

The Effect of Retroactive Interference on Anatomical Learning  
using a Virtual Learning Environment

BartJan Dieperink  
University of Twente

December 11, 2006

Examiners:  
Prof. Dr. Ing. W. B. Verwey  
J. M. Luursema

### Abstract

The aim of the present study was to test whether retroactive interference (RI) effects might occur in anatomical learning using a virtual learning environment, in order to give recommendations for adapting such an environment to the needs and capacities of its users to minimize possible negative RI effects. RI refers to situations in which subsequent learning negatively affects the recall of earlier learned material. Participants studied the human abdomen in a virtual study phase after which they were tested on this knowledge. Then, three groups were formed with similar visuospatial ability and each group was assigned to a different experimental condition. The first condition included studying the human eye. The second condition involved studying an imaginary three-dimensional figure, the Sphere. The third, control, group did not study additional material but watched a short movie instead. At the end of the experiment, all participants were again tested on their knowledge of the human abdomen. It appeared that relative to the first test accuracy of the Sphere group improved, and accuracy in the Eye and Control groups reduced. These results are in line with RI in that anatomical knowledge of the human eye reduced abdominal knowledge in the Eye group more than in the Sphere group, and that performance of the Control group reduced because they were less skilled than both other groups in test-specific skills. Visuospatial ability (VSA) did not significantly affect RI.

**Keywords:** retroactive interference, anatomical learning, virtual learning environments, visuospatial ability.

## 1 Introduction

This study addresses a phenomenon in memory research and applies it to a modern technology. The modern technology is computer-assisted learning of anatomy, the phenomenon is *retroactive interference (RI)*. A longstanding theory of forgetting stipulates that learning something new can interfere with, and hence cause the forgetting of, knowledge of things already acquired. Indeed, interference effects are a well-known cause of forgetting. Because interference impairs memory performance, it is important that the causes of interference are isolated and that is specified how these effects can be diminished or eliminated.

### 1.1 Retroactive Interference

Generally speaking, in interference in human memory by learning other information, two forms can be distinguished. First, *proactive interference (PI)* refers to situations in which prior learning negatively affects the later learning of new material. Second, *retroactive interference (RI)* refers to situations in which subsequent learning negatively affects the recall of earlier learned material (Marsch et al., 1998). Typically, interference occurs in situations where learning materials or memory tasks are somewhat similar (Crowder, 1976). RI is classically investigated in a paradigm that requires subjects to learn material, mostly verbal, on a first list, after which a test assesses subjects' memory of the learned associations. Then material on a second, or interpolated list is learned; subjects are then tested for the second time on the material from the original list. Results typically show poorer recall of the first list as compared with a control condition in which no second list is given. RI is indicated by a drop in performance on recall of the first list between the first and second test, compared to the performance on recall of the control condition. RI is usually strongest on an immediate test. Over time, the RI group's production of the original responses becomes more similar to that of a control group, a phenomenon known as *spontaneous recovery*.

A factor known to affect RI is similarity between the original and the interpolated list in terms of content (Baddeley & Dale, 1966). Performance on a memory test tends to be more accurate for events that are followed by dissimilar events (control condition) rather than by similar events (experimental condition). RI is consistently obtained with recall procedures, but is typically weak or absent when recognition tests are used (Crowder, 1976).

While the phenomenon of RI has been well documented, most of the studies on this topic have dealt with verbal materials. Only a small number of studies have focused on RI with materials, such as 2-dimensional (2D) images, that call upon *visuospatial ability (VSA)*. Visuospatial ability is the capacity to construct, manipulate and retrieve mental, visuospatial representations of objects and environments. It allows people to transform 2-dimensional images to visuospatial 3D mental representations and to mentally rotate these representations (Gordon, 1986). Studies have shown that retention of visuospatial items in working memory is impaired when subjects perform interpolated

tasks requiring visuospatial attention (Beech, 1984; Logie, Zucco & Baddeley, 1990), requiring memory of additional visuospatial items (Logie et al., 1990; Hole, 1996) or when subjects are exposed to task-irrelevant visuospatial material during the memory delay (Toms, Morris & Foley, 1994; Hole, 1996; Quinn & McConnell, 1996). In the present study, focus is on differences between people of high and low VSA, to find out if one of these groups is especially vulnerable to retroactive interference effects. Alternative strategies and/or codes used in relation to visuospatial working memory have previously been reported for people of different general cognitive ability (Gevins & Smith, 2000) and for people of different gender (Weiss et al., 2003). Maybe people of low and high VSA also use different strategies when dealing with visuospatial problems, perhaps in the way stimuli are encoded, for example in a more semantic or visuospatial form.

## 1.2 Anatomical Learning

Success on many clinical tasks (i.e. physical examination, radiology interpretation, surgery) is dependent on visuospatial anatomical knowledge. Medical students traditionally learn human anatomy by self-education with the help of anatomical atlases, with the help of anatomical manikins in the dissection room, or with a combination of these two methods. These methods have limitations. Books are not interactive and are not able to portray anatomy in three dimensions. In vitro training models made of synthetic materials can be useful, but it would be difficult to maintain a library of models with all important pathologies and anatomical variations. Training in the dissection room offers the possibility of studying real human anatomy, but this is expensive and takes a lot of maintenance. Computer-based training with the use of virtual learning environments has many potential advantages. It is interactive and an instructor's presence is not necessary, so students may practice in their free time. Any pathology or anatomical variation needed for training can be created. Students can also look at anatomy from perspectives that would be impossible in real life, even during surgery (Tendick & Cavusoglu, 1997). More recent developments in the field of virtual learning environments offer excellent opportunities for studying human anatomy in 3D (Jastrow & Vollrath, 2003). Various virtual reality (VR) applications for teaching anatomy are based upon data sets of the Visible Human Project (Ackerman, 1998). Examples are AnatQuest from the National Library of Medicine (NLM) (AnatQuest, 2004), and a recent development of a customizable web-based three-dimensional (3D) anatomy training system, W3D-VBS (Temkin et al., 2006). Despite all these efforts at building virtual anatomical learning environments, basic questions with respect to the transfer of skills and/or knowledge from a virtual learning environment to the real world are still largely unanswered. The studies that do exist report contradicting results. Stanney (2002) found that knowledge can be obtained with the use of a virtual learning environment, while Garg et al. (1999, 2001, and 2002) concluded over a series of experiments that 3D anatomical models do not seem to benefit anatomical learning more than the use of traditional atlases.

The present study attempts to determine whether retroactive interference effects could occur in virtual anatomical learning. Contrary to anatomical atlases in which students can determine their own pace and order in which they learn different anatomical parts, computerized learning environments may lack this kind of individual adaptations. It is therefore important to study if and when RI effects occur in VR learning environments too, in order to optimize the use of virtual learning environments. The possible influence of visuospatial ability was also taken into account when studying RI effects. Visuospatial ability is important for medical students who are learning anatomy (Garg, Norman & Sperotable, 2001). It is also related to competency and quality of results in complex surgery (Wanzel et al., 2002). If there would be an effect of visuospatial ability, adjustments can be made to the study program to make it suitable for the needs and capacities of the individual user. Taken together, the goal of this study is to provide recommendations for development of virtual learning environments for teaching human anatomy.

Recently, the University of Twente (UT) has become involved in the Distributed Interactive Medical Exploratory (DIME) Project. One of the goals of this project is to create a medical training support system for learning and exploration in virtual reality. The role of the University of Twente is to optimize the presentation of virtual visualizations to the trainees. The question is: how these virtual learning environments can be made more efficient, so that the rate of learning is optimal? One thing needed is a better adaptation to the needs and capacities of the user.

Earlier DIME research at the UT indicated that offering material for studying human anatomy in a virtual learning environment that supports both *stereopsis* and dynamic exploration has a positive effect on subsequent tests (Luursema et al., 2006). Stereopsis is a mechanism of the human visual system that contributes to the perception of depth. The patterns of light projected on the retinas of the left and the right eye are slightly disparate, since the eyes themselves are not in the exact same location. The degree of disparity of the two images provides a basis for the judgment of distance and depth (Wheatstone, 1838). Participants of differing visuospatial ability learned about human abdominal organs via anatomical 3-dimensional reconstructions using either a stereoptic study phase (involving stereopsis and interactivity) or using a biocular study phase that involved neither stereopsis nor interactivity. Subsequent tests assessed the acquired knowledge. Results showed that the stereoptic group performed significantly better on both tasks and that participants of low visuospatial ability benefited more from the stereoptic study phase than those of high visuospatial ability.

### 1.3 The Present Study

The purpose of the present study was to ascertain whether retroactive interference effects are prevalent in anatomical learning using a virtual learning environment. More specifically, there were three questions: First, are retroactive interference effects prevalent in virtual anatomical learning? If so, virtual learning environments have to be designed in such a way to keep RI effects at a minimum. Second, what is the effect of the content offered after the material to be studied on recall of that

original material? Does content that is more similar to the original material cause more RI than content that is less similar? Third, does visuospatial ability play a part in the extent of the RI effect?

To study these questions, the following experiment was executed. First, the visuospatial ability of the participants was assessed, and the results were used to create three equal groups with respect to VSA. All three groups then studied human abdominal anatomy in a 3D virtual environment. Subsequently, acquired knowledge of the abdominal anatomy was assessed in the first knowledge test phase by two different tests. The *localization task* required participants to select the correct horizontal level of a Computed Tomography (CT) based anatomical cross-section in a frontal view of the studied abdomen. The *identification task* required participants to identify a highlighted anatomical structure of the human abdomen in a CT-based anatomical cross-section. After the first knowledge test, a same-domain group (the *Eye group*) studied a 3D anatomy model of an eye using the same virtual environment as used when studying the abdomen anatomy. The other-domain group (the *Sphere group*) studied a 3D model of an imaginary object using that virtual environment, and the *Control group* studied no additional material. Finally, all three groups were tested again on their knowledge of the human abdominal anatomy in a second knowledge test phase using the same two tests as in the first test phase (with different items though).

We expected that due to different degrees of RI the Control group would outperform the Sphere group on the second knowledge test, which in turn would outperform the Eye group.

## 2 Method

### 2.1 Participants

A total of 48 participants took part (32 women and 16 men). Participants were students from the University of Twente, who either received extra course credit or were paid for their participation in the study. Participants were between 18 and 30 years of age ( $M = 21.87$  years). All reported normal or corrected to normal vision.

### 2.2 Procedure and design

The study consisted of two parts: a pre-test and the actual experiment. The pre-test (described later) was administered to four groups of participants. Based on the results of this pre-test three groups of participants were formed of equal size and with an equal distribution of visuospatial ability. This was done with the help of a list of ascending pre-test scores from which participants were selected and assigned to an experimental group. Each group was assigned to a different condition for the actual experiment (see Figure 1). During the experiment proper, participants were tested individually.

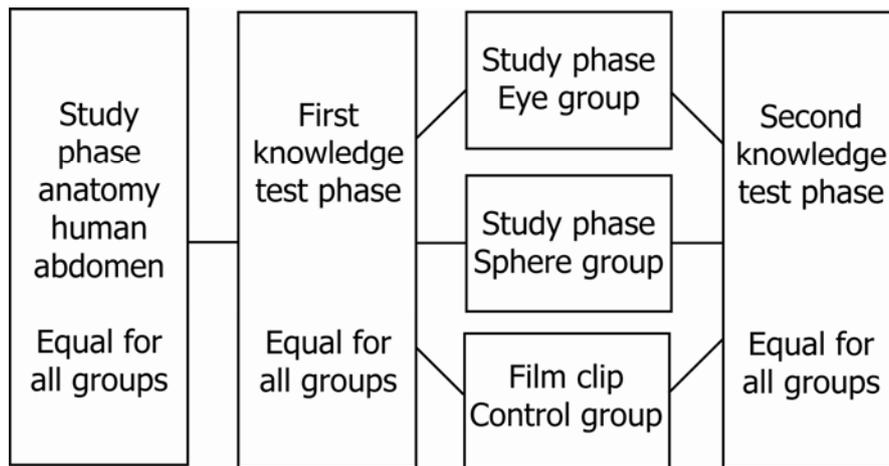


Figure 1. Overview of the phases for the three groups.

The participants first completed a stereoptic study phase of the human abdomen. After that, their obtained knowledge was measured by a knowledge test that consisted of ten items of a localization task and ten items of an identification task (see paragraph 2.5). In the next study phase, the procedure differed for the three experimental groups. The same-domain group completed a study phase on the human eye. The other-domain group was given a study phase about a Sphere (Figure 2), an imaginary 3D object.

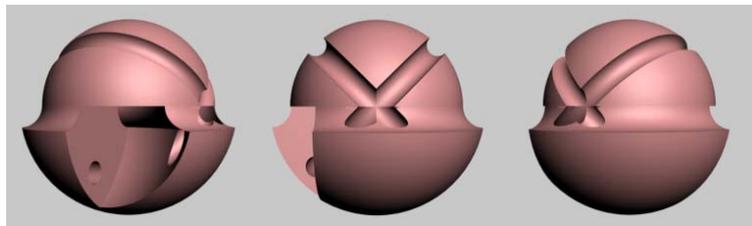


Figure 2. The Sphere object seen from three different angles.

The control group received no extra study phase. Instead, they watched a short film clip. After this, participants' knowledge of the human abdomen was again assessed with the help of the test consisting of ten items of a localization task and ten items of an identification task. These ten items of both tasks differed from those that were used in the first knowledge task. The twenty items of the knowledge task were randomized to control for the effect of order. The dependent variables, measured by the localization and identification task, were latency (reaction time) and accuracy. The different study phases are described in more detail in section 2.6.

The study phase and both knowledge tests were administered to each participant individually in a specially prepared cubicle at the University of Twente. The cubicle contained the hardware and software that were necessary for the experiment and was shut off from possible disturbances during the experiment. All explanations were provided in an instruction booklet and on the computer screen. The cubicle was equipped with a camera so that the experimenter could monitor the participant, to check whether the experiment was carried out correctly.

### 2.3 Pre-test

Prior to the actual experiment, 48 participants took part in a pre-test to test their visuospatial ability which was to be taken as an independent variable in the study. The pre-test is a group test, for which the total number of participants was split up into four groups. It consists of a translated version of the redrawn version of Vandenberg and Kruse's mental rotation test (MRT; Vandenberg & Kuse, 1978; Peters et al., 1995; Luursema, 2004). The mental rotation test consists of a series of test items with a target stimulus and four sample stimuli, of which two are identical but rotated copies of the target stimulus (Figure 3). The participant has to indicate which two of the four sample stimuli represent rotated copies. There are 24 items in the test, and each item has two and only two correct matches. A score between 0 and 24 can be obtained by the participant. In the pre-test a second test was used, the Guay-Lippa Visualization of Viewpoints (VVT; Guay & McDaniels, 1976; as modified by Lippa, Hegarty & Montello, 2002). However, the results showed that this test did not correlate much with performance and it is therefore not mentioned any further in this article.

2.a

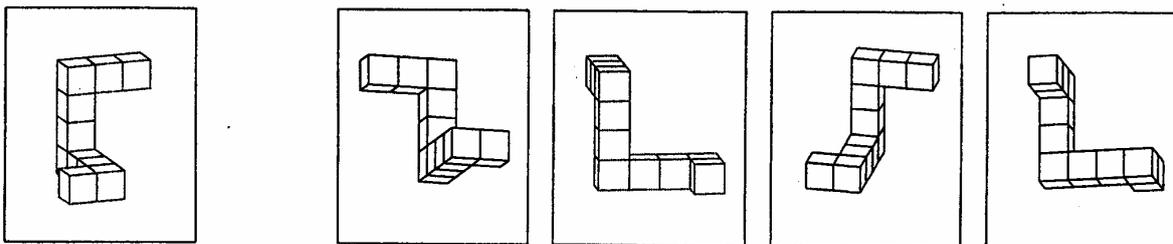


Figure 3. An item of the Mental Rotation Test used in the first pre-test. Participants are to indicate which two of the four stimuli on the right are rotated versions of the target stimulus shown on the left.

In this case, number two and three are the correct answers.

### 2.4 Human abdomen study phase

The first phase of the experiment was the same for all three groups. Participants were provided with an instruction booklet and oral instructions at the beginning of the experiment, after which they watched a self-rotating 3D-reconstruction of the human abdomen (Figure 4) using shutter-glasses to perceive

stereoptical depth. Each model was presented for the same amount of time and rotated in the same direction, so each participant spent exactly the same time watching each model, and was exposed to the same views of the model. The participants had no influence on the reconstruction during the study phase; mouse and keyboard input were not effective. As shown in Figure 4, the 3D-reconstruction was accompanied by labeled reference figures for the eleven anatomical parts of the abdomen relevant to the anatomical study phase and the localization and identification tasks. The names of the eleven anatomical parts are listed in Appendix A. Participants were given three minutes to learn the shape and location of the eleven anatomical parts of the human abdomen. After the three minutes, the reconstruction automatically stopped, and participants were instructed to start the first knowledge test.

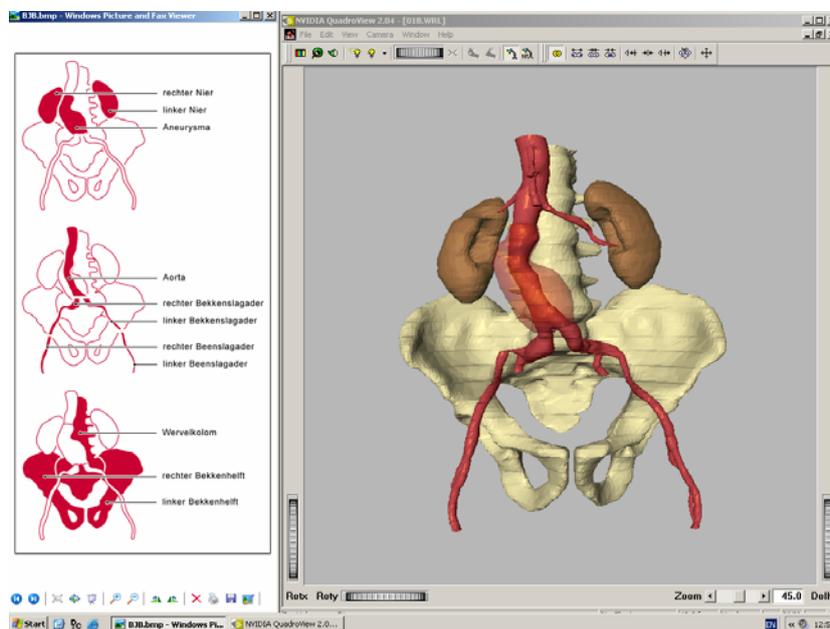


Figure 4. Screenshot of the human abdomen study phase.

## 2.5 Knowledge tests

After the human abdomen study phase, the obtained knowledge of the participants was assessed with a localization task and an identification task. These two knowledge tests have been shown to correlate well with spatial ability (Luursema et al., 2006). This phase was the same for all three experimental groups. Participants were given three practice trials of both tests, after which 20 test trials were presented. These consisted of ten localization and ten identification trials, which were all presented to participants in random order. The two tasks were not blocked. Every trial had a maximum duration of fifteen seconds. Participants were instructed that it was more important to get the answer right than to answer as quickly as possible. Accuracy was judged more important than latency, in line with Lohman

(1988), who found that accuracy, rather than speed, is an important source of individual differences on most visualization tests.

### 2.5.1 Localization task

The localization task required participants to select the correct horizontal level of a Computed Tomography (CT) based anatomical cross-section in a frontal view of the studied abdomen (Figure 5). No shutter-glasses were used during this task. The cross-sections were taken from the scans that had been used to develop the material for the human abdomen study phase. The trial started when the cross-section appeared on the screen. Participants were instructed to identify the location of the cross-section, and then select the corresponding horizontal line on the frontal abdominal model by pressing the key of the letter that labeled the line. There were 23 possible horizontal levels. The order of these letters was randomized on each next trial, so that participants could not predict the letter corresponding with a certain line from previous items. The order of letters per items was the same for all participants however.

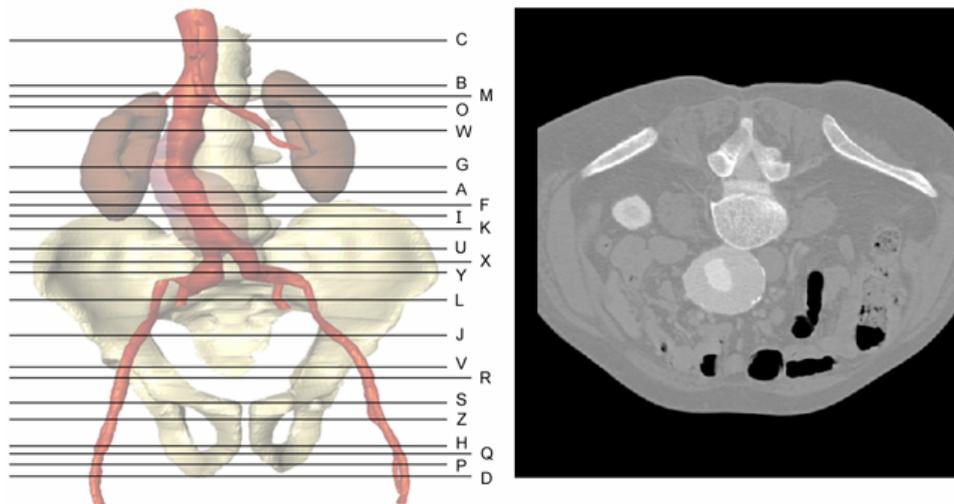


Figure 5. Example of an item of the localization task. Here the line with the letter ‘I’ on the left corresponds with the cross-section shown on the right.

If the participants had not responded within fifteen seconds the cross-section automatically disappeared from the screen and an error was scored. Reaction time was defined as the time between the start of the trial and when the participants pressed a key. A correct answer was defined as pressing the letter key corresponding exactly with the cross-section. Error feedback was supplied after each trial.

### 2.5.2 Identification task

The identification task required participants to identify a highlighted anatomical structure of the human abdomen in a CT-based anatomical cross-section (Figure 6). No shutter-glasses were used during the task. The trial started when a CT-based anatomical cross-section with the highlighted anatomical structure appeared. Also visible was a list with names of the eleven anatomical structures that had been shown during the human abdomen study task. This list was presented to the right of the cross-section. Again, the cross-sections were taken from the scans that had been used to develop the material for the human abdomen study phase. Participants were instructed to identify the highlighted structure in the cross-section, and then to press the key of the letter that labeled the corresponding name in the list on the right at their own pace. As with the localization task, the order of these letters was randomized per item.

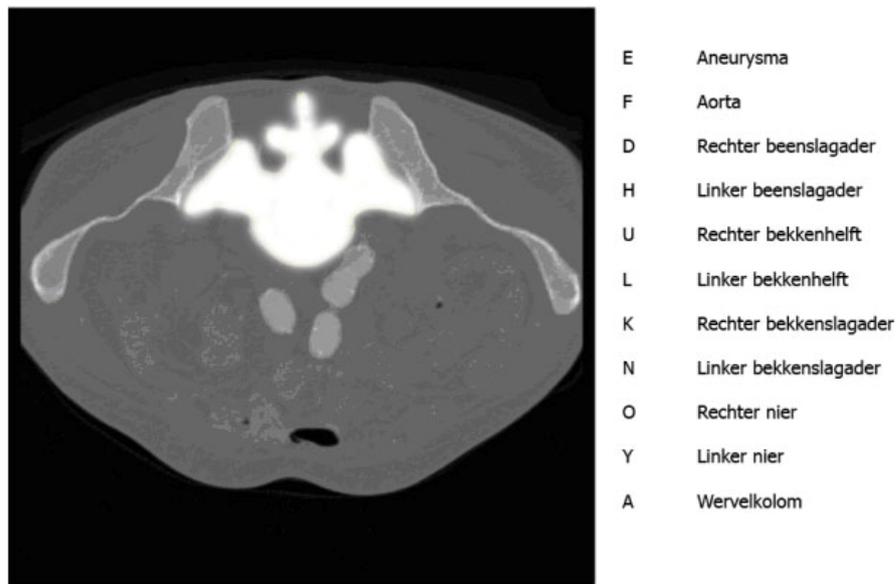


Figure 6. Example of an item of the identification task. Here ‘A’, the spinal column (wervelkolom) is highlighted.

After fifteen seconds the cross-section automatically disappeared from the screen even if the participant had not pressed a key, and an error was scored. Reaction time was defined as the time between the start of the trial and when the participants pressed a letter key. A correct answer was defined as pressing the letter key corresponding exactly with the cross-section. Error feedback was supplied after each trial.

## 2.6 Second study phase

The third phase of the experiment consisted of an extra study phase for the Eye and Sphere groups, while the Control group watched a film clip.

### 2.6.1 Eye group study phase

After the obtained knowledge of the human abdomen was assessed with the knowledge task, the Eye group studied another 3D-representation, in this case of the anatomy of the human eye (Figure 7). The 3D-reconstruction was accompanied by a labeled reference figure for seven anatomical parts of the eye (Appendix B). Similar to the study phase of the abdomen, shutter-glasses were used to create the perception of depth (stereopsis). Again, this was a self-rotating reconstruction.

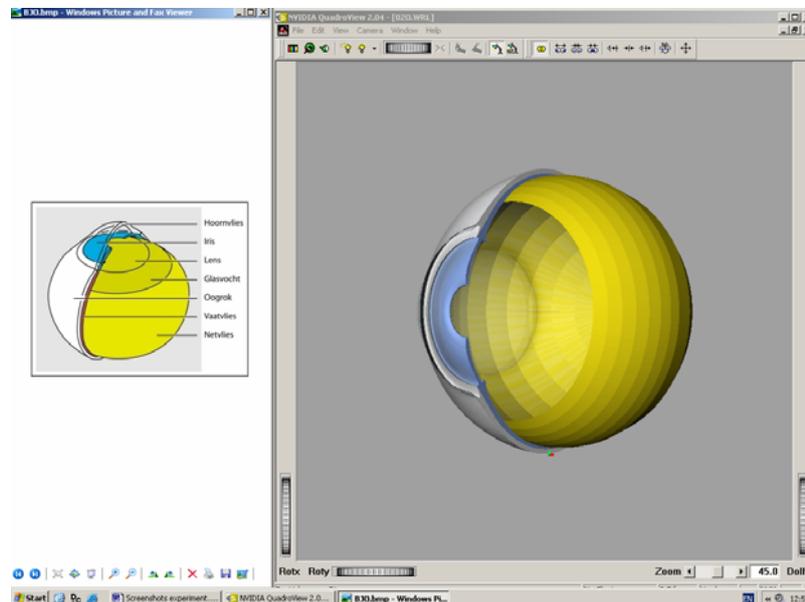


Figure 7. Screenshot of the Eye group study phase.

Participants received three minutes time to form a mental representation of the 3D anatomy and to learn the location of seven anatomical parts of the human eye. To ensure that they would put the same amount of effort in studying the eye as they did in studying the abdomen, participants were made to believe that they would later be tested about their obtained knowledge of the anatomy of the eye. After the three minutes, the reconstruction automatically stopped, and participants were instructed to start the second knowledge test.

### 2.6.2 Sphere group study phase

The Sphere group also received additional study material after the knowledge task. They studied a 3D-representation of an imaginary object, a so-called Sphere (Figure 8). It was accompanied by a labeled

reference figure for seven parts of the imaginary object (Appendix C). The names of the seven parts were entirely made up. Again, the participants wore shutter-glasses to perceive depth (stereopsis), and the representation was self-rotating. Participants were given three minutes to study the object and to learn the names and positions of the seven parts of the Sphere and were also made to believe that they would later be tested about their obtained knowledge of the seven parts. After the three minutes, this representation automatically stopped too, and participants were instructed to start the second knowledge test (i.e. the third phase).

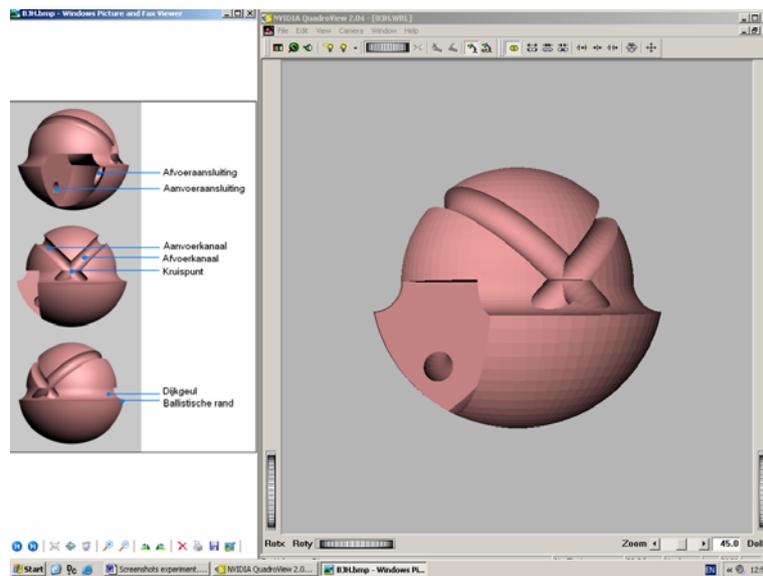


Figure 8. Screenshot of the Sphere group study phase.

### 2.6.3 Second phase for the Control group

The Control group did not receive extra study material. To control for the effect of decay, the participants watched a film clip taken from the National Geographic Documentary 'Free Flight New Zealand' with a duration of three minutes, equal to the study phases of the two experimental conditions. The sound of the film clip was switched off, in order to not offer this extra modality to the participant, since the two other conditions did not contain sounds either. After the three minutes, the video window automatically closed down, and participants were instructed to start the second knowledge test.

### 2.7 Third phase: second knowledge test

The final phase of the experiment was equal again for all three experimental groups. Their knowledge of the human abdomen was assessed again the localization and identification tasks. The stimuli for this second test were different from those used in the first knowledge test phase. Participants were given

three practice trials of both types of task, after which 20 test trials were presented. These consisted of ten localization and ten identification trials, which were randomized over participants to control for order effects. Every trial had a maximum duration of fifteen seconds. Participants were again instructed to focus on getting the answer right, and not on answering as quick as possible. After the participants had answered all 20 test trials they were thanked for their cooperation and the experiment ended. After they had left the cubicle, they were asked if they had used any strategy while answering the items of the knowledge tests (see Appendix D). At the end they were debriefed about the purpose of the study and their questions regarding the study were answered.

## 2.8 Apparatus

All participants used a Pentium 4 computer running Windows XP, a 19" CRT-monitor (Iiyama Vision Master Pro 454) with a refresh rate of 140 Hz, and a PNY-Quadro 4 580XGL videocard. During the different study phases, stereoptic vision was created using Stereographics CrystalEyes CE-3 active shutter-glasses and an E-2 Emitter and StereoEnabler. The refresh rate of 140 Hz of the monitor allowed having an alternating 70 Hz refresh rate for the shutter-glasses. This ensured that participants were not hindered by noticeable flicker of the monitor. The 3D anatomic model of the human abdomen was the same as used by Luursema et al. (2006). It was based a patient suffering from an abdominal aortic aneurysm. The 3D model of the eye was developed by Luursema. The other-domain model for the Sphere group was developed by a fellow student as part of a master thesis (Geerlings, 2006). During the various study phases, the 3D models were self-rotating, and were shown with use of the Nvidia QuadroView 2.04 application. The study phases were set up using AceMacro, and the E-Prime software program, developed by Psychology Software Tools Inc., was used to create the knowledge tasks, and log files for each participant necessary for the analysis of the data.

## 3 Results

The scores of the localization and identification tests were normalized on a scale between 0 and 1. Because accuracy proportions are usually not normally distributed, additional arcsine transformations were conducted on the accuracy proportions to stabilize the variance (Winer, Brown & Michels, 1991). Since the accuracy of one participant differed significantly from the mean accuracy of the control group, this participant was omitted from further analysis.

### 3.1 The Pre-test

The data showed highly significant correlations between the Mental Rotation Task (MRT) and accuracy on the first localization task,  $r = .47, p < .005$ ; and the first identification task,  $r = .30, p < .05$ ; confirming that visuospatial ability (VSA) is a relevant factor in anatomical learning.

Consequently, VSA, as measured by the MRT in this study, was treated as a covariate during further analysis of the data. This analysis also indicates that the localization task correlated better with VSA than the identification task, and therefore in the analyses more emphasis was put on the localization task.

### 3.2 The knowledge tests

As shown in Table 1, accuracy was lower for the first localization task than for the first identification task. Besides, latency was higher for the localization task than for the identification task.

Table 1. Proportion of correct answers and latency (in s) for localization and identification tasks, across all three groups.

	Localization task 1	Identification task 1
Mean Accuracy (SD)	.24 (.15)	.60 (.18)
Mean Latency (SD)	8.8 (2.2)	6.4 (1.6)

A possible latency-accuracy trade-off was ruled out since no significant correlations were found between reaction time and correct-answer proportions on both tasks (localization task:  $r = -.08$ ,  $p = .63$ ; identification task:  $r = -.21$ ,  $p = .16$ ). Analysis of variance (ANOVA) of the latency showed no remarkable results, as was expected with the experimental setup, which judged accuracy more important than latency. Therefore, only analyses of accuracy are reported. In addition, the identification task showed no remarkable effects, and since the localization task also correlated higher with the MRT, the description of the analyses will emphasize those of the localization task.

### 3.3 Group differences

To study the effects of the experimental manipulation, a repeated measures ANCOVA with the accuracy on the first and second task (i.e. block) as a dependent within-subjects factor, group as an independent between-subjects factor and VSA as a covariate was conducted for both the localization task and the identification task. For the localization task, the analysis revealed some significant results. The main effect for group was not significant,  $F(2,43) = 2.6$ ,  $p = .084$ . But the interaction in Figure 9 between group and block proved significant;  $F(2,43) = 5.4$ ,  $p < .05$ .

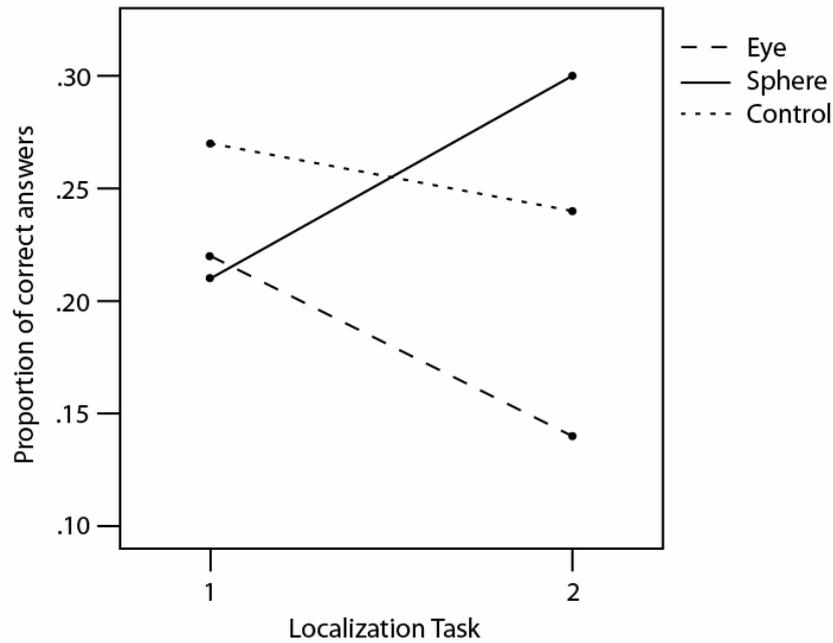


Figure 9. Overview of mean accuracy on the first and second localization tasks for the three groups.

Planned comparisons of the interaction between group and block for pairs of groups revealed significant differences between the Sphere and Eye group:  $F(1,29) = 9.4, p < .005$ , indicating that the Sphere group improved significantly more than the Eye group. The Sphere group also improved significantly more than the Control group:  $F(2,28) = 6.0, p < .05$ ; while there was no significant difference between the Eye and Control group:  $F(1,28) = 0.3, p = .57$ .

For the identification task the repeated measures ANCOVA revealed no significant effects. No significant main effect for group,  $F(2,43) = 3.2, p = .052$ ; and the interaction between group and block also proved not significant;  $F(2,43) = 0.2, p = .82$ . See Figure 10 for an overview.

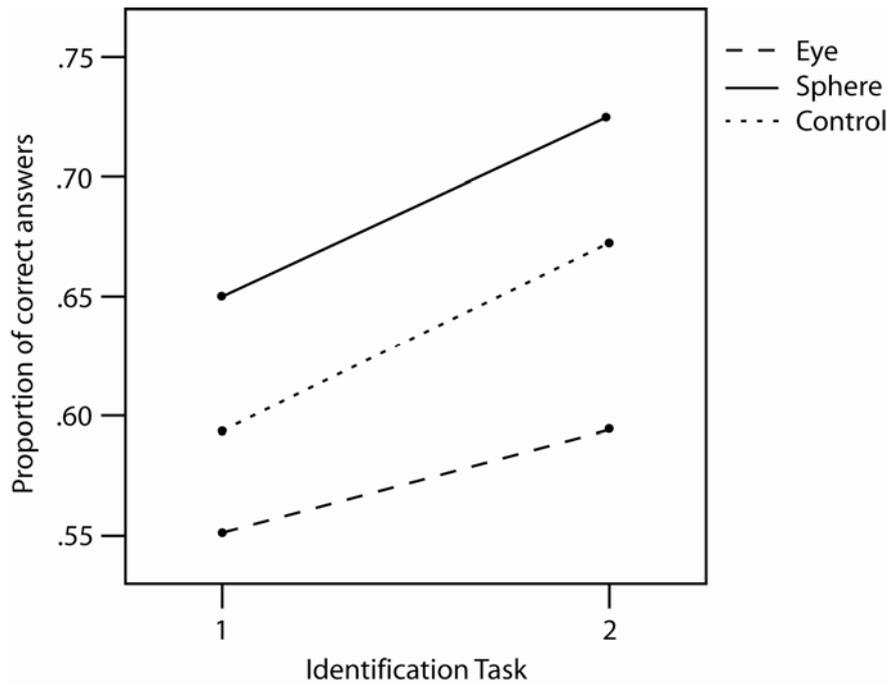


Figure 10. Overview of mean accuracy on the first and second identification tasks for the three groups.

An ANOVA with planned comparisons showed that the differences between the groups on the first identification task were not significant:  $F(2,46) = 1.2, p = .32$ . Histograms of the scores on the identification task revealed a ceiling effect, indicating that the identification task was probably too easy.

### 3.4 Influence of visuospatial ability (VSA)

Accuracy on the localization task was analyzed to see whether individual differences between participants regarding VSA had an influence on their accuracy in that participants of low VSA (as measured by the MRT) were affected more by the experimental treatment than participants of high VSA. On basis of the MRT score, participants in each group were divided into eight high and eight low VSA participants. The low VSA class had a mean MRT score of around .31 (range .13 - .46), while the high VSA class had a Mean MRT score of around .64 (range .46 - .83). A repeated measures ANOVA with block as a dependent variable, and VSA class and experimental group as independent variables did not reveal a significant main effect for VSA class;  $F(2,41) = 1.0, p = .35$ , or any other effects. It seemed that the number of participants used in the study was too small to get a clear view of the influence of VSA on possible retroactive interference effects.

## 4 Discussion

The aim of the present study was to test whether retroactive interference (RI) effects might occur in anatomical learning using a virtual learning environment. This was done in order to develop recommendations for optimally adapting such an environment to the needs and capacities of its users. In the present study, an anatomical study phase was offered to the participants, which was followed by one of three tasks expected to induce different degrees of RI. Participants were tested on their obtained anatomical knowledge using a localization task and an identification task. It was expected that RI would be largest in the Eye group, smallest in the Control group, and the Sphere group would be in between the both others. The influence of visuospatial ability (VSA), as measured by the Mental Rotation Test (MRT), was also taken into account.

### 4.1 Indications for Retroactive Interference

Relative to the first knowledge test phase, accuracy on the second localization task improved for participants in the Sphere group, whereas it decreased for participants in the Eye group and Control group (Figure 9). Accuracy of the Sphere group improved significantly more than accuracy of both the Eye and Control group. There was no difference between the Eye group and the Control group.

The results reveal an unexpected effect. As noted before they do not seem to show RI. On the contrary, compared to the Control group the Sphere study phase seemed to help participants to improve their anatomical knowledge. The effect is remarkable, since there have been reports of memory for visuospatial items being impaired by subsequently offering irrelevant visuospatial material (Toms et al., 1994; Hole, 1996; Quinn & McConnell, 1996).

At first sight, these results seem to lead to the rejection of the hypothesis that RI would be present in the Eye and Sphere group. Yet, a closer look at the data suggests that RI may still have occurred. Perhaps the fact that the human abdomen study phase and the Sphere study phase were similar in terms of presentation mode improved skills that were used in the second knowledge task by the Eye and Sphere groups. Both groups benefited in their second study phase from the extra practice of the mental rotation that was necessary for the knowledge task. But for the Eye group, this positive effect of identical procedure was masked by a negative effect for the similar, but not identical anatomical content. Those skills did not develop to the same degree in the Control group (who only watched a video clip). So, in line with RI, the Eye group performed more poorly than the Sphere group. The relatively poor performance of the Control group on the second knowledge test phase can be attributed to limited experience with the experimental tasks. More research should be conducted to further explain this finding, and to provide implications for virtual learning environments for learning anatomy. The present study did suggest that the necessary mental rotation for learning in the virtual learning environment is facilitated by subsequent studying of an unrelated object which is offered in the same presentation mode as the anatomical structure. Perhaps the necessary time interval between consecutive anatomical structures could be filled with studying these kinds of objects.

#### 4.2 Individual differences

No significant differences between participants of low and high VSA were found. Possibly, the number of participants used in the study was too small to get a clear view of the influence of VSA on possible retroactive interference effects.

#### 4.3 Differences between the localization and identification task

There were some differences regarding the two tasks that together comprised the knowledge task. The localization task proved to be of better use than the identification task, similar to previous use of the tasks (Luursema et al., 2006). One difference was that the localization task correlated better with VSA than the identification task. An explanation could be that the localization task is a task that requires more strict visuospatial processing by the participant, whereas the identification task also has a semantic aspect in naming the various anatomical parts of the abdomen. Graphical representations of the scores on the identification task revealed a ceiling effect, indicating that the identification task was too easy. Besides, participants also stated that the identification task was easier. Also, the identification task did require some mental rotation of the cross-section. Participants had to realize that in the cross-section, the left one of two bilateral organs was actually on the right side, and vice versa. When this was understood, participants made it into a rule for answering the items; left is in fact right, and right is left. This was also stated by participants after the experiment and may have made the identification task a poorer indicator for visuospatial ability.

Hence, the results suggest possible retroactive interference effects when studying anatomical structures in quick succession using a virtual learning environment. More research should be conducted to further explain this finding. The present study did suggest that the necessary mental rotation for learning in the virtual learning environment is facilitated by subsequent studying of an unrelated object which is offered in the same presentation mode as the anatomical structure.

#### 4.4 Suggestions for further research

Some suggestions for further research can be given to extend and further clarify the results of the present study. The first suggestion is to include a different Control condition with a different second study phase. This second phase should contain similar content (i.e. human anatomy) but not a similar procedure (i.e. self-rotating 3D-reconstruction) as the first phase, the human abdomen study phase. In this way, a better insight could be obtained of the effects of similar content and procedure on possible RI effects.

Another suggestion is to isolate the effect of time by varying the time between the various study phases and knowledge tests. In this way more clarity can be obtained when the RI effects are strongest, and after which period of time they are no longer present or negligible.

The results could also be extended by changing the task and its complexity. Altering the complexity of the anatomical reconstructions can be used to discover if the reported findings also apply to more difficult tasks, since in this study only simple study models were used.

The fourth suggestion is to get a better insight in the influence of visuospatial ability. This can be done by using more pretests to determine the level of VSA of the participants. In this study the Mental Rotations Test was used. The MRT mainly loads on the Visualization (VZ) factor of visual perception, as determined by Carroll (1993). Carroll distinguishes five major first-order factors in the domain of visual perception: Visualization (VZ), Spatial Relations (SR), Closure Speed (CS), Flexibility of Closure (CF) and Perceptual Speed (P). A characteristic of tests of the VZ factor is the requirement that the participant apprehends a spatial form, and rotates it in two or three dimensions before matching it with another spatial form (Eliot & Smith, 1983). Also, VZ tests are designed to measure accuracy over latency. To obtain a better picture of the level of VSA of the participants, all five factors could be taken into account. For tests that can be used to measure each factor, see Carroll (1993) or Eliot & Smith (1983). More recent research has successfully identified one additional factor, the Imagery (IM) factor (Burton & Fogarty, 2003). This factor was defined as “the ability in forming internal mental representations of visual patterns and in using such representation in solving spatial problems.” Since the anatomical learning process also requires participants to form these internal representations, the IM factor could also be taken into account. A better view of the influence of VSA could optimize the extent to which the learning environment can be adapted to the needs and capacities of the individual user.

The final suggestion is to check the ecological validity of the study by repeating it with actual medical students and more elaborate anatomical models. Another comparison that could be made in such a study is to check whether such a virtual reality application of anatomical learning facilitates learning better than the traditional way, using anatomical atlases.

#### 4.5 Limitations to the study

Finally, a limitation of this study is that only students from the faculty of behavioral sciences of the University of Twente were used. This reduces the extent to which the obtained results can be generalized.

#### Acknowledgment

I would like to thank my supervisors Jan-Maarten Luursema and Willem Verwey of the department of Cognitive Psychology and Ergonomics (CPE) of the University of Twente, for their help with, and feedback on, the experiment and this thesis.

## References

- Ackerman, J. M. (1998). The visible human project. *Proceedings of the IEEE*, 86, 504–511.
- Anatquest (2004, August 23). Retrieved March 17, 2006, from <http://anatquest.nlm.nih.gov/>
- Baddeley, A. D., & Dale, H. A. (1966). The effect of semantic similarity on retroactive interference in long- and short-term memory. *Journal of Verbal Learning & Verbal Behavior*, 5, 417-421.
- Beech, J. R. (1984). The effects of visual and spatial interference on spatial working memory. *Journal of General Psychology*, 110, 141-149.
- Burton, L. J., & Fogarty, G. J. (2003). The factor structure of visual imagery and spatial abilities. *Intelligence*, 31, 289-318.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, New Jersey: Erlbaum.
- Eliot, J., & Smith, I. M. (1983). *An international directory of spatial tests*. Windsor, England: NFER-Nelson.
- Garg, A. X., Norman, G. R., Spero, L., & Maheshwari, P. (1999). Do virtual computer models hinder anatomy learning? *Academic Medicine*, 74, S87-S89.
- Garg, A. X., Norman, G., & Sperotable, L. (2001). How medical students learn spatial anatomy. *The Lancet*, 357, 363-364.
- Garg A.X., Norman G. R., & Spero L. (2002). Is there any real virtue of virtual reality? *Academic Medicine*, 77, S97-S99.
- Geerlings, H., (2006). *Factors influencing the storage of mental representations of objects in memory*. Master thesis, University of Twente.
- Gevins, A., & Smith, M. E. (2000). Neurophysiological Measures of Working Memory and Individual Differences in Cognitive Ability and Cognitive Style. *Cerebral Cortex*, 10, 829-839.
- Gordon, H. W. (1986). The cognitive laterality battery: tests of specialized cognitive function. *International Journal of Neuroscience*, 29, 223-244.
- Guay, R. & Mc Daniels, E. (1976). *The visualization of viewpoints*. The Purdue Research Foundation. West Lafayette, IN. (as modified by Lippa, Hegarty & Montello, 2002).
- Hole, G. J. (1996). Decay and interference effects in visuospatial short-term memory. *Perception*, 25, 53-64.
- Jastrow, H., & Vollrath, H. (2003). Teaching and learning gross anatomy using modern electronic media based on the Visible Human Project. *Clinical Anatomy*, 16, 44-54.
- Logie, R. H., Zucco, G. M. & Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica*, 75, 55-74.
- Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*, vol. 4. Hillsdale, NJ: Lawrence Erlbaum.
- Luursema, J. M. (2004). Mentale Rotatie Test. University of Twente.

- Luursema, J. M., Verwey, W. B., Kommers, P. A. M., Geelkerken, R. H. & Vos, H. J. (2006). Optimizing conditions for computer-assisted anatomical learning. *Interacting with Computers*, 18, 1123-1138.
- Marsh, R. L., Landau, J. D., Hicks, J. L., & Bink, M. L. (1998). On reducing retroactive interference. *American Journal of Psychology*, 111, 175-190.
- Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna R., & Richardson C. (1995). A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance. *Brain and Cognition*, 28, 39-58.
- Quinn, J. G. & McConnell, J. (1996). Irrelevant pictures in working memory. *The Quarterly Journal of Experimental Psychology*, 49A, 200-215.
- Stanney, K. M. (ed.) (2002): *Handbook of virtual environments: design, implementation, and applications*. Lawrence Erlbaum Associates.
- Temkin, B., Acosta, E., Malvankar, A., & Vaidyanath, S. (2006). An interactive three-dimensional virtual body structures system for anatomical training over the internet. *Clinical Anatomy*, 19, 267-274.
- Tendick F., & Cavusoglu M. C. (1997). Human-machine interface for minimally invasive surgery. *Proceedings 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, IL, Nov 1997.
- Toms, M., Morris, N. & Foley, P. (1994). Characteristics of visual interference with visuospatial working memory. *British Journal of Psychology*, 85, 131-144.
- Vandenberg, S. G., & Kuse A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599-601.
- Wanzel, K. R., Hamstra, S. J., Anastakis, D. J., Matsumoto, E. D., & Cusimano, M. D. (2002). Effect of visual-spatial ability on learning of spatially-complex surgical skills. *The lancet*, 359, 230-231.
- Weiss, E., Siedentopf, C. M., Hofer, A., Deisenhammer, E. A., Hoptman, M. J., Kremser, C. et al. (2003). Sex differences in brain activation pattern during a visuospatial cognitive task: A functional magnetic resonance imaging study in healthy volunteers. *Neuroscience Letters*, 344, 169-172.
- Wheatstone, C. (1838) Contributions to the physiology of vision. Part I. On some remarkable, and hitherto unobserved, phenomena of binocular vision, *Royal Society of London Philosophical Transactions*, 128, 371-394.
- Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design*. New York: McGraw-Hill.

## Appendix A: Anatomical Structures of the Abdomen

List of the 11 anatomical structures of the human abdomen that were used in the human abdomen study phase of the experiment. The Dutch terms were used in the experiment, the English translation is given as a service to the non-Dutch reader.

English:	Dutch:
Aorta	Aorta
Aneurysm	Aneurysma
Spinal Column	Wervelkolom
Left half of pelvis	Linker bekkenhelft
Right half of pelvis	Rechter bekkenhelft
Left kidney	Linkernier
Right kidney	Rechternier
Left iliac artery	Linker bekkenslagader
Right iliac artery	Rechter bekkenslagader
Left femoral artery	Linker beenslagader
Right femoral artery	Rechter beenslagader

## Appendix B: Anatomical Structures of the Eye

List of the 7 anatomical structures of the human eye that were used in the same-domain study phase of the experiment. The Dutch terms were used in the experiment, the English translation is given as a service to the non-Dutch reader.

English:	Dutch:
Cornea	Hoornvlies
Iris	Iris
Lens	Lens
Vitreous humor	Glasvocht
Sclera	Oogrok
Choroid	Vaatvlies
Retina	Netvlies

### Appendix C: Structures of the Imaginary Object

List of the 7 structures of the imaginary object that were used in the other-domain study phase of the experiment. The Dutch terms were used in the experiment, the English translation is given as a service to the non-Dutch reader.

English:

Dutch:

Drainconnection

Afvoeraansluiting

Supplyconnection

Aanvoeraansluiting

Drainchannel

Afvoerkanaal

Supplychannel

Aanvoerkanaal

Crossing

Kruispunt

Diketrench

Dijkgeul

Ballistic edge

Ballistische rand

Appendix D: Questions asked to the participants after the experiment.

1. What did you think of the experiment?
2. Was it hard to answer the items of the knowledge tasks?
3. What task did you like more, the localization task or the identification task, and why?
4. Did you use a strategy to answer the items of the knowledge tasks?
5. Do you have any questions regarding the study?