Does being smart matter? Cognitive abilities and training of simulated minimally invasive surgery tasks

Master Thesis Psychology

University of Twente

Isabel Sippel   s0196371

Supervisors

Dr. Pascal Wilhelm

Department of Instructional Technology

Prof. Dr. Jan Maarten Schraagen

Department of Cognitive Psychology and Ergonomics

11. 01. 2013

Table of Contents

Abstract

Minimally invasive surgery (MIS) has benefits for patients, but places additional burdens on the surgeons, because they have to work with reduced tactile feedback and loss of 3D vision. High cognitive ability might aid surgeons in learning MIS, insofar that surgeons with high cognitive ability acquire MIS skills more quickly and make fewer errors, which is why this study investigated whether this is the case and if so, which cognitive abilities are particularly helpful. The role of experience with MIS was also explored. To do this, 34 participants completed several cognitive ability tests and trained an endovascular procedure on a simulator. The results showed that cognitive abilities alone could not predict learning speed or errors, but an interaction between experience and cognitive abilities was visible. While experts with high and low abilities did not differ from each other, novices with high abilities learned significantly faster and made fewer errors than novices with low abilities. This shows that cognitive abilities might be especially important for the initial phase of learning.

Introduction

Since the introduction of minimally invasive surgery (MIS), many open procedures have been replaced by MIS because of the benefits for patients. Overall, the minimal incisions reduce post-operative pain and lead to earlier mobilization of patients and therefore shorter hospital stays (Van Empel et al., 2012) However, MIS is challenging for the surgeons performing the operation, because they have to work with reduced tactile feedback and a loss of 3-dimensional (3D) vision, which happens because the 3D real world patient is represented on a 2D screen (Ahlberg et al., 2007). Furthermore, surgeons have to deal with changing camera angles, which are not necessarily aligned with their sight, and with the fulcrum effect, which means that a surgeon's movement to the right will cause an actual surgical instrument to move to the left and vice versa (Luursema, 2010).

It has been proposed that some surgeons are able to learn MIS faster than others because they have higher cognitive abilities (Luursema, 2012). If this is indeed the case, cognitive abilities can be used to assess training needs, so that even early in medical education, individual training programs can be developed. Patient safety will be enhanced, because surgeons have been trained according to their needs. The research question of this study is therefore to explore whether *medical trainees with high cognitive abilities are able to learn MIS faster, and if so, which cognitive abilities best predict training efficiency*. A second question is *which role experience with MIS plays in the learning process.*

To do this, it is important to understand how (surgical) skills are acquired and practiced. Fitts and Posner (1967) proposed a three - stage theory of skill acquisition. First, in the cognitive phase, people execute the necessary action in steps and observe closely what they are doing. In the second stage, the integrative or associative phase, they integrate performance feedback and convert this into action, so that the execution of the task at hand gradually becomes more fluent. Finally, in the autonomous phase, the action becomes automated and can be carried out without much awareness. Ackerman (1992) suggests that in the different phases of skill acquisition, different cognitive abilities play a role. The cognitive phase demands general abilities such as reasoning, spatial ability and perceptual speed, the associative phase calls upon visuo – spatial ability, and the automation phase requires psychomotor abilities. The strength of the correlation between these cognitive abilities and the execution of the action also depends on the task complexity and on the consistency of training, the latter meaning that the correlation is higher when people train regularly and repeatedly. The three – stages theory of skill acquisition states that only consistent tasks can become fully automated (Anderson, 1982). In inconsistent tasks, the demand on cognitive and

attentional resources stays high and the task does not become easier over time (Shiffrin & Schneider, 1977). This is the case with surgeries, because each operation is different and remains unpredictable. Some surgical skills can therefore become automated, but the combination and accurate execution of these skills will always require attention, which is why general abilities and visuo – spatial ability remain important. People cannot fully move to the automated phase and therefore still have to use the cognitive abilities associated with the other phases as well (Luursema, 2010).

To help surgeons practice MIS, simulator trainings have been introduced, in which surgeons can practice new techniques in a safe environment without compromising patients' safety (Lee et al., 2010; Schreuder, 2010). Simulators allow for errors and mistakes, which can be useful learning points (Lewis, Aggarwal, Rajaretnam, Grantcharov & Darzi, 2011). In addition to this, simulators provide opportunities for feedback and assessment, as well as a standardized experience for trainees (Palter & Grantcharov, 2010). The most important advantages are that a surgeon's performance can be measured objectively and that technical skills can be evaluated in a transparent way (Gallagher & Traynor, 2008).

Research shows that training success on these simulators can be predicted by cognitive aptitude. Satava, Gallagher and Pellegrini (2003) estimated that a typical surgical procedure calls upon- approximately- 75 percent cognitive abilities and 25 percent technical skill. They observed that this is because surgery tasks cannot become fully automated and always call for attention, as also stated above. Therefore, it seems evident that measuring these abilities could aid in the selection of surgical trainees (Francis, Hanna, Cresswell, Carter & Cuschieri, 2001) and could help to assess the amount of training a surgeon needs (Ritter, McClusky, Gallagher, Enochsson & Smith, 2006).

There are several important cognitive abilities relevant to learning MIS, which will be discussed here. Firstly, visuo-spatial ability refers to a set of cognitive functions related to processing and representing spatial information. This includes visualization, spatial orientation and relations, visual working memory, mental rotations and perceptual speed (Keehner et al., 2004; Luursema, 2010). Several studies have shown the importance of visuo-spatial ability for surgery and surgical skill acquisition. Hedman, Klingberg, Enochsson, Kjellin and Felländer-Tsai (2007) observed that visual working memory span correlated with the total time needed to complete a procedure and the efficiency with which a person does so. In addition, visuo-spatial ability scores are positively related to performance in learning and completing surgical procedures, and the correlation is higher when the complexity of the procedure is higher (Hedman et al., 2006). Furthermore, it seems that spatial skills remain

important throughout all phases of learning, because surgery cannot become fully automated, as pointed out before. In tasks with inconsistent processing demands, such as MIS, the correlation between performance and reasoning ability remains stable, and the correlation between performance and spatial ability even increases during the training (Keehner, Lippa, Montello, Tendick & Hegarty, 2006). This means that as surgeons get more advanced and experienced in carrying out surgery tasks during their training, differences in performance become more evident between people with high and low spatial ability. Visuo-spatial ability is therefore important for initial and final performance of training sessions, and strongly predicts the duration of training that is needed to reach proficiency (Ritter, McClusky, Gallagher, Enochsson & Smith, 2006). Luursema (2011) observed that participants with high visualization ability improved faster on time-on-task measures, which means that they needed fewer rehearsals to complete the surgical task faster than at the initial training. Another study (Keehner, Lippa, Montello, Tendick & Hegarty, 2006) also showed that participants with high spatial ability reached proficiency faster. After each participant had reached proficiency, high spatials were still somewhat faster at completing the surgical tasks. However, it should be observed that all studies showing effects of visuo-spatial ability on skill learning are restricted to training situations and have not investigated the effects of visuo-spatial ability on surgical performance with practicing surgeons in a real world setting.

A second set of relevant cognitive abilities are psychomotor abilities, which consist of bi-manual dexterity, depth perception, and the ability to cope with reduced tactile feedback and the fulcrum effect (Schreuder, 2010). It was found in several studies that these abilities correlate with surgical performance as well. Fine motor dexterity predicts performance mostly at initial training sessions, but this was still the case at final sessions, even though the correlation was weaker (e.g. Van Herzeele et al., 2010). Furthermore, Gallagher, Leonard and Traynor (2009) found that psychomotor ability, assessed via depth perception, correlates with operative skill and the ability to learn on a simulator, because participants with lower psychomotor abilities found it more difficult to reach a pre-set level of proficiency.

Overall, it seems that the above mentioned cognitive abilities predict surgical performance during training sessions and that those with less aptitude might require more training to reach proficiency, which is why the first hypothesis is:

H1: Participants with high cognitive aptitude perform better on the simulator and reach proficiency faster than those with less aptitude.

However, do skills learned from simulator training transfer to real operation theatres or do people trained on a simulator only get better in performing simulator tasks (Van Herzeele et al., 2008)? Several studies show that simulator training is beneficial for real world surgery. For example, Kundhal and Grantcharov (2009) found that on the simulator, significant skills can be rehearsed, and that simulator training does have predictive validity for real world surgery. In line with this, Lee et al. (2010) also observed that subjects receiving simulation-based training demonstrated superior performance of complex surgical skills. Moreover, Van Herzeele et al. (2008) observed that experienced surgeons also benefit from simulator training. In their study, expert endovascular surgeons received a simulator training course, after which they showed shorter real surgery time and fewer errors. They also felt more competent to conduct the procedure. Also, group consistency was higher after the course; they all performed the task about as fast and as safe. Thus, there is evidence that skills acquired in a simulator are indeed transferable to reality and lead to reduction of errors in the operation theatre (Crochet et al., 2011) and an improvement in overall performance (Lewis, Aggarwal, Rajaretnam, Grantcharov & Darzi, 2011).

Still, the relationship is not that straightforward, because in addition to the benefit of simulator training and the influence of cognitive abilities, the effects of previous experience or expertise also have to be taken into account. Van Herzeele, et al. (2010) found that trainees show significantly better surgical performance after training than non-trainees, and that participants with previous experience with MIS performed better than participants without experience, thereby making evident that experience is an important factor in predicting MIS performance.

Research shows that training periods eliminate differences in performance for novice groups (Hedman et al., 2006). If participants have no experience with simulator training, they are highly inconsistent in performance at the initial training session, suggesting that differences in cognitive abilities play a role here. However, they get highly consistent after some training sessions (Hedman, Klingberg, Enochsson, Kjellin & Felländer-Tsai, 2007). It could thus be said that with higher levels of experience, the importance of cognitive abilities diminishes, as was also discussed in the section about automation of skills. However, this is only partly true. Keehner, Lippa, Montello, Tendick and Hegarty (2006) showed that surgical task content is equally important. If a task is highly spatial, the correlation remains stable even after practice, therefore suggesting that content based abilities, meaning cognitive abilities directly related to the content of the surgery, stay important also after proficiency is reached. However, Maagaard et al. (2011) showed that this skill retention only holds for a limited

period of time. For novices, acquired skills were the same after a six months period without training, but after 18 months, they were back to the pre-training level, which shows that regular training is necessary to maintain adequate performance levels. Furthermore, Gallagher, Cowie, Crothers, Jordan-Black and Satava (2003) showed that a relationship between laparoscopic surgery performance and perceptual ability exists for novices as well as experienced surgeons. Using the PicSOr test (Pictorial Surface Orientation; Gallagher, Cowie, Crothers, Jordan-Black & Satava, 2003), they found that perceptual skills associated with the recall of information about depth from pictures are directly relevant to differences in laparoscopic performance in both novice and expert groups. Still, expert surgeons do perform better on simulators (Schreuder, 2010) and reach proficiency faster than novices (Grantcharov, Bardram, Funch-Jensen & Rosenberg, 2003), even when correcting for spatial ability (Keehner, Lippa, Montello, Tendick & Hegarty, 2006). This is to be expected, because master surgeons have higher levels of eye-hand coordination and manual dexterity. They are also better at coping with stress, compared with the reference population (Francis, Hanna, Cresswell, Carter & Cuschieri, 2001). These are all important skills needed for the completion of simulator tasks.

All in all, there seems to be an interaction between cognitive skills and surgical experience, but the nature of this interaction is not clear yet. Research in aviation shows that cognitive abilities and expertise both influence the initial level of performance at flight simulator tasks (Yesavage et al., 2011). Wickens, Barnett, Stokes, Davis and Hyman (1989) proposed that the cognitive skills that predict better performance for experts would be different from the cognitive skills that predict better performance for novices. Indeed, they found that novice performance was influenced greatly by spatial abilities and working memory capacity, while expert performance could not be predicted by any of the tests used in their study. In line with this, Ericsson and Smith (1999) suggest that experts rely on different abilities than novices to guide their performance, and mostly use pattern − recognition mechanisms and memory of earlier conducted similar tasks.

Cognitive abilities and expertise thus seems to interact in predicting performance on a simulator. Therefore, the second hypothesis which will be explored, is:

H2: Expertise and cognitive abilities interact such that novices with higher cognitive abilities will perform better on a simulator task and will reach proficiency faster than novices with low cognitive abilities. Experienced participants will perform better overall than novices and

while a difference between experienced participants with high and low cognitive abilities is significant, the effect size will be smaller.

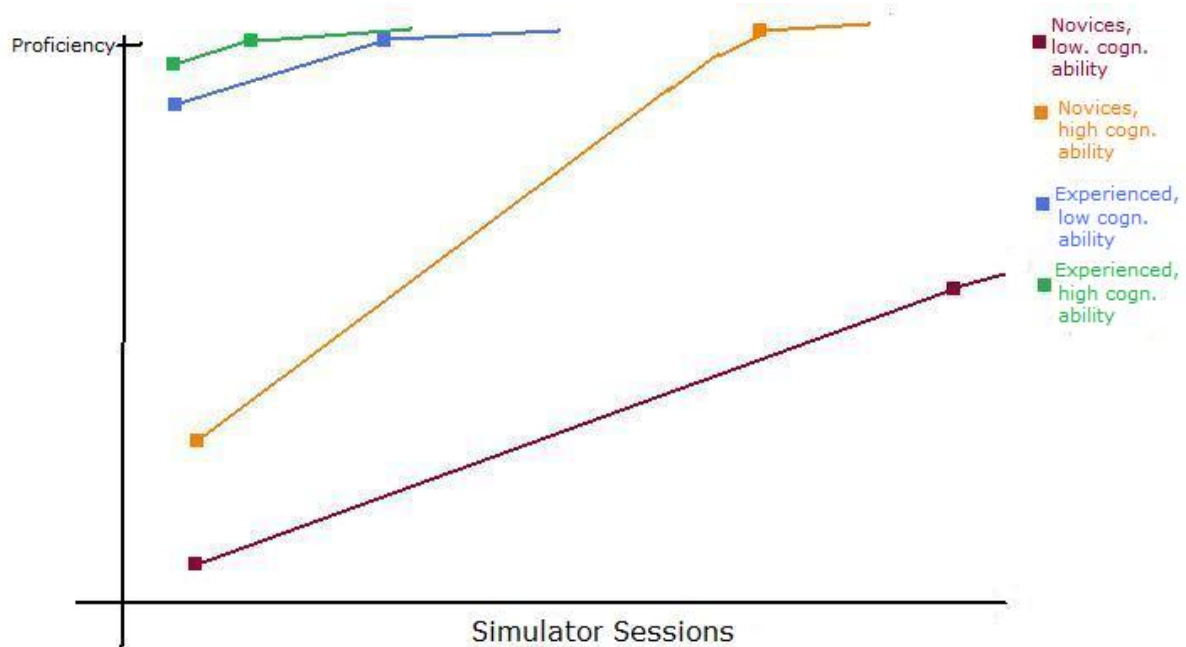A schematic rendition of these relationships can be found in Figure 1.

Figure 1 *Schematic rendition of Hypothesis 2*

Methods

Participants

In this study, 34 participants took part. Of these, 13 were technical medicine students, 7 were medical students, and 14 were medical PhD students. A total of 17 of these participants did not have experience with MIS procedures, while 17 did. Their experience ranged from having performed two MIS procedures to 150 procedures, with a mean of 15.5 procedures and a standard deviation of 34.29. In comparison to expert surgeons, it can be seen that these participants are still beginners, because a surgeon with 10 years of work experience has conducted about 2500 MIS procedures. In the following, participants without any experience in performing MIS procedures will be referred to as novices, whereas all participants who have any experience with MIS, regardless of how many procedures they have conducted, will be labelled experienced participants. In total, the participants were between 20 and 39 years old with a mean age of 27.6 years and a standard deviation of 6.04. Sixteen of the participants were male, four were left – handed and 19 had impaired eyesight corrected with lenses or glasses.

Instruments

Demographic variables questionnaire. This questionnaire consisted of 13 questions concerning for example the participants' age, gender, eyesight and experience with MIS (cp. Appendix 1). It took about two minutes to complete and was pen – and – paper administered.

Cognitive aptitude tests. Several cognitive aptitude tests, assessing spatial insight (Rotating Shapes, Mental Rotation, Paper Folding, and PicSOr), reasoning ability (Raven), spatial memory (Corsi) and processing speed (Identical Pictures), were administered. The completion of these tests took about 60 minutes and took place on a computer. For every test, the number of items was the maximum score that could be reached. The number of items a participant answered correctly was the score the participant received. The following tests were used.

1. *Rotating Shapes (Ekstrom, French, Harman, & Dermen, 1976)*
   On each trial, participants compare two figures that are rotated in the frontal plane only. Participants indicate whether the two figures are identical but rotated or different by pressing a key. The test consists of 128 items and the time limit is four seconds per trial.

2. *Mental Rotation Test (Vandenberg & Kuse, 1978)*
   The Mental Rotation Test consists of 96 items. On each trial, participants compare two figures made of cubes that are rotated around the vertical axis. Participants indicate whether the two figures are identical but rotated or different by pressing a key. The time limit is six seconds per trial.

3. *Paper Folding (Ekstrom et al., 1976)*
   For this test, participants have to imagine folding and unfolding pieces of paper that have been punched. Participants indicate which item of five items shows the correct positions of the holes after the paper has been folded, punched, and completely unfolded again. The time limit is 25 seconds per trial and the test consists of 20 items.

4. *PicSOr (Gallagher, Cowie, Crothers, Jordan – Black & Satava, 2003)*

   The PicSOr (Pictorial Surface Orientation) test consists of 35 items and takes about five minutes to complete. Participants view a geometrical figure on which an arrow is superimposed. Participants have to position the arrow in such a manner that it is placed perpendicular on the surface of the figure.

5. *Corsi Block Tapping (Corsi, 1972)*

   The Corsi Block Tapping test consists of 36 items. On each trial, participants are required to repeat the sequences of blocks that light up on the screen by clicking on these blocks with their mouse pointer. The number of blocks that needs to be tapped (i.e. clicked on with the mouse pointer) is restricted from four to nine. There is no time limit for this test.

6. *Raven's Progressive Matrices (Raven, Raven, & Court, 1998)*

   For this test, participants are given a figure (with three rows and columns) with the lower right-hand cell cut out, along with eight possible alternative solutions. Participants choose the solution that correctly completes the figure (across rows and columns). The limit is 60 seconds per trial, with 18 trails in total.

7. *Identical Pictures (Ekstrom et al., 1976)*

   This test consists of 96 items. On each trial, participants have to compare one figure on the left with five figures on the right of the screen and indicate which of the five figures is the same as the figure on the left. The time limit is two minutes for the complete test.

   Simulator Procedure. Case 1 of the renal module of the ANGIO mentor (Simbionix, 2011) was used. This case is a left ostial lesion, whereby participants have to navigate a guidewire to the accurate position in the artery without damaging any vessels, so that a stent can be placed over it. The participants followed a proficiency – based program. This means that they completed as many trials as they needed to accurately perform the procedure. The level they had to reach was based on earlier sampled expert parameter values. To obtain these values, two expert vascular surgeons and four interventional radiologists were asked to complete the procedure twice. The mean value of these different parameters was used to determine the level that the participants had to reach. The parameters used were total time,

fluoroscopy time, total amount of contrast liquid used, total number of roadmaps (pictures of the vessel) and numbers of guidewires and catheters used (cp. Table 1). After each completed trial in the actual experiment, the participants received feedback on how well they had done and to what extent they had come close to the expert values. If they had reached the expert values, the experiment was terminated.

Table 1   *Means of Expert Simulator Parameter Values*

	Total time	Fluoroscopy time	Contrast liquid	Roadmaps	Guidewires	Catheters
Expert value to be reached	2:40 min	1:30 min	4.5 ml	1	2	2

Procedure

The study consisted of two parts.

First, participants signed an informed consent form that explained the study. Participation was completely voluntary and they could stop at any moment without having to specify reasons and without any consequences. After this, the participants filled in a demographic variables questionnaire. They then proceeded to complete the cognitive aptitude tests.

Second, the participants completed simulator sessions, while their performance was measured using the aforementioned simulator parameters. These sessions took place directly after the cognitive ability tests or on a different day. This was up to the participants' preferences. One session took 30 minutes and consisted of as many trials as the participant could complete in that time period. Since there were time limits to the study, the number of simulator sessions was limited. If they could not perform the procedure accurately after five sessions, they were still asked to stop. This value was based on previously conducted pilot – tests, which showed that five was the maximum number of sessions needed.

Scoring and Data - Analyses

To assess the participants' performance, several variables were used: number of trials needed to reach expert values, amount of time needed for each trial, and number of errors made. These were assessed via log file analysis (Simbionix, 2011). A comparison was made between the participants' performance and adequate performance, assessing how much the participants

digressed from the optimal course of action. A distinction was made between severe and non-severe errors. Severe errors were actions that resulted in harm to the patient or that made continuation of the procedure impossible. These were scored with two error points. Non-severe errors were actions that were not optimal and differed from the experts' actions, but did not harm the patient and made continuation of the procedure still possible. These were scored with one error point. Both were taken together to make up a total error score for each participant on each trial.

Scatter plots and descriptive statistics were used to find any outliers in the data. Scores that differed from the other participants by more than two standard deviations were treated as outliers. Also, Pearson correlations were calculated between the cognitive ability test scores to check for possible overlap in abilities measured.

Participants were split into a high cognitive ability and a low cognitive ability group using median split on the overall score on the cognitive ability tests, and into experienced participants and novices, as explained above. Ultimately, four groups were produced: Experienced participants with high or low cognitive abilities and novices with high or low cognitive abilities.

To test the first hypothesis, regression analysis was used to find out if the independent variables could predict the outcome variables. The predictors were cognitive ability scores and the number of MIS procedures the participants had conducted before the experiment. The dependent variables were the errors a participant made, and the number of trials each participant needed to reach the expert values. Errors reflect the participants' performance and the number of trials relates to learning speed.

To test the second hypothesis, a repeated measure ANOVA with one within-subjects factor was conducted. This within-subjects factor was the time a participant needed for each trial to complete the procedure, and had five levels, which were the five trials. The novice/experienced and high/low ability distinctions served as independent variable, so that it could be tested if cognitive abilities and MIS experience could predict the outcome variables and if they interacted. The same was also done with the number of errors participants made. Bonferroni confidence intervals and contrast analyses were then used to analyse group differences and the Kruskal Wallis test was conducted to find out at which trials the groups differed from each other. Effect sizes of the significant differences were also conducted.

Results

Descriptive Statistics

Scatter plots and descriptive statistics showed that one participant scored more than two standard deviations higher than the other participants on four of the cognitive skills tests (Mental Rotation, Corsi Block Tapping, Paper Folding and Rotating Shapes) as well as on two of the performance measures (time and errors). Another participant scored more than two standard deviations lower on three of the cognitive skills tests (Mental Rotation, Paper Folding and PicSOr) and higher on two of the performance measures (time and errors). These participants were therefore excluded from all further analyses, which left the following distribution into the four groups: Nine participants were experienced with high cognitive abilities, six were experienced with low cognitive abilities, ten were novices with high cognitive abilities, and seven were novices with low cognitive abilities.

To present a first overview of the data, descriptive statistics were conducted, which already showed differences in the means between these groups on most of the measures, as can be seen in Table 2.

Table 2   *Differences between Groups on Cognitive Ability Tests and Simulator Values*

	Experienced High ability (N = 9)		Experienced Low Ability (N = 6)		Novices High Ability (N = 10)		Novices Low Ability (N = 7)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Raven	6.57	0.68	7.17	1.22	8.60	1.04	6.86	1.79
Paper Folding	11.97	0.89	11.67	0.83	12.75	0.80	11.09	1.31
Mental Rotation	56.29	5.83	24.67	5.88	55.40	4.51	28.86	7.43
Identical Pictures	67.00	6.13	62.83	4.28	62.50	2.82	56.71	4.15
Rotating Shapes	102.67	7.53	75.33	14.18	99.45	9.43	43.83	24.98
Corsi	1.57	0.20	2.17	0.16	2.20	0.24	1.71	0.18
PicSOr	0.95	0.00	0.85	0.04	0.85	0.04	0.86	0.03
Cogn. Ab. Total	232.54	14.34	184.50	9.65	241.45	6.97	156.23	14.61
Trials	2.00	0.21	2.00	0.79	3.20	0.41	3.29	0.18
Total Time	7.53	1.69	8.31	1.13	13.20	1.63	17.00	3.01
Total Errors	7.50	3.20	9.86	5.54	10.60	2.13	20.14	4.27

T-tests revealed that the differences between the groups on the total score on the cognitive ability tests were significant. The mean differences between experienced participants with high and low ability was 48.04 (t (13) = 3.18, p < 0.05) and between novices with high and low abilities it was 85.22 (t (15) = 5.79, p < 0.001). There were no significant differences between novices and experienced participants with high abilities (t (17) = 0.11, p > 0.05), nor between novices and experienced participants with low abilities (t (11) = 1.55, p > 0.05). For the Rotating Shapes test, the same picture emerged. The mean difference between experienced participants with high and low abilities was 27.33 (t (13) = 4.30, p < 0.05). Between novices with high and low abilities it was 55.62 (t (15) = 6.71, p < 0.001). No significant differences were found between novices and experienced participants with high ability (t (15) = 0.72, p > 0.05), nor between novices and experienced participants with low abilities (t (13) = 2.41, p > 0.05). The group division thus worked as intended.

To get a further overview of the data, correlation analyses were conducted to see if the outcome variables were related to each other, as well as to find out if they correlated with the cognitive ability tests, and if the cognitive ability tests correlated with each other. The significant correlations can be found in Table 3.

Table 3   *Correlations between Cognitive Ability Tests and Outcome Variables*

	Total Time	Errors	Trials	Cogn. Ab. Total	Rotating Shapes	Mental Rotation	MIS Experience
Total Time		0.67**	0.74**	- 0.44*	- 0.42*	- 0.31	- 0.31
Errors	0.67**		0.56*	- 0.27	- 0.25	- 0.27	- 0.28
Trials	0.74**	0.56*		- 0.11	- 0.13	- 0.03	- 0.44*
Cogn. Ab. Total	- 0.44*	- 0.27	- 0.11		0.87**	0.78**	- 0.10
Rotating Shapes	- 0.42*	- 0.25	- 0.13	0.87**		0.47*	- 0.06
Mental Rotation	- 0.31	- 0.27	- 0.03	0.78**	0.47*		- 0.10
MIS experience	- 0.31	- 0.28	- 0.44*	- 0.10	- 0.06	- 0.10	

Note. *p < 0.05 **p < 0.001

Regression analysis

Regression analyses with a stepwise method and the number of trials and the number of errors as dependent variables were conducted to test the first hypothesis. The independent variables were chosen on basis of the correlation analysis, which showed that only the rotating shapes

tests, the total score on the cognitive ability tests and experience with MIS correlated significantly with the outcome variables.

Regression analysis showed that the cognitive ability tests could not significantly predict the number of trials a participant needed to learn the procedure accurately. It could be observed, though, that the experience a participant had with minimally invasive surgery, thus the number of procedures a participant had conducted before the experiment, did predict how many trials they needed on the simulator. This is summarized in Table 4.

The same picture emerged for the total number of errors the participants made, as a dependent variable. Here, cognitive ability also did not significantly predict the outcome variable. What is worth noting here, is that the number of MIS procedures conducted beforehand also failed to predict the number of errors made (t (30) = - 1.39, $p > 0.05$), which confirmed the outcome of the correlation analyses. Therefore, the regression analysis showed that Hypothesis 1 could not be supported by the data.

Table 4 *Regression with Number of Trials as Dependent Variable*

Predictor	B	SE B	β	t	Significance
Rotating Shapes	- 0.01	0.01	- 0.13	- 0.69	0.495
Cogn. Ab. total	- 0.00	0.00	- 0.11	- 0.60	0.555
MIS experience	- 0.02	0.01	- 0.44	- 2.65*	0.013

Note. *p < 0.05

Repeated Measures ANOVAs

To test Hypothesis 2, a repeated measures ANOVA was conducted.

First, the time a participant needed to complete each trial was used as a within-subjects factor and ability (high and low) and expertise (novices and experienced participants) were the between-subjects factors. This revealed that while the cognitive ability tests alone could not significantly predict the time for each trial, an interaction between cognitive ability and experience with MIS procedures did occur, as can be seen in Table 5.

Table 5 *Repeated Measures with Time for Each Trial as Within-Subject Factor*

Between-Subjects Factors	F value	Significance
Cogn. Ab. Total	2.60	0.118
MIS experience	14.47*	0.001
Cogn. Ab. Total*MIS	5.80*	0.003

Note. *p < 0.05

Bonferroni confidence intervals further revealed where the differences between the groups originated. First, it could be observed that experienced participants with high cognitive abilities did not score significantly different on the time needed to complete each trial compared to experienced participants with low cognitive abilities (mean difference = 0.34, $p > 0.05$). Furthermore, novices with high cognitive abilities also did not score differently from both experienced groups (mean difference = 1.17, $p > 0.05$ and 0.83, $p > 0.05$, respectively), but novices with low cognitive abilities did need significantly more time to complete the trials than the other three groups. The mean difference with the experienced participants with high aptitude was 2.05 ($p < 0.05$), with the experienced participants with low aptitude it was 1.71 ($p < 0.05$) and with the novices with high aptitude it was 1.19 ($p < 0.05$). However, the Kruskal Wallis Test showed that all of these differences are only significant at the first three trials (H (3) = 13.67, $p < 0.05$; H (3) = 16.70, $p < 0.05$ and H (3) = 19.56, $p < 0.001$, respectively). The differences were not significant at trial four and five (H (3) = 6.66, $p > 0.05$ and H (3) = 4.54, $p > 0.05$, respectively). This can be seen in Figure 2.

Furthermore, the effect size for the difference between the novice groups was $\omega^2 = 0.006$. The effect size for the difference between the experienced groups was not conducted, because this difference was not significant.
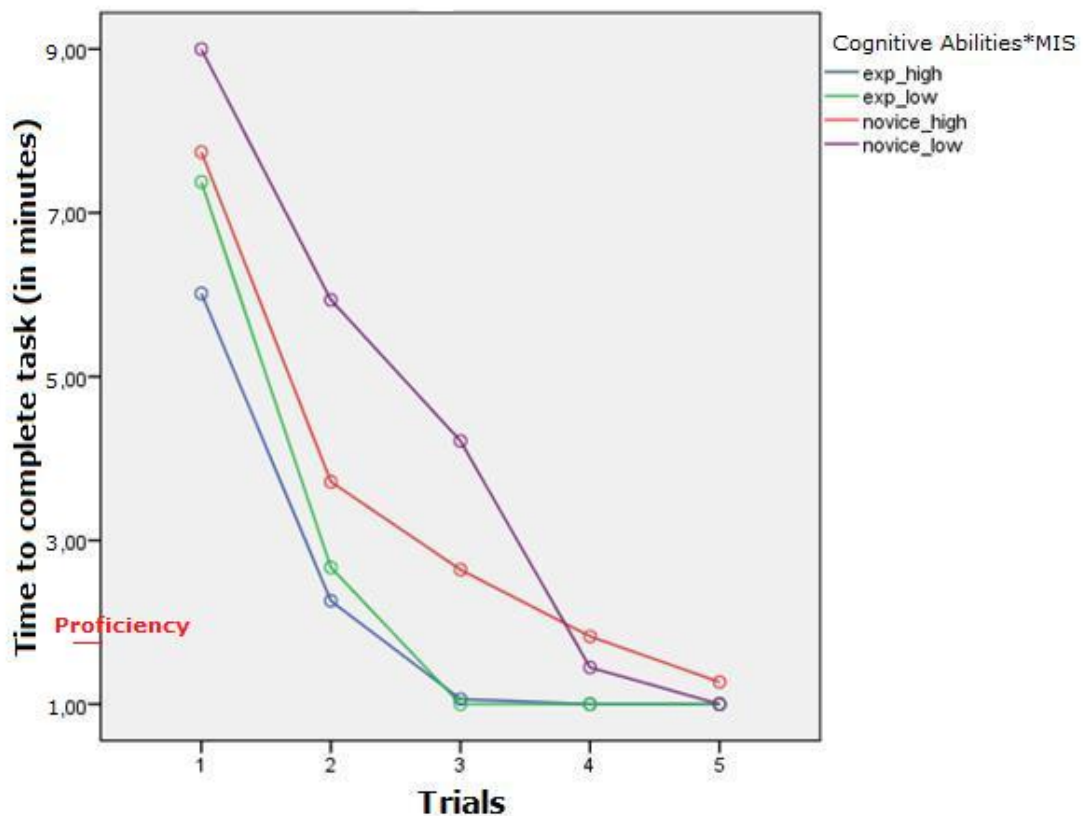

Figure 2 *Group differences in time to complete the task for each successive trial*

Secondly, the number of errors a participant made at each trial was used as a within-subjects factor and the cognitive ability tests and experience with MIS procedures were the between-subjects factors. As can be seen in Table 6, only one of the cognitive skills tests could in interaction with MIS experience significantly predict the number of errors made.

Table 6   *Repeated Measures with Errors at Each Trial as Within-Subjects Factor*

Between-Subjects Factors	F value	Significance
Rotating Shapes	5.76*	0.023
MIS experience	14.47*	0.001
Rotating Shapes*MIS	3.26*	0.036

Note. *p < 0.05

Since only one of the cognitive ability tests had predictive value for the number of errors made, group differences on this test were further analysed. Bonferroni confidence intervals showed how the groups differed from each other regarding the Rotating Shapes test. It could be observed that the experienced participants with a high score on this test had a significantly lower error score than the three other groups. The mean difference with the low scoring experienced group was - 2.56 ($p < 0.05$), with the high scoring novice group it was - 2.10 ($p < 0.05$), and with novices with a low score it was - 3.53 ($p < 0.05$). Furthermore, the two novice groups also significantly differed on how many errors they made at the trials; novices with a high score made fewer errors than novices with a low score (mean difference = 2.16, $p < 0.05$). What could further be observed was that experienced with a low score did not differ significantly from novices with a high score (mean difference = 0.07, $p > 0.05$), nor from novices with a low score (mean difference = 1.50, $p > 0.05$). Again, the Kruskal Wallis test showed that the differences were only significant at the first three trials ($H (3) = 9.11, p < 0.05$, $H (3) = 15.39, p < 0.05$ and $H (3) = 16.34, p < 0.05$, respectively). At trial four and five, the differences were not significant ($H (3) = 6.41, p > 0.05$ and $H (3) = 3.94, p > 0.05$, respectively).

Furthermore, the effect size for the difference between the novice groups was $\omega^2 = 0.003$, while the effect size for the difference between the experienced groups was $\omega^2 = 0.002$. What is worth noting is, that while experienced participants with a low score did not differ overall from novices with a high score, they did make more errors on the first trial (mean difference = 3.86, $p < 0.05$). These results can be seen in Figure 3. All in all, this indicates that Hypothesis 2 is partly supported by the data.
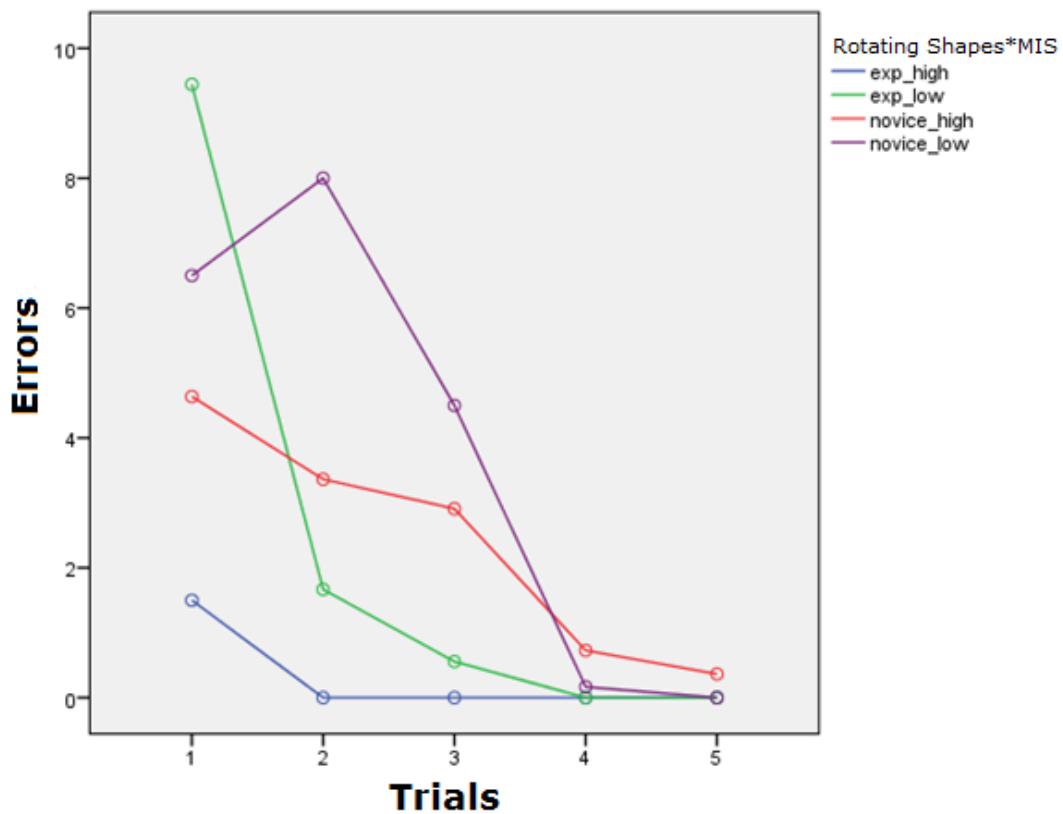
Figure 3 *Group differences for errors*

Conclusions and Discussion

The goal of this study was to investigate the relative influence of cognitive ability and experience on simulated endovascular surgery performance. To do this, an experiment was conducted with experts and novices in different cognitive ability groups. This study tested two hypotheses, which will be discussed in the following, and the results will be put into the context of earlier studies of this type. Furthermore, limitations of this study will be discussed, and recommendations for applied settings and for further research will be given.

The first hypothesis was concerned with the influence of cognitive abilities on simulator performance, and stated that participants with higher abilities would perform better and learn faster than participants with low abilities. This hypothesis was not supported by the data, as it was found that while the cognitive ability tests did correlate with the time needed to complete the trials, cognitive ability could predict neither the number of trials a participant needed to reach proficiency, nor the total number of errors they made. One explanation for this is proposed by Crotchet et al. (2011), who state that it is the training that matters, not the cognitive abilities per se. Good performance does not necessarily depend on innate abilities,

but rather on the nature of the training one receives. In this study, all participants received the same training on the simulator, with feedback on how good they did it, which is why the cognitive abilities might not have been that important. Ritter, McClusky, Gallagher, Enochsson and Smith (2006) support this by saying that similar performance for all groups can be expected after adequate training.

A second explanation is offered by Keehner, Lippa, Montello, Tendick and Hegarty (2006), who found that the task content is very important. The difficulty of the task, as well as spatiality, plays a role in the correlation with cognitive ability. So it might be that the chosen task was too simple to show any correlation with cognitive ability, because it seems that novices could learn it as well as experienced participants, irrespective of ability level.

The second hypothesis concerned the interaction between cognitive abilities and experience with MIS. It was expected that novices with high cognitive abilities would perform better than novices with low abilities, and that for the experienced groups there would also be differences, but these would be smaller. In addition, experienced were expected to perform better than novices overall, regardless of cognitive ability. This was partly supported by the data, and will therefore be discussed in several parts.

Firstly, it was found that novices with high cognitive abilities performed as fast as both experienced groups, meaning that they could complete the procedure in the same time an experienced could. Novices with low cognitive abilities, on the other hand, performed significantly slower. This could indicate that the importance of cognitive abilities diminishes with experience (Keehner et al., 2004), meaning that while novices need to rely on their cognitive abilities to complete a task, and therefore perform better if they have higher abilities, experienced participants do not need to do this, because they can rely on their experience and knowledge of the task. In addition to this, there were no differences in the total time to complete the task found between the two experienced groups, which could have to do with the concept of automation mentioned in the introduction. It was expected that surgery tasks cannot become fully automated because they are inconsistent and every case is different, and therefore cognitive abilities would stay important (Shiffrin & Schneider, 1977). But in this study, the participants practiced the same task several times, and it was a navigating task virtually every endovascular surgery contains, so it might have actually been automated for the experienced participants, and possibly therefore there were no differences between the two experienced groups.

Secondly, because a time measure alone has limitations, which means that a too fast performance might be unsafe (Schreuder, 2010), the number of errors was also investigated. Here, it could be observed that scores on only one of the cognitive tests was related to the errors made at different trials in interaction with MIS experience. Experienced participants with a high score on this particular test, which measured visuo – spatial ability, made significantly fewer errors than participants in all other groups. Besides, novices with high visuo – spatial ability made fewer errors than novices with low ability, while experienced participants with low ability did not differ from both novice groups. It can thus be seen that experience and visuo – spatial ability both matter for the number of errors made. This is in line with earlier research by Hedman et al. (2006), who state that visuo – spatial ability has more impact than general cognitive ability, especially if the task is spatially complex. This possibly explains why the other cognitive tests failed to predict the number of errors, while this test could. The task used in this study was a navigation task, which while it appeared to be relatively simple and common, as discussed above, still called for visuo – spatial ability. Furthermore, Luursema (2010) offers a possible explanation for the fact that experienced participants with low visuo – spatial ability and novices with high ability can perform at the same level. He states that visuo – spatial ability correlates with learning rate, which means that high scorers reach proficiency faster than low scorers, and therefore make fewer mistakes. Now, experience has to be taken into account, and therefore, experienced participants with low ability did not perform worse overall than novices with high ability, but at the same level, because they can compensate for their lack of visuo – spatial ability with their experience. Still, experienced participants with low ability did make more errors than novices with high ability on the first trial. Hoffman, Marx, Amin and McDermott (2010) state that initial trials show how long someone needs to get to know a new system, in this case the simulator. Therefore, it is possible that experienced participants with low ability needed longer to figure out how the simulator works exactly, and therefore made more mistakes than novices with high ability on the first trial. On the following trials, they could then use their experience, as stated above. Furthermore, experienced participants with high ability made fewer errors than all other groups because they have their experience and their highly developed visuo – spatial ability to rely on (Keehner, Lippa, Montello, Tendick & Hegarty, 2006). In addition to this, the effect size of the differences between the two novice groups was bigger than between the two experienced groups, which again shows that experience seems to be more important than cognitive abilities.

However, all these effects could only be observed at the initial trials and diminish after three trials. First, this shows that the measurements did work as intended, but that the task might have been too simple, as discussed above. Second, it again demonstrates that while initially, there are significant differences between the groups, these vanish after training, which links back to the earlier conclusion that training seems to matter more than cognitive abilities per se.

All in all, it thus can be seen that while cognitive abilities alone cannot predict endovascular surgery training performance, they still play a role in interaction with experience. Cognitive ability, assessed by different subtests, seems to be especially important for novices regarding the time needed to complete a trial, and visuo – spatial ability seems so play a role for both novices and experienced participants in the number of errors made. For applied settings, this implies that especially tests for visuo – spatial ability can be used to detect who needs more extensive training than others. It could be seen that experience can compensate for a lack of visuo – spatial ability, which means that low scorers should receive more training to be able to perform as safe as high scorers. Still, these results have to be regarded with care, since there are several limitations to this study.

It is important to mention, that only a small number of participants took part. The results indicate that there are differences between the groups, but for the most part these were small and could be unreliable. It is therefore recommended to use a bigger sample size in following studies. What could also be interesting to investigate is how experts with even more experience behave. In this study, the most MIS procedures completed beforehand were 150, and it could be seen that experienced participants with low abilities did not make fewer errors than novices with high abilities. So it could be investigated if more experience matters even more, to clarify the role of cognitive abilities and expertise even further.

Furthermore, as mentioned above, the nature of the task could have played a role. In this study, a navigation task was used, which was chosen for its spatial complexity, and which was presumed to have a relation with all the cognitive abilities investigated in this study. However, it ultimately only seemed to call for visuo – spatial ability and appeared to be relatively common and seemingly easy to learn. That is why it is recommended to use a more complex and maybe also longer task. In following studies, for example not only navigating the guidewire to the correct place, but also placing a stent in a safe and accurate position could be done. This would then involve not only spatial relations, but also making decisions about which stent to use and how to get it to the right place, which should correlate with

reasoning ability, and deciding this relatively fast to not harm the patient, which should relate to processing speed. These tasks could be used to show if differences in cognitive ability play a more pronounced role then.

Finally, it can be said that this study shows promising starting points for the future, but that there is still work to be done to fully determine the role of cognitive abilities for endovascular surgery.

Acknowledgements

References

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: dynamics of ability determinants. *Journal of Applied Psychology, 77 (5),* 598-614.

Ahlberg, G., Enochsson, L., Gallagher, A. G., Hedman, L., Hogman, C., McClusky, D. A., Ramel, S., Smith, C. D., & Arvidsson, D. (2007). Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *The American Journal of Surgery, 193,* 797-804.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369-406.

Crochet, P., Aggarwal, J., Singh, S., Ziprin, P., Rajaretnam, N., Grantcharov, T. P., Ericsson, K. A., & Darzi, A. (2011). Deliberate practice on a virtual reality laparoscopic simulator enhances the quality of surgical technical skills. *Annals of Surgery, 253 (6),* 1216-1222.

Ericcson, K. A., & Smith, J. (1999). Prospects and limits of the empirical study of expertise: an introduction. In K. A. Ericsson & J. Smith (eds.), *Towards a general theory of expertise: prospects and limits* (pp. 1-36). New York, NY: Cambridge University Press.

Fitts, P. M., & Posner, M. I. (1967). *Human performance.* Belmont, CA: Brooks/Cole.

Francis, N. K., Hanna, G. B., Cresswell, A. B., Carter, F. J., & Cuschieri, A. (2001). The performance of master surgeons on standard aptitude testing. *The American Journal of Surgery, 182,* 30-33.

Gallagher, A. G., Cowie, R., Crothers, I., Jordan – Black, J. – A., & Satava, R. M. (2003). PicSOr: An objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopic performance. *Surgical Endoscopy, 17,* 1468-1471.

Gallagher, A. G., Leonard, G., & Traynor, O. J. (2009). Role and feasibility of psychomotor and dexterity testing in selection for surgical training. *ANZ Journal of Surgery, 79,* 108-113.

Gallagher, A. G., & Traynor, O. J. (2008). Simulation in surgery: opportunity or thread?. *Irish Journal of Medical Science, 177,* 283-287.

Grantcharov, T. P., Bardram, L., Funch-Jensen, P., & Rosenberg, J. (2003). Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *The American Journal of Surgery, 185,* 146-149.

Hedman, L., Klingberg, T., Enochsson, L., Kjellin, A., & Felländer-Tsai, L. (2007). Visual working memory influences the performance in virtual image–guided surgical intervention. *Surgical Endoscopy,21,* 2044-2050.

Hedman, L., Ström, P., Andersson, P., Kjellin, A., Wredmark, T., & Felländer-Tsai, L. (2006). High-level visual-spatial ability for novices correlates with performance in a visual-spatial complex surgical simulator task. *Surgical Endoscopy, 20,* 1275-1280.

Hoffman, R. R., Marx, M., Amin, R., 5 McDermott, P. L. (2010). Measurement for evaluating the learnability and resilience of methods of cognitive work. *Theoretical Issues in Ergonomics Science, 11 (6),* 561-575.

Keehner, M. M., Lippa, Y., Montello, D. R., Tendick, F., & Hegarty, M. (2006). Learning a spatial skill for surgery: how the contributions of abilities change with practice. *Applied Cognitive Psychology, 20,* 487-503.

Keehner, M. M., Tendick, F., Meng, M. V., Anwar, H. P., Hegarty, M., Stoller, M. L., Duh, Q. (2004). Spatial ability, experience, and skill in laparoscopic surgery. *The American Journal of Surgery, 188,* 71-75.

Kundhal, P. S., & Grantcharov, T. P. (2009). Psychomotor performance measured in a virtual environment correlates with technical skills in the operating room. *Surgical Endoscopy, 23,* 645-649.

Lee, T. J., Qiu, M., Teshome, M., Raghavan, S. S., Tedesco, M. M., & Dalman, R. L. (2010). The utility of endovascular simulation to improve technical performance and stimulate continued interest of preclinical medical students in vascular surgery. *Journal of Surgical Education, 66 (6),* 367-373.

Lewis, T. M., Aggarwal, R., Rajaretnam, N., Grantcharov, T. P., & Darzi, A. (2011). Training in surgical oncology - The role of VR simulation. *Surgical Oncology, 20,* 134-139.

Luursema, J. M. (2010). *See me, touch me, heal me – the role of visuo-spatial ability in virtual anatomical learning and endoscopic simulator training.* Tilburg: Medialaan BV.

Maagaard, M., Sorensen, J. L., Oestergaard, J., Dalsgaard, T., Grantcharov, T. P., Ottesen, B. S., & Larsen, C. R. (2011). Retention of laparoscopic procedural skills acquired on a virtual-reality surgical trainer. *Surgical Endoscopy, 25,* 722-727.

Palter, V. N., & Grantcharov, T. P. (2010). Simulation in surgical education. *Canadian Medical Association Journal, 182 (11),* 1191-1196.

Ritter, E. M., McClusky, D. A., Gallagher, A. G., Enochsson, L., & Smith, C. D. (2006). Perceptual, visuospatial, and psychomotor abilities correlate with duration of training required on a virtual-reality flexible endoscopy simulator. *The American Journal of Surgery, 192,* 379-384.

Satava, R. M., Gallagher, A. G., & Pellegrini, C. A. (2003). Surgical competence and surgical proficiency: definitions, taxonomy, and metrics. *Journal of the American College of Surgeons, 196 (6),* 933-937.

Schreuder, H. (2010). *Advances in training for laparoscopic and robotic surgery.* Enschede: Ipskamp Drukkers.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84 (2),* 127-190.

Van Empel, P. J., van der Veer, W. M., van Rijssen, L. B., Cuesta, A. M., Scheele, F., Bonjer. J., Meijerink, W. J. (2012). Mapping the maze of minimally invasive surgery simulators. *Journal of Laparoendoscopic & advanced surgical Techniques, 22 (1),* 51-60.

Van Heerzele, I., Aggarwal, R., Neequaye, S., Hamady, M., Cleveland, T., Darzi, A., Cheshire, N. J. W., & Gaines, P. (2008). Experienced endovascular interventionalists objectively improve their skills by attending carotid artery stent training courses. *European Journal of Vascular and Endovascular Surgery, 35,* 541-550.

Van Herzeele, I., O'Donoghue, K. G. L., Aggarwal, R., Vermassen, F., Darzi, A., & Cheshire, N. J. W. (2010). Visuospatial and psychomotor aptitude predicts endovascular performance of inexperienced individuals on a virtual reality simulator. *Journal of Vascular Surgery, 51,* 1035-1042.

Wickens, C. D., Barnett, B., Stokes, A., Davis, T., & Hyman, F. (1989). Expertise, stress and pilot judgement. *Human Behaviour in High Stress Situations in Aerospace operations: AGARD Conference Proceedings, 458,* 89-97.

Yesavage, J.A., Jo, B., Adamson, M.M., Kennedy, Q., Noda, A., Hernandez, B., Zeitzer, J.M., Friedman, L.F., Fairchild, K., Scanlon, B.K., Murphy, G.M. Jr., & Taylor, J.L. (2011). Initial cognitive performance predicts longitudinal aviator performance. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 66(4),* 444-453

Appendix

Appendix 1: Demographic variables questionnaire

Welkom!

Dit is het eerste onderdeel van het onderzoek Validatie Cognitieve Vaardigheden Test - Endovasculair.

Dit onderdeel bestaat uit een korte demografische vragenlijst.

Als je vragen hebt over de vragenlijst tijdens het invullen, vraag dan de onderzoeker om hulp.

1. **Wat is je geslacht?**
   Man
   Vrouw

2. **Wat is je geboortedatum?**

3. **Wat is je nationaliteit?**

4. **Geef aan tot welke groep je behoord**
   Technische Geneeskunde Studenten
   Geneeskunde Studenten
   Basisartsen (in opleiding)
   AIOS vaatchirurgie
   AIOS (andere specialisme)
   Vaatchirurgen (niet in opleiding)
   Anders, nml

5. **Ben je (voornamelijk) links- of rechtshandig?**
   Links
   Rechts

6. **Ben je slechtziend?**
   Nee
   Ja, ik draag een bril
   Ja, ik draag contactlenzen
   Ja, maar ik draag geen bril of lenzen
   Niet bekend
   Anders, nml

7. Ben je kleurenblind?

Ja

Nee

Weet niet

8. Heb je dyslexie?

(Dyslexie is een aandoening waardoor je niet vloeiend kunt lezen of je vermogen tot het begrijpen van woorden/zinnen verminderd is.)

Ja

Nee

Weet niet

9. Heb je ervaring met het spelen van computerspelletjes (bijv. op de PC of Wii)?

Ja

Nee

10. Zo ja, hoeveel uur per week speel je gemiddeld computerspelletjes?

11. Heb je wel eens eerder een cognitieve vaardigheden test gemaakt (bijv. een IQ test)?

Ja

Nee

Weet niet

12. Heb je ervaring met het assisteren bij of uitvoeren van een minimaal invasieve procedure?

Ja

Nee

13. Zo ja, hoeveel procedures heb je ongeveer uitgevoerd/geassisteerd?

Dankjewel voor het invullen van de vragenlijst.

Vraag de onderzoeker om verdere instructies.

Appendix 2: Photos of simulator design