition to the already described manipulations of the feedback, making errors and testing effects, is
the manipulation of practice schedules (Bjork; Johnson, 2003; Schmidt & Bjork, 1992).

One possible manipulation of practice schedules involves giving *blocked* versus *random* practice. In blocked practice, one variation of the task is practiced in one block or session and another variation in a different one, while in random practice all variations are possible within the same block. Blocked practice has been shown to produce better performance during training but random practice produces better results at transfer and retention tests. These results have been replicated in a variety of fields, including motor tasks (Shea & Morgan, 1979), acquisition of typing skills (Baddeley & Longman, 1978), problem solving (Carlson & Yaure, 1990) and acquisition of logical rules (Schneider, Healy, Ericsson & Bourne, 1995). A proposed explanation for the advantage at transfer of random scheduling is the *contextual interference hypothesis* (Battig, 1979), which suggests that changes in the experimental context inter-trials result in a more elaborate processing independent from the context, which promotes transfer.

Another manipulation of the practice schedule is that of *massing practice*, such that a few sessions consisting of many trials each are followed, versus distributed practice, in which the training is distributed along more sessions with fewer trials each. Massed practice produces better short-term performance but distributed practice seems to show a long-term advantage (Bjork, 1994). The beneficial effect of distributed practice on transfer performance has been shown in such different fields as motor skill acquisition (Lee & Genovese, 1989), retention of mathematical knowledge (Rohrer & Taylor, 2006) and verbal recall tasks (Cepeda, Pashler, Vul, Wixted & Rohrer, 2006). There is some evidence, however, showing that the effect might interact with the type of task: In continuous tasks, massed practice has a negative effect on retention but with discrete tasks, this effect might be positive (Lee & Genovese, 1989; see also Donovan & Radosevich, 1999). Bjork proposes that the spacing effect, described in experiments with verbal tasks (Melton, 1967; Pashler et al., 2003) is another example of advantageous transfer due to distributed practice: The massing of repetitions produces better performance during the training session but poorer transfer than when the repetitions are more spaced. In the case of the spacing effect, a commonly assumed explanation is that spaced repetitions increase the number of retrieval cues. That is, the retrieval cues associated with each repetition are more likely to be different if they are separated from each other in time, and thus context (Glenberg, 1979). Having more cues to retrieve a given association means greater access to it and easier generalization to different contexts.
Schmidt and Bjork (1992) adopted a transfer-appropriate processing approach to explain the effects of manipulating practice schedules. Manipulations of the scheduling, feedback or variability of the task increase the performance at transfer because they increase the overlap between the type of information processing exercised at training and that necessary for performance at testing situations. Although the short-term effect of the manipulations is an increase in the difficulty of the training situation, nevertheless they prepare the learner better for the type of processing that will be required at transfer. Based on the work of Bjork and colleagues, McDaniel and Einstein (2005) have recently proposed a material appropriate difficulty (MAD) framework to determine when difficulties should be introduced in the learning situation. The model tries to relate learning materials, learner characteristics and desirability of difficulty for learning enhancement. A basic assumption is that difficulty is desirable when it stimulates processing that complements that required by the training task.

Individual Differences in Learning

A last group of variables that might affect learning is not composed of variables that are manipulated by the experimenter but variables in which the participants already differ, that is, individual differences. Ackerman (1988) proposed a model in which the role different cognitive abilities play in performance depends on the stage of skill acquisition. The stages of skill acquisition are respectively, cognitive mediation, an associative phase, and automatic performance (Fitts, 1964). The initial stage of cognitive mediation involves the acquisition of declarative knowledge, and it therefore places demands on memory, reasoning and knowledge retrieval. During the associative phase, learners develop production rules and, as speed develops, performance becomes more dependent on task-specific associations and less on declarative knowledge. During the last stage, performance has become automatized and relatively independent of attention, declarative knowledge is not needed and performance is related with psychomotor ability.

The model of Ackerman (1988) allows clear predictions that can be tested through the comparison of groups with different levels of ability. Matthews, Jones and Chamberlain (1992) compared the performance of high- and low-ability individuals in a mail-coding task and found that, in contrast to predictions based on Ackerman’s model, differences in ability did not involve changes in the correlation with cognitive and perceptual-speed tasks. A measure related with fluid intelligence (digit span) was predictive of performance at all ability levels, and therefore, at different stages of skill acquisition. Ackerman (1992) also found that general, and not just perceptual, abilities
Contributions to Learning

correlated with performance at high levels of skill acquisition in a simulation of air-
traffic control. These findings suggest that when complex tasks are involved, a
complete independence of general ability might never be reached due to the
involvement of working memory and executive control in performance (Johnson,
2003). On the other hand, if a simple task can be easily automatized, a complete
independence of general intelligence may be achieved. Reber (1993) has argued that
implicit learning, being based in an evolutionary older system than explicit learning,
should be more robust, relatively independent of age and development, and display
smaller population variances and tighter distributions. Moreover, the independence of
implicit learning from control and attentional resources should be reflected in a lack
of correlation between performance in implicit learning tasks and measures of
cognitive ability such as intelligence. This idea is supported by studies in automaticity
(Hasher & Zacks, 1979, 1984) showing that coding of basic information such as
frequency is independent of developmental differences. Aslin, Saffran and Newport
(1996) have also demonstrated that 8-month-old infants can learn conditional
probability statistics at a pace equal to that of adults. A number of studies have also
shown independence between performance in implicit learning tasks and cognitive
measures such as IQ (Gebauer & Mackintosh, 2007; Maybery, Taylor & O’Brien-

Individual Differences in Working Memory Capacity (WMC)

Working memory can be defined as a limited capacity system that temporally
maintains information and provides an interface for perception, long-term memory
and action (Baddeley, 2003; Miyake & Shah, 1999). Working memory is intimately
linked with long-term memory, and more specifically, with declarative memory. For
example, Eichenbaum and Cohen (2001) describe working memory as a form of
declarative memory because it involves relational judgments, carries out conscious
processing of information and its contents can be explicitly expressed. Budson &
Price (2005) argue that working memory is a declarative and explicit system of
memory because it requires active and conscious involvement of the individual.
Furthermore, although different working memory tasks can invoke different parts of
the brain, the main areas consistently involved in working memory functioning are the
frontal lobes and especially the prefrontal cortex, which also have an important role in
declarative memory (Fletcher & Henson, 2001). The link between working memory
and declarative memory is so strong that some models consider working memory as
that part of long-term (declarative) memory under the focus of attention (Cowan,
1995).
In addition to the link between working memory and declarative memory, working memory capacity has been related with attention. Working memory capacity (WMC) is assumed to reflect domain-general executive attention, as opposed to short-term memory capacity (STMC) which reflects domain-specific storage (Engle, Tuholski, Laughlin & Conway, 1999). Working memory capacity is commonly measured with WM span tasks, which combine the presentation of target stimuli to be remembered with the presentation of a secondary task that demands concurrent processing. Engle and colleagues argue that WMC differences should influence performance when the task at hand demands attention, but not when it can be carried out with automatic processing (Tuholski, Engle & Baylis, 2001). It can then be expected that WMC will not influence performance when carried out under implicit learning conditions. As WMC is a construct that correlates highly with measures of intelligence (Conway, et al., 2005), this prediction is consistent with the described hypothesis of Reber regarding the independence of implicit processing from measures of cognitive ability (Reber, 1993).

Not all the evidence supports the hypothesis that performance in implicit learning tasks is independent of WMC. For example, Karpicke and Pisoni (2004) found that subjects with a greater auditory digit span were better able to learn an artificial grammar. These results can be explained assuming that, as Reber notes (Reber & Allen, 2000), participants in the experiment might engage in some explicit learning during the artificial grammar task, and this explicit learning might be related with cognitive measures. Frensch and Miner (1994) found a correlation between short-term memory span and sequence learning in a SRT task under intentional instructions with a short RSI (Experiment 1) and dual-task conditions (Experiment 2). The results of Frensch and Miner are more difficult to explain because, assuming that: 1) implicit and explicit learning coexist in the task, 2) an attentional load would interfere especially with explicit learning and, 3) WMC is related with explicit learning and attention, then the secondary task should load attention and interfere with working memory, cancelling any possible differences in performance due to differences in WMC. However, Frensch and Miner measured STMC and, as explained, this construct fails to capture the aspect of maintenance of active information while performing concurrent processing reflected by WMC and measured with WM span tasks. The hypothesis that WMC should correlate with learning when all attentional resources are available was tested in the experiments described in Chapter 3 of this thesis.

Conclusion
The manipulation of variables that make training more difficult, such as delaying feedback, introducing tests, introducing errors or increasing the variability of the task, seems to promote performance at transfer and delayed tests (Schmidt & Bjork, 1992; see also Bjork, 1994; Goodman, 1998; Hesketh, 1997). Bjork argued that the extra difficulty of training pushes the learner into processing additional context-related information and facilitates the development of more flexible representations that are responsible for long-term retention and improved transfer performance. Transferable, relational representations are characteristic of declarative memory (Eichenbaum & Cohen, 2001; Squire, 2004; Schacter & Tulving, 1994). It can thus be hypothesized that an increase in the difficulties of training promotes the use of declarative memory. The next section describes the predictions derived from this hypothesis and the studies designed to test them.

Outline of the thesis

The main goal of this thesis is to test the hypothesis that the manipulation of variables that increase difficulty during training promotes the use of declarative memory. Based on this hypothesis, it can be predicted that encountering difficulties during learning will promote memory for relations and transfer performance. Additionally, it is expected that manipulations that interfere with declarative memory will be especially detrimental for the learning of relational representations.

The fields chosen to test these predictions ranged from a simple sequence-learning task to complex training tasks in which foreign languages and maps were used as stimuli. The logic behind the selection of tasks was twofold. First, a general theory of memory should apply to fields that differ in their degree of complexity. Second, it is expected that the material will influence the degree to which the different memory systems influence learning within the strict context of an experimental situation. Complex stimuli (i.e., a foreign language or a map) might be more prone to be processed explicitly than simple stimuli whereas simple stimuli as those used in the sequence learning studies might be more readily processed by the procedural system than complex stimuli.

The studies in Chapter 3 describe two experiments in implicit learning with the SRT task. These experiments investigated the effect of an attentional load in declarative and procedural memory. In addition to the availability of attention, the stimuli used in the SRT task were also manipulated in order to present different levels of complexity. It was expected that the attentional load would interfere with declarative memory and deteriorate the learning of complex sequences of stimuli. Furthermore, a measure of WMC was taken with the expectation that it would
correlate with performance when attention was available but not when attention was loaded through a secondary task.

The first study of Chapter 4 describes an experiment in second language acquisition (SLA) in which the probability of making errors was manipulated. The expectation was that making errors would facilitate transfer to novel stimuli whereas avoiding errors would promote speed of processing and performance during a timed test. In the second experiment of Chapter 4, the probability of making errors was combined with a manipulation of the focus of training. Participants training with a focus on form learned under conditions that promoted memory for rules whereas participants training with a focus on meaning learned under conditions that promoted memory for examples. The hypotheses in this study were that making errors and learning with a focus on form would promote active hypothesis-testing behavior and the use of explicit memory, whereas learning that avoided errors and focused on meaning would promote memory for examples, passive learning of stimulus-response associations and implicit memory use. The results are interpreted considering theories of procedural and declarative memory within the field of SLA (Pinker & Ullman, 2002).

The last empirical chapter describes one experiment in map learning in which the role of making errors and receiving feedback from those errors was investigated. In line with the hypothesis of Schmidt and Bjork (1992) that introducing difficulties during learning promotes transfer performance, it was predicted that making errors and delaying feedback would result in declarative memory use, formation of relational representations and improved performance at later tests. Theories of implicit and explicit knowledge from the field of geography learning (Friedman & Brown, 2000; Postigo & Pozo, 2004; Tversky, 1993) are considered in the interpretation of the results.