Transfer and Interference in skill acquisition

Dekker, B.

Citation for published version (APA):
Interference in Letter-Digit Substitution

A series of four experiments is reported, addressing questions about the occurrence of negative transfer and retroactive interference in skill acquisition. In Experiment 2.1 negative transfer was studied in the learning of different variants of a letter-digit substitution task (coding skill), implemented as a choice-reaction time task. Subjects learned one variant of the coding skill and transferred to a coding skill in which stimuli and responses were re-paired (A-B, A-Br design). Negative transfer was found. In Experiment 2.2, the coding task included a retroactive test phase in which the first learned skill was retested after interpolated coding. Negative transfer and retroactive interference were found. The negative transfer effect was, however, smaller and more transient than the effect Parreren (1951) found. Therefore, a closer replication of the Parreren study with a paper and pencil coding instead of a computer version was conducted in Experiment 2.3. Experiment 2.4 investigated also the occurrence of retroactive interference. Strong negative effects were found. The effects were found in response times, as well as in error rates (less consistent). It is concluded that, in contrast with some of the current ideas on this topic (Anderson, 1995), interference does occur in learning of skills.

Positive transfer of training, i.e. the positive effect of a preceding activity (training task) upon the learning of a new task (transfer task), is thought to be possible when tasks have elements in common (Thorndike, 1906). According to this identical elements theory, the amount of positive transfer increases when similarity between tasks increases. Negative transfer is not possible according to the identical elements theory because the least possible transfer is when tasks have no elements in common. Studies on verbal learning (e.g., learning paired associates) have shown that an increase in similarity can also lead to negative transfer (e.g., Bilodeau & Schlosberg, 1951). Thus, increases in similarity can both lead to positive and negative transfer.

In contrast with results found in verbal learning, negative transfer and interference have seldom been found in skill learning. Some studies emphasise the possibility of positive transfer by showing that in research in which interference is expected, i.e. situations where stimulus conditions are the same but the responses required are re-paired, A-B, A-Br design, no negative transfer was found (Singley & Anderson, 1989). This led Anderson (1995) to conclude that interference such as found in memory for facts, is nonexistent in skill learning. The only
exception, according to Anderson, is negative transfer that is due to transfer of nonoptimal methods, also referred to as set or Einstellung. Transfer is perfect in this situation, but the subjects do not consider other and better solutions to a problem than the ones already learned. The nonexistence hypothesis as proposed by Anderson can be questioned for several reasons.

1. A possible reason for the absence of negative transfer in many experiments is that transfer effects are often measured as an aggregate of several effects resulting from different sources. Even when an A-B, A-Br design does not lead to negative transfer on an aggregate level, it is still possible that the positive effects (e.g., training effect) mask the negative effects. Siipola (1941) who did not find negative transfer in a transfer task with reversed responses argues that positive influences, such as familiarity with the broad nature of a task, overshadowed the negative influence of the reversions. To gain insight into the occurrence of interference, it is important to look at the component processes and to isolate the interference effects. To circumvent the problem of aggregate measures, it is necessary to differentiate within the transfer task between learning of the new procedures and negative transfer. This can be accomplished by comparing an experimental group with a control group for whom the transfer task contains new stimuli instead of a reversal of the first learned stimuli. Thus, by comparing these groups the effects of the negative factors alone can be measured.

2. It seems that a reliable assertion about the occurrence of interference in skill learning is only possible when the task complexity and amount of training is somewhat comparable with that in the verbal learning experiments. The nonexistence hypothesis builds on a comparison between the results of learning simple (not always easy however) tasks in verbal learning (e.g., learning of paired associates) and the results of learning complex tasks in skill acquisition (e.g., learning of text-editors). It is concluded on basis of this comparison that findings from the verbal learning literature could not be generalised to skill learning (Anderson, 1995; Olson & Olson, 1990). Thus, although the use of real life skills is a useful way to obtain information on practical issues concerning learning, it is not always suitable for answering questions on a theoretical level.

A simple task is more suitable for answering the question whether negative transfer occurs in skill acquisition for several reasons. First, simple tasks allow more control over the characteristics of the stimulus and response sets, as well as the mappings between them (Proctor, Reeve, Weeks, Dornier, & Van Zandt, 1991). Second, a simple task contains fewer elements than a complex task. The negative transfer effect will be larger in a simple task in comparison with a complex task with the same number of response reversals because in a simple task there are fewer factors working in the positive direction. Therefore, the chances for finding negative transfer at an aggregate level will be higher when using a simple task. Third, a simple task becomes automated faster. To investigate negative transfer in skill acquisition, knowledge of the entry levels of training is needed, as well as behaviour in well-trained situations. It is evident that a simple task makes this easier and less time-consuming.

3. The statement that negative transfer and interference are absent in the case of skill acquisition is not in accordance with other studies in this area. Negative transfer can be found in text editing using the appropriate paradigm (Heydemann, Hoffmann, & Schmidt, 1991; Schmidt, Fischer, Heydemann, & Hoffmann, 1991). In a more simple skill domain, such as reversal learning in animals (e.g., Chittka & Thomson, 1997) or learning SR relations
Interference in Coding Skills

(Fendrich, Healy, & Bourne, 1991; Proctor & Dutta, 1995; Schneider & Shiffrin, 1977), the negative influence of response reversal is well known. There are also earlier examples of negative transfer and retroactive interference in skill acquisition (e.g., Lewis & Miles, 1956; Ryans, 1936; Parreren, 1951). One of the experiments conducted by Parreren (1951) raised our interest in negative transfer in skill learning because its results were so much in contrast with those of Singley and Anderson (1989). Parreren used a variant of the letter-digit substitution test. According to Ryans (1936), this is a very successful paradigm to yield positive as well as negative transfer of training. In the experiment by Parreren, participants were trained to substitute letters with digits according to a key, consisting of a set of letter-digit combinations. The actual substitution consisted of writing digits underneath the presented letters. The training task was followed by the transfer task, which was identical to the training task except for the letter-digit combinations within the key. In the transfer task the letters from the training task were re-paired with digits from the training task (e.g., from N-1, R-2, B-3, D-4 in the training task to R-1, D-2, N-3, B-4 in the transfer task). Robust net negative transfer was found with this task, with participants performing worse on the transfer task than their entry level on the training task.

Unfortunately, Parreren's experimental design has three major flaws. First, no control group was used to measure the time needed to learn the second key. Therefore, interference could only be established in a comparison between the first and the second skill. Unfortunately, this comparison was not informative because of the second flaw: the letter-digit combinations used in the training phase and the transfer phase were not counterbalanced. Therefore, the worse performance in the transfer phase could be the result of a difference in difficulty between the first and the second set of combinations. Third, only data on response times were reported, which made it impossible to establish whether there was evidence for a speed-accuracy trade-off. The shortcomings of this study make it impossible to draw definite conclusions about the occurrence of negative transfer in the coding task.

In response to Singley and Anderson, we push the question whether negative transfer occurs in skill acquisition even further to the limit. In the current experiments, a design producing maximal interference (A-B, A-Br) is used, instead of an approximation to such a design as used by Singley and Anderson. In an A-B, A-Br design, the same stimuli and responses are used in the training phase and the transfer phase, but the combinations in which they have to be learned differ across phases. In a comparison of various designs, the A-B, A-Br design appeared to lead to the greatest amount of negative transfer (Deese & Hulse, 1967). Another difference with the experiment by Singley and Anderson (1989) is that a relatively simple task was used. These experiments were conducted to verify the statement that negative transfer and retroactive interference are nonexistent in skill learning. If interference occurs, the magnitude and duration of the effect will be investigated as well.

The present chapter describes a series of four experiments. The first experiment is a forced choice reaction time version of the Parreren experiment. The main goal of Experiment 2.1 is to investigate the occurrence of interference in the acquisition of a second similar skill. In Experiment 2.2, the first learned task is repeated after the transfer task, to also investigate an eventual interfering effect of the second task on the performance of the first task. Experiment 2.3 is a replication, with necessary modifications, of the original paper and pencil experiment.
by Parreren (1951). Experiment 2.4 is also a paper and pencil experiment, in which the first learned task is repeated after the transfer task to investigate interference between the second task and the first task.

**Experiment 2.1**

The main goal of Experiment 2.1 is to examine whether training on a training task (T1) could impede the learning and performance in a transfer task (T2). Therefore, this is an attempt to replicate the findings of van Parreren (1951), who did find negative transfer in a coding task. In Experiment 2.1, the letter-digit coding was implemented as a computerized choice reaction time task. This experiment was designed to answer several questions. The first question is whether a reversal of responses hinders performance in comparison with a situation in which a new set of responses is learned. The second question is whether response time or accuracy, or both are affected. Finally, the third question is how strong and long lasting the effect is.

In the current experiment, two conditions in letter-digit coding were compared. After training of the first task, the experimental group switched to a key with letter-digit combinations consisting of the same letters and digits as in the training phase, but in other combinations. The control group switched to a key consisting of different letters paired with the same digits. The control group was used to differentiate between time needed for learning of the new key, and time lost due to negative transfer. The expectation is that performance in the transfer phase is worse in the experimental group than in the control group.

**Method**

**Participants**

Fifty-nine students of the University of Amsterdam participated in this experiment in exchange for course credit. All participants were tested individually.

**Design and Materials**

The participant’s task was to press the correct button in response to the presentation of a letter. The letter-digit combinations (letters above the digits) were shown in the upper middle part of the screen and were constantly available during the task, except for during the pauses between blocks. In this experiment the key consisted of four letters linked to the numbers 1 through 4 depicting the four buttons. The keys used are presented in Table 2.1.

In this first experiment we focused on the transfer phase and chose to make transfer keys identical for each subject and use different training keys without counterbalancing across both phases. Because of this choice the comparison between phases cannot be made and it is not possible to verify whether the two groups perform similarly in the training stage. On the basis of the results obtained in the current experiment, the keys were counterbalanced across the phases and the conditions in all further experiments.

Two groups were distinguished, the re-paired group and the new group. The re-paired group worked with letters that were the same across the phases but differed in position (e.g.,
NRDS in T1, DSNR in T2) and thus in the corresponding buttons. The new group worked with different letters in the training phase and the transfer phase (e.g., GKWL in T1, DSNR in T2). Each subject worked with one training key and one transfer key. Stimuli were arranged in 10 blocks, of 60 trials each, with each letter appearing 15 times in a pseudo-random order. After 5 blocks participants switched to the transfer phase.

**Table 2.1.** The keys and their presentation order in the re-paired and the new group in Experiment 2.1.

<table>
<thead>
<tr>
<th>Button</th>
<th>Training</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>Re-paired Group</td>
<td>N R D S</td>
<td>D S N R</td>
</tr>
<tr>
<td></td>
<td>S N R D</td>
<td>R D S N</td>
</tr>
<tr>
<td></td>
<td>D S N R</td>
<td>N R D S</td>
</tr>
<tr>
<td></td>
<td>R D S N</td>
<td>S N R D</td>
</tr>
<tr>
<td>New Group</td>
<td>G K W L</td>
<td>D S N R</td>
</tr>
<tr>
<td></td>
<td>L G K W</td>
<td>R D S N</td>
</tr>
<tr>
<td></td>
<td>W L G K</td>
<td>N R D S</td>
</tr>
<tr>
<td></td>
<td>K W L G</td>
<td>S N R D</td>
</tr>
</tbody>
</table>

**Apparatus**

The experiment was conducted on a PC with a monochrome display and a custom-made button box. Four push-button switches were used. The switch caps were round, approximately 2.2 cm (0.87 inch) in diameter, and labelled with a digit ranging from 1 to 4. The switches were arranged in a half circle around a button indicating the starting position. Hand travel between starting position and the numbered keys was held constant at 5.5 cm (2.17 inch). Response timing was accurate to 1 ms.

**Procedure**

Participants were instructed to make their key-press responses as rapidly and accurately as possible, using the index and middle finger (held together as if it was one finger) of their dominant hand to press the buttons. After each response the fingers had to be brought back to the starting position to prevent decrease of reaction time due to dwelling of the hand before the response. The experimenter checked whether each participant performed the task according to these instructions.

The stimulus display of each trial consisted of a warning stimulus, a plus sign, and a target stimulus, which was one of the four letters from the key. When an error was made a tone sounded as feedback. Between the response and the presentation of the next stimulus was an interval of 1 sec.
Chapter 2

Results

Response Times

Repeated measures ANOVAs were performed on the response time (RT) variable. The factors for the RT ANOVAs were block and group. The analyses were performed on the median RTs per block for every participant. The median RTs were based on the correct responses only.

We first address the data from the training phase (Blocks 1-5). The significant main effect for block, $F(4, 228) = 51.62, p < .0001$, $MSE = 2265$, confirmed the reduction in RTs with practice that is represented in Figure 2.1. No main effect for group, and no significant Group x Block interaction was found, indicating that the two groups were comparable in speed and learning rate in the training phase of the experiment.

![Figure 2.1. Mean correct response time and standard error of the mean as function of block and group in Experiment 2.1. The vertical line indicates a key change.](image)

The RT data for the transfer phase, Blocks 6-10, showed a main effect for group, $F(1, 57) = 4.01, p < .05$, $MSE = 48,285$ and a main effect for block, $F(4, 228) = 33.92, p < .0001$, $MSE = 2261$. Figure 2.1 shows that the re-paired group was slower in the first half of T2 compared with the new group. This difference was confirmed by a Group x Block interaction, $F(4, 228) = 5.29, p < .001$, $MSE = 2261$. Further analyses with t-tests comparing the two groups per block revealed that the re-paired group was significantly slower than the new group on the first three blocks after switching of keys.
Error Rates

The overall percentage of errors made was 1.3%. Repeated measures ANOVAs were performed on the error rates. The factors for the ANOVAs were group and block. We first address the data from the training phase, Blocks 1-5. There were no significant main effects, nor was there a significant Group x Block interaction. This indicates that the two groups were comparable in speed and learning rate in the training phase of the experiment. There was also no increase in accuracy across blocks, which might be due to the high accuracy at the start of the experiment.

The error rate data for the transfer phase, Blocks 6-10, showed only a main effect for group, $F(1, 57) = 5.19$, $p < .05$, $MSE = 2.35$. This result indicates that the re-paired group made more errors than the new group. The pattern of the errors across blocks was less consistent than the pattern of the RTs. Some caution is needed as the analysis on the errors is less informative because the overall number of errors was very small. So, although statistically significant, the practical significance is limited.

Transfer Index

In many studies of transfer an index of transfer is given by a simple formula (see for reviews on transfer formulas, Gagné, Foster, & Crowley, 1948; Murdock, 1957; Singley & Anderson, 1989). As an illustration of such a transfer index, the transfer effects in this chapter are calculated with Formula 1:

$$\%\text{Transfer} = \frac{C_b - E_b}{C_b} \times 100$$

In Formula 1, $E_b$ is the performance of the re-paired group in the transfer phase and $C_b$ is the performance of the new group in the transfer phase. With Formula 1, a standardised percentage of the amount of savings is obtained where negative values represent negative transfer. Using the mean RTs over the total transfer phase, the amount of transfer given by Formula 1 is $-7.2\%$.

Discussion

The results of Experiment 2.1 partly confirm the results found by Parreren (1951). The re-paired group performed worse than the new group. Negative transfer, i.e. the impeding effect of prior learning on new learning or performance was found in RTs and in error rates. Both learning and performance are mentioned because it is impossible to distinguish which of the two processes is affected. The total number of errors was very small, which may have been the result of the simple character of the task and the eagerness of the participants to prevent errors. This experiment demonstrated that procedural knowledge does interfere with the learning or the retrieval of the new responses. The next question is whether a newly learned skill may
deteriorate the performance on a previously learned skill. To address this question, an experiment was conducted in which the effect of interpolated learning on the performance of a previously learned similar task was investigated. In addition, a design was chosen that also allowed a comparison of the phases instead of only between the conditions in the transfer phase as in this experiment.

**Experiment 2.2**

The main goal of Experiment 2.2 was to investigate the occurrence of interference between two previously learned skills. To do this, a third phase was introduced to retest the training task. This experiment was also an attempt to replicate Experiment 2.1. The main question in the current experiment is whether learning of the transfer task interferes with the memory for training tasks or the reinstatement of it.

**Method**

**Participants**

Fifty-six students of the University of Amsterdam participated in this experiment in exchange for course credit. Every second participant was assigned to the control condition. All participants were tested individually.

**Materials**

This experiment was in many respects the same as Experiment 2.1, apart from the following changes. A different design was used, and as a consequence the use of the keys across phases was different. Four different keys were used (GZVM, ZMGV, JPQT, PTJQ). Each letter was linked to the digit corresponding with its position in the key (e.g., in the first key G-1, Z-2, V-3, M-4).

The letters used for the re-paired group were the same across the phases but differed in position (e.g., ZMGV in T1, GZVM in T2 and ZMGV in T3), and therefore in the corresponding button. The letters used for the new group were different in T2 (e.g., ZMGV in T1, PTJQ in T2 and ZMGV in T3). The keys were counterbalanced across phases.

**Procedure**

The trials were identical to the trials in Experiment 2.1. Stimuli were arranged in 15 blocks, and each block consisted of 40 trials with each letter appearing 10 times in a pseudo-random order. After 5 blocks participants switched to the transfer task and switched back to the original task after 10 blocks. In this experiment, feedback about the mean time of the correct responses was given after each block.
Results

Response Times

Repeated measures ANOVAs were performed on the response time (RT) variable. The factors for the RT ANOVAs were block and group. The analyses were performed on the median RTs per block for every participant. The median RTs were based on the correct responses only.

We first address the data from the training phase (Blocks 1-5). The significant main effect for block, $F(4, 216) = 96.80, p < .0001, MSE = 1945$, confirmed the reduction in RT with practice that can be seen in Figure 2.2. There was no main effect for group, and no significant Group x Block interaction, indicating that the two groups were comparable in terms of speed and learning rate in the training phase of the experiment.

Figure 2.2 Mean correct response time and standard error of the mean as a function of block and group in Experiment 2.2. The vertical lines indicate key changes.

Although the RT data for the transfer phase, Blocks 6-10, showed no main effects for group, $F(1, 54) = 2.52, p = .118, MSE = 42,196$, Figure 2.2 clearly shows that the re-paired group was slower in the first half of T2 in comparison with the new group. This difference was confirmed by a Group x Block interaction, $F(4, 216) = 6.42, p < .0001, MSE = 2171$. A closer inspection of the data within T2 shows that the groups differed in the first two blocks after switching keys (Blocks 6 and 7) according to a t-test, $t(54) = 2.35, p < .05$ for Block 6; $t(54) =$
2.68, $p < .05$ for Block 7. This indicates that the re-paired group was slower in responding after the first switch. This effect lasted two blocks and then the groups were comparable in speed again. The training effect for block was significant, $F(4, 216) = 37.26, p < .0001, MSE = 2171$. The amount of transfer given by Formula 1 is -5.7%.

The RT data for the retention phase, Blocks 11-15, showed no main effect for group, $F(1, 54) = 1.97, p = .166, MSE = 29,976$, indicating that the re-paired group was as fast as the new group when comparing mean RTs of T3. Closer inspection of the data shows that in the first block after switching back to the original task (Block 11), the re-paired group was slower. According to a t-test, Block 11 showed the only significant difference, $t(54) = 2.24, p < .05$, between re-paired letter and new group within T3. The effect for block, $F(4, 216) = 13.13, p < .0001, MSE = 1802$, and the Group x Block interaction, $F(4, 216) = 3.33, p < .05, MSE = 1802$, were significant. Using the mean RTs on T3, the amount of transfer given by Formula 1 is -4.5%.

**Error Rates**

The overall percentage of errors made was 2%. Repeated measures ANOVAs were performed on the error rates. The factors for the ANOVAs were group and block. In none of the three phases of the experiment was there a significant main effect of group or block, or a Group x Block interaction.

**Discussion**

In Experiment 2.2, negative transfer was found and the results confirm the findings of Experiment 2.1. In addition, strong retroactive interference was found. Thus, the interfering effect is not limited to learning alone. When phases are compared, the transfer is obviously positive, whereas the comparison of the groups in T2 and T3 reveals that there is also an interfering component source. Thus, the findings of Experiment 2.2 as well as 2.1 support the hypothesis of masking of negative factors by positive factors, suggested by Siipola (1941).

Although the pattern of results is in line with the earlier experiments of Parreren (1951), the negative transfer and the retroactive interference effect are smaller and more transient than the effects found by Parreren. The smaller effect is possibly caused by the larger influence of positive factors. For the more transient aspect, suggesting a quicker learning rate we have no explanation at this stage. In an attempt to replicate the results of Parreren (1951), a paper and pencil version of the coding task was introduced.

**Experiment 2.3**

The goal of Experiment 2.3 is to replicate the findings of Parreren. Although interference was found in the first two experiments, the effects were relatively short-lived and the initial performance on the transfer task was always better than initial performance on the training task. By conducting an experiment that more closely resembles the Parreren experiment we expect to find net negative transfer that is less transient. There are several important differences between the first two experiments and Experiment 2.3. The first important difference is that a paper and
pencil version of the task was used instead of a computerised version. The second difference is that in the current experiment responses were given in continuous series instead of one at a time.

The task components were not formally stated because most description models (e.g., GOMS) of tasks or skills were not specific enough to make useful differential predictions for our versions of the coding task. Despite the absence of a formal description of the task, it seems likely that the computerised version of the coding task is more difficult than the paper and pencil version. The stimulus presentation and the use of the button box in the computer task were relatively new in comparison with the reading and writing procedures in the paper and pencil test. In the computerised experiment, for example, the exact placement of the buttons had to be learned to operate well. It is likely that the number of factors working in a positive direction is smaller in the easier paper and pencil version. In a simple task, the masking effect of positive factors may decrease and therefore the negative effect will become more visible.

As in Experiment 2.1, the main goal of Experiment 2.3 is to examine whether negative transfer from a training task (T1) to a transfer task (T2) occurs in skill acquisition. To investigate this a modified version of the coding task as described by Parreren (1951) was developed. In our experiment the keys were counterbalanced to rule out the alternative hypothesis that findings could be due to differences in difficulty between the training and transfer key. In addition, a control group was used to differentiate between a decrease in RT due to learning the new key or due to negative transfer. Experiment 2.3 was designed to give answers to several questions. First, does a reversal of responses hinder the performance in comparison with a situation in which a new set of responses is learned? Second, has the training task the same effect on the learning rate in T2 in both conditions? Third, which measure is affected, response times, accuracy, or both? Finally, how large and long lasting is the effect?

In this experiment two conditions in letter-digit coding were compared. After training, the experimental group switched to a key with re-paired letter-digit combinations, whereas the control group switched to a key consisting of different letters paired with the same digits. It was expected that the experimental group would perform worse in the transfer phase than the control group.

Method

Participants

Sixty-four students of the University of Amsterdam participated in the experiment in exchange for course credit. Each participant was assigned randomly to one of two conditions and was tested individually.

Materials

In this paper and pencil experiment participants had to write digits under letters in accordance with a key. The key consisted of four letter-digit combinations (capital letters above digits) and was printed at the top of the form (size A4). The letters in the key were linked with
digits 1 to 4. Four keys were used in this experiment (NRBD, RDNB, WHKF, HFWK). These four keys were counterbalanced across phases (T1 and T2) to eliminate possible influences of differences in difficulty between the keys.

The response field consisted of six rows of twenty letters. Under each letter was a space to fill in the digit. Eight forms that differed in the presentation order of the to be substituted letters were used. The orders used in the eight forms were identical for the four keys. The presentation order of the forms was counterbalanced to investigate the training effect without possible influences of the differences in difficulty between forms. Each subject completed 14 forms consisting of 8 training forms and 6 transfer forms. To make the two phases better comparable the presentation order of the forms in the training phase was identical to that in the transfer phase.

Procedure

An example of the form without the letters was used for the instructions. Participants were instructed to fill in the form as rapidly and accurately as possible. Digits had to be filled in from left to right. The experimenter measured the total time (TT) required to complete each form. The stopwatch was started when the participant received the form, and was stopped when the last digit was filled in. Between each form was a short break during which feedback was given about the time. During these breaks no key was present.

The critical difference between the training task and the transfer task were the letter-digit combinations used. After eight forms participants were informed that the key would change. The experimental group, hereafter referred to as re-paired group, switched to a key with the same letters as in the training task, but in other combinations (e.g., from NRBD into RDNB). The control group, hereafter referred to as new group, switched to a key consisting of letters, which were different from the letters in the training task (e.g., from NRBD to HFWK). The session ended after completion of six transfer forms.

Results

Response Times

Repeated measures ANOVAs were performed on the total times (TT). The factors for the TT ANOVAs were group (new group, re-paired group) and form.

We first address the data from the training phase (Forms 1-8). The significant main effect for form, $F(7, 434) = 168.62, p < .0001, MSE = 86.14$, confirmed the reduction in TTs with practice that can be seen in Figure 2.3. There was no main effect for group, and no Group x Form interaction, indicating that the re-paired and the new group were comparable in terms of speed and learning rate in the training phase of the experiment.

The TT data for T2, Forms 9-14, show main effects for group, $F(1, 62) = 13.29, p < .001, MSE = 2882$, and form, $F(5, 310) = 86.27, p < .0001, MSE = 113.42$. The main effect for group indicates that the re-paired group was slower in T2 than the new group in T2. The main effect for form indicates that there is an increase in the performance across the forms. In
addition, there was a significant Group x Form interaction, $F(5, 310) = 3.43, p < .01, MSE = 113.42$. The interaction shows that this increase is not the same for the groups, which is possibly due to the large difference in Form 10.

The difference between the groups, as can be seen in Figure 2.3, remains significant across the forms. Testing of the simple effects revealed that the means of group for every transfer form differed at $p < .01$. In short, although there is an interaction, the negative transfer effect remains relatively stable. As an illustration of the negative impact of the re-pairing in this very simple task, the TTs on the first form of T1 and T2 were compared for the re-paired group. The marginally significant difference given by the t-test, $t(31) = 1.73, p = .093$, demonstrates that the negative effect is at least as large as the positive effect. Using the mean TTs over the total transfer phase, the amount of transfer given by Formula 1 is $-14.7\%$.

**Error Rates**

The overall error rate was very low (0.3%), indicating that participants were working very accurately. Repeated measures ANOVAs with the factors group and form were performed on the error rates. The only significant effect was a main effect for form in the training phase, $F(7, 434) = 5.94, p < .0001, MSE = .51$, indicating that the decrease in errors during the training was reliable.
Discussion

The results of Experiment 2.3 confirmed the results found by Parreren (1951). The repaired group performed worse than the new group. Massive negative transfer was found. This effect was so large that the re-paired group was slower at the start of the transfer phase than at the start of the training phase. The difference in TTs between the groups remained significant across all forms in the transfer phase. The relatively constant difference between the two groups is an indication that the speed of learning in the transfer phase did not differ between groups. The negative transfer was only found in TT, and not in error rate. The total number of errors was very small, which may have been the result of the simple character of the task.

Negative transfer, the impeding effect of prior learning on new learning, is evident in the transfer phase. The acquisition of the new skill is impeded by a procedure learned before. Although training time was short, the tasks were relatively well automated. The automaticity is supported by the relatively small angle of inclination at the end of the training and transfer curve. In addition, some participants remarked spontaneously that the task was easier when they did not think about their responses. Although anecdotal, these remarks suggest that their behaviour was relatively automatic. Experiment 2.3 clearly shows that procedural knowledge interferes with the learning of new responses. It is not clear whether the negative transfer in the transfer phase is only caused by the hindering influence of the first-learned task on the acquisition of the second task, or whether there is also a problem in automatic response selection. These two processes can not be differentiated because both lead to the same outcome, an increase in response time. By adding a third phase in which the already learned skill is re-tested retroactive interference can be studied. This may give more insight into interference in situations where both responses are well trained.

A possible explanation for the quicker disappearance of the negative transfer effect in the computerised task than in the paper and pencil tests, may be that the responses in the paper and pencil test are better automated than the responses in the computerised test and may therefore lead to stronger interference. The training consisted of 200 responses in Experiment 2.1 and 960 responses in Experiment 2.3. However, in comparing the computerised experiment with the paper and pencil test some caution is needed because the tasks were different in many respects.

Experiment 2.4

The main goal of Experiment 2.4 was to investigate the occurrence of negative transfer and retroactive interference in a paper and pencil coding skill. The main question added to the questions already asked in Experiment 2.3 is; does learning of the transfer task interfere with the memory for or the performance of the original training task? Or, in other words, do two well-learned procedures interfere with each other?
Method

Participants

Forty students of the University of Amsterdam participated in this experiment in exchange for course credit. Each participant was assigned randomly to one of two conditions and was tested individually.

Procedure

In Experiment 2.4, identical stimuli and procedures were used as in Experiment 2.3. There were two differences between Experiment 2.3 and 2.4. The first difference was the addition of a third phase, a retest phase (T3), to investigate retroactive interference. The second difference was a shortening of the training (T1) and the transfer phase (T2) to five forms, which was necessary to keep the time of testing within reasonable limits.

Results

Response Times

Repeated measures ANOVAs were performed on the total times (TTs). The total time was the time needed for completion of a form. The factors for the TT ANOVAs were group and form.

We first address the data from T1 (Forms 1-5). The significant main effect for form, $F(4, 152) = 72.23, p < .0001$, $MSE = 99.40$, confirmed the reduction of TT with practice that can be seen in Figure 2.4. There was no main effect for group, nor was there a Group x Form interaction, indicating that the two groups were comparable in terms of speed and learning rate in T1. Note that in T2 the re-paired group was slower than the new group, whereas in T1 the re-paired group was somewhat faster than the new group. Thus, the negative transfer and the retroactive interference effect in T2 might be somewhat underestimated.

The TT data for the transfer phase, Forms 6-10, show main effects for group, $F(1, 38) = 9.24, p < .01$, $MSE = 1983$, and form, $F(4, 152) = 76.66, p < .0001$, $MSE = 101.98$. These results indicate that the re-paired group was slower than the new group in T2 and that both groups became faster during the course of T2. There was no Group x Form interaction, $F(4, 152) = 2.26, p = .065$, $MSE = 101.98$, indicating that the negative transfer effect remained constant over the forms. The amount of transfer given by Formula 1 is -15.0%, which is comparable with the amount of transfer found in Experiment 2.3. When the TTs on the first forms of the training phase and the transfer phase for the re-paired group were compared, there is according to a t-test, an increase in total time, $t(19) = 4.37, p < .001$. Thus, there is a clear net negative effect of digit-letter re-pairing in this experiment.

The TT data for the retention phase, Forms 11-14, show no main effect for group, $F(1, 38) = 1.05, p = .31$, $MSE = 2482$, indicating that the re-paired group was as fast as the new group when comparing the mean times of T3. Closer inspection of the data shows that the re-
paired group was slower in the first form after switching back to the original task (Form 11). According to a t-test, form 11 showed the only significant difference, $t(38) = 18.85, p < .05$, between the re-paired group and the new group within T3. The effect of form, $F(3, 114) = 48.61, p < .0001, MSE = 61.44$, confirms that there is still a decrease in TTs across forms. The Group x Form interaction, $F(3, 114) = 8.94, p < .0001, MSE = 61.44$, was significant, indicating that there was a different learning rate in T3 for the groups, with the re-paired group having a faster decrease of TT. This faster decrease of the re-paired group is caused by the slower TT in form 11. Using the mean TTs over the total retention phase, the amount of transfer given by Formula 1 is -7.1%.

![Figure 2.4](image)

**Figure 2.4** Mean total time per form split by condition with standard error of the mean in Experiment 2.4. The vertical lines indicate key changes

**Error Rates**

The overall error rate in Experiment 2.4 was very low (0.4%). Repeated measures ANOVAs were performed on the error rates. The factors for the ANOVAs were group and form. There were no main effects for group or form, nor was there a Group x Form interaction in error rates in the three phases.

**Discussion**

As in Experiment 2.3, a clear negative transfer effect was found in Experiment 2.4. The reversal of letters led to a negative transfer effect on an aggregate level, meaning that the re-paired group was even slower than its entry level in the training phase. The amount of negative
Interference in Coding Skills

transfer and the number of errors were almost the same for Experiments 2.3 and 2.4. As in Experiment 2.3, the difference in speed was relatively constant across the transfer blocks, and no substantial release of negative transfer was present in the transfer phase.

In addition, an effect of retroactive interference was found. The retroactive interference had a considerable impact, but was slightly shorter in duration than the negative transfer effect. There are several plausible explanations for the shorter duration. A possible explanation is that the first learned responses are so strong that these are quickly restored and the second learned responses are too weak to interfere strongly with them. A related and complementary explanation is that the transfer task was less well trained than the training task because of the acquisition interference. Because the responses of the transfer task were possibly weaker the interfering effect could be smaller and less dramatic.

A Word on Sex Differences in Coding

One of the main characteristics of standard symbol-digit substitution tests (e.g., in the WAIS) is that women perform better than men (Harrison & Whissell, 1980; Laux & Lane, 1985; Royer, 1978; Wong & Gilpin, 1991). In a computerised substitution test, where letters were used instead of symbols, however, Laux and Lane (1985) did not find a sex difference. In our experiments, sex differences in a paper and pencil (81 women and 22 men) and a computerised version (81 women and 34 men) of a letter-digit substitution test could be studied as a byproduct of the interference study. For every subject a mean response time and error rate was computed across the first five training forms. According to t-tests, there were no sex differences in response times or error rates for the computerised and the paper and pencil versions of the coding task (all four p’s > .5) Thus, there is no apparent sex difference in both variants of the letter-digit substitution

Table 2.2. Means and Standard Deviations of Total Times (in seconds for paper and pencil test and milliseconds for the computerized test) and Error Rates as a function of Sex

<table>
<thead>
<tr>
<th>Forms</th>
<th>Women TT</th>
<th>Men TT</th>
<th>Women Error</th>
<th>Men Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>131.5 (22.7)</td>
<td>127.9 (23.2)</td>
<td>.52 (.62)</td>
<td>.49 (.73)</td>
</tr>
<tr>
<td>Comp.</td>
<td>751.0 (106.8)</td>
<td>744.4 (98.4)</td>
<td>.78 (.69)</td>
<td>.75 (.63)</td>
</tr>
</tbody>
</table>

This result is in concordance with the absence of a sex difference in letter-digit substitution in the young adult sample in the study by Laux & Lane (1985). Our results show that the sex difference in symbol-digit substitution is not universal and cannot be generalised to all forms of substitution tasks. The finding that the sex difference is absent in letter-digit substitution is consistent with the hypothesis of Estes (1974) that the superior performance of
women on the symbol-digit test is due to better verbal recoding of the abstract symbols. In letter-digit substitution no verbal recoding is necessary because the well-known letters serve as codes themselves. As Estes concluded the difference in ability to recode symbols in verbal labels is a major candidate for explaining sex difference in symbol-digit substitution. A crucial experiment (comparison of sexes after each subject coded with letters and symbols) is needed to confirm the role of stimulus type in coding. This is, however, beyond the scope of this project and will not be further elaborated.

General discussion

Negative transfer was found in all four experiments. Comparisons of the entry levels of the training and the transfer phase showed a marginal net negative effect in Experiment 2.3 and a clear net negative effect in Experiment 2.4. This means that the negative transfer effect was so large that the performance of the participants in the re-paired groups during the transfer phase was worse than their own entry level. Thus, the negative effect of the re-pairing of responses in the transfer phase was larger than the positive effects caused by training. The effect of the negative transfer in the paper and pencil experiments remained relatively constant across the total transfer phase.

In addition, retroactive interference was found. Experiments 2.2 and 2.4 showed that responses of two well-learned tasks do interfere with each other. Thus, interpolated learning of a similar task is interfering with the execution of a well-learned task. The retroactive interference effect was, however, smaller than the acquisition interference effect but lasted at least one form or block. The disruption in performance in the transfer and retest phase of the re-paired groups was always found in the response times, but in error rates the interference effect was only apparent in Experiment 2.1.

The decrease in performance after switching to a different but similar task is often interpreted as a consequence of the learning of the new information. Our experiments clearly show that the relapse in performance is not only caused by learning of the new skill but also by interference. In short, a well-learned task can influence the learning of a similar task in a negative way. In situations where stimuli are identical and responses are varied negative transfer is obtained. The interference effects in skill acquisition can be large when using simple tasks that are comparable in complexity with verbal learning. So the statement that interference and negative transfer in learning of skills is nonexistent except for Einstellung (Anderson, 1995) is not supported by these experiments.

Can this negative transfer effect be interpreted as a form of Einstellung? The concept of Einstellung is used in research on problem solving and one of its characteristics is that subjects are not aware of the easy alternative and keep on using the previously learned nonoptimal method. Einstellung disappears at the moment the experimenter points at an easier solution (e.g., the “don't be blind” instruction in the water jug task). In most interference studies, and certainly in the here reported experiments, the subjects are fully aware of the different and reversed response alternatives. This knowledge, however, does not lead to a disappearance of the interference effect. Interpreting the negative transfer in our experiments as Einstellung is therefore unjustified.
Interference in Coding Skills

The findings of the reported experiments show that learning one skill can disrupt the learning and retrieval of another similar skill. It is often suggested that negative transfer is a limited problem, which exists only between related tasks (Annett & Sparrow, 1985). Yet, negative transfer and interference do exist in learning of skills and, as our experiment show, can be obtained in a short time with relatively little training. Although the problem of negative transfer may be limited in incidence, it should be of great concern in situations where closely related tasks are performed (e.g., flying different types of aircraft). Surprisingly there is almost no research published on the subject of negative transfer and aviation (Rayman, 1982). Thus, stating that negative transfer and interference in skill acquisition do not exist can have dangerous implications for learning in applied fields. Although our experiments did not show an increase in error rate it is possible that in situations with more workload and stress the error rate would increase more in the re-paired condition.

In the current experiments the focus was on negative transfer and retroactive interference on the short term. Therefore, no sound conclusions for the occurrence of negative transfer on the long term are possible. According to Schmidt and Bjork (1992), performance during training is not always representative for performance on tests after a longer interval. For instance, in contextual interference research a random schedule of training that degrades performance during acquisition compared to a blocked schedule, enhances performance at retention. This could mean that the re-paired group in our experiments, which performed worse after a short interval, would perform better than the new group when there is a longer interval between the transfer and retention test.

Based on these results we can only speculate about the mechanism that leads to the negative transfer and interference. According to some researchers negative transfer effects in motor skills are essentially cognitive rather than motor (Blais, Kerr, & Hughes, 1993; Magill, 1989). This is likely as long as cognitive is not the same as declarative. In our view, the mechanism that leads to interference takes place on a very low cognitive level that is not accessible for introspection. The cause for the interference (reversal or re-pairing of responses) or the outcomes (the slowing down of the responses) can be stated, the interference process itself is implicit. We suggest that by using an A-B, A-Br design, interference occurs in verbal learning as well as skill learning. We suggest that the underlying mechanisms for the replacement of old responses with new responses are the same for skill learning and verbal learning. To have a flexible and at the same time robust memory, the storage of new responses in combination with old stimuli is necessary, but without the immediate destruction of the previous combinations.

It seems that the supposed difference in sensitivity to interference in verbal learning and skill learning experiments is due to differences between variables such as task complexity, context, and amount of training. Before reaching conclusions on mechanisms themselves, the outcomes of these mechanisms have to be thoroughly studied and stated in clear language. Until now there is for reasons of parsimony no ground to assume qualitatively different transfer mechanisms mediating verbal learning and skill learning.