# Individual Learning Curves in Bronchoscopy:

# Exploring Skill Acquisition with a Low-Fidelity Prototype

Ace Küpper

Thesis for the degree: M.Sc.

University of Twente

Faculty of Behavioral Management and Social Sciences,

Course of Studies: Human Factors and Engineering Psychology

1st supervisor: Dr. Martin Schmettow (CPE)

2nd supervisor: Dr. Marleen Groenier (TG)

Abstract

**Introduction:** Minimally invasive surgery is a standard procedure and endoscopes are standard devices in modern medicine. Nevertheless, handling of these devices led to restricted fields of vision, more demanding hand-eye-coordination and dexterity skills. Therefore, adjusted training and assessment methods became necessary to determine the skill levels of surgeons. The current study joins the paradigm shift from traditional apprenticeship to modern virtual-reality simulator training and assessment. For this purpose a low-fidelity, low-budget prototype for flexible endoscopy (EndoProto) was created to explore bronchoscopic skill acquisition.

**Method:** Twenty-four students executed two basic scope insertion tasks with the EndoProto. First, they trained bronchoscopic skill from an allocentric and then form an egocentric perspective, defined by the visualization through a low-budget endoscopic camera. We made use of a within-subjects design. We focused on the exploration of learning with the estimation of learning curves. No learning curves could be estimated. Instead, explored individual performance and differences between the allocentric and the egocentric tasks.

**Results:** We applied a multi-level mixed-effects model with a binominal distribution, where trials became exchangeable repeated measures. (1) Results showed group effects for wall contacts conditional on ToT. (2) Participants performed better in the allocentric compared to the egocentric perspectives regarding the performance parameter wall contacts and ToT. (3) We found an interaction effect of routes conditional on tasks. Results were based on 95% CI.

**Discussion:** The validation of the EndoProto still needs to be proven in the context of the resemblance spectrum. Exchangeable trials helped to investigate possible causes for the missing learning effects of participants, as for example the influence of the routes. Finally, observations and expert interviews helped to interpret possible causes for the missing learning curves and helped to adjust the EndoProto for possible future investigations.

*Keywords*: MIS; bronchoscopy; VR-simulator; surgical dexterity; MIS prototype; pulmonary specialist; exponential learning curves; multi-level mixed-effects models

Resumé

**Introductie:** Minimaal invasief chirurgie is een standard procedure en endoscopen zijn de bijhoorende apparaten in de moderne geneeskunde. Toch het gebruik van deze apparaten beperkt de velden van visie en nieuwe eisen zoals meer ingewikkeld hand-oog-coordinatie en handvaardigheden war aan de orde. Aangepast methoden van training en evaluatie waren nodig om de niveaus van de vaardigheden van chirurgen te bepalen. Dit onderzoek sluit aan an het wissel van het traditionele naar het moderne training en evaluatie door simulatoren, die gebruik maken van virtuele-realiteiten. Op grond van dit onderwerp werd een goedkoep en laag-nauwkeurig prototype voor de endoscopie ontwikkelt. Dit thesis is gefocusseerd op het vakgebied van de bronchoscopie, een specifieke vorm van endoscopie. Ons doel was het exploreren van bronchoscopisch vaardigheden met hulp van de ontwikkelde prototype.

**Methode:** Vierentwintig studenten heben twee standard taken voor het invoeren van de endoscope met de EndoProto uitgevoerd. Eerst hebben ze handvaardigheden van een allocentrisch perspectie getraind en dan van een egocentrisch perspectie, gekenmekt door een goedkoep endoscopisch camera. Voor de leercurves hebben wij gebruik van een tijd-reeks design gemaakt. Het was niet mogelijk om leercurves te schatten. Dus hebben wij een within-subjects design toegepast, om de prestaties (gemeten door tijd-van-opgave en muur aanrakingen) van proefpersonen te exploreren.

**Resultaten:** Een muliti-level mixed-effects model met binominal distributie was toegepast, met de variable trials als uitwisselbaar herhaalde meeting. (1) Resultaten tonden group effecten voor muur aanrakingen conditionel on ToT aan. (2) In vergelijking tussen de condities, scoorden de proefpersonen beter in de allocentrische conditie. (3) Resultaten lieten een interactive effect voor routes conditioneel on taken zien. Resultate waren gebaseerd op een 95% CI.

**Discussie:** Het EndoProto maakt uitwisselbare herhaalde metingen mogelijk, zodat onafhankelijke vragen konden onderzocht worden, zoals de invlued van de rotues op de taken. De resultaten van de regressie analyse hebben de geobserveerde problemen van de participanten tijdens het traineren kunnen bevestigen. Verklaringen voor het uitblijven van de leerbochten zijn onderwerp van de discussie. Tot slot hebben qualitatief methoden zoals observaties en interviews met longartsen geholpen om gefundeerd redenen voor het aanpassen en verbeteren van de EndoProto voor toekomstige onderzoeken te geven.

*Sleutelwoorden:* Minimaal invasief chirurgie; bronchoscopie; VR-simulator; chirurgisch handvaardigheden; MIS prototype; longartsen, exponentieel leercurves; multi-level mixed-effects modellen

## Table of Contents

# Author Note

## Original Focus of the Study

The original idea for this thesis was the investigation of the effect of dexterity (expected to be trained with the EndoProto) on the learning curve parameter (ToT). The sample of current study should have served as experimental group, conducting the EndoProto training and a VR-simulator session. The sample of a colleague (Westerhof, 2018) should have served as a control group, only conducting the BRONCH Mentor simulator training. We planned to compare learning curves on individual and group levels between the two samples. In this context we aimed for getting deeper insights in possible consequences of the different training tools and the related hand and finger movements. We planned to use a sensor-glove for this purpose. Nonetheless, technical problems prevented us to use the sensor-glove. Furthermore a breakdown of the BRONCH Mentor and problems with re-availability of the system caused an adjustment of the current study, resulting in a solely focus on the EndoProto training. For reasons of transparency and completeness the partly conducted simulator session is described in Appendix A. For further interest in learning curves regarding the BRONCH Mentor training, we refer to a related master thesis (Westerhof, 2018).

**Individual Learning Curves in Bronchoscopy:**

**Exploring Skill Acquisition with a Low-Fidelity Prototype**

In the 1950s a new standard in modern medicine was set. Accordingly, open surgery was replaced through minimally invasive surgery (MIS). This MIS enabled saver and more qualitative ways in diagnosing and treating human diseases (Alvarado & Reichelderfer, 2000; Gmeiner, et al., 2007; Darzi & Munz, 2004). Nowadays, MIS is a standard procedure used around the globe for all main surgery fields (Degani, Choset, Zubiate, Ota, & Zenati, 2008). The bronchoscopy, defined as the examination of the human lung and internal bronchia, is expected to be the most executed procedure in MIS. Already in the year 1998 more than 200,000 bronchoscopies were executed in Germany. (Rogalla, Rückert, Schmidt, Witt, & Meiri, 2001).

MIS is defined by its endoscopic way of accessing the human body. Accordingly, endoscopic procedures require high-mechanical fiber-optic tools, called endoscopes (Degani, Choset, Zubiate, Ota, & Zenati, 2008). As a consequence of using endoscopes in surgery, new challenges arose for surgeons. In open surgery the perspective of a surgeon was focused directly on the field of operation. As a comparison, in MIS the perspective of a surgeon is focused on a monitor, which provides the recorded procedures in real-time. As a result, MIS led to limited fields of vision and adjusted surgical dexterity skills. Medical examinations became more demanding, time intense and error-prone (Tendick & Cavusoglu, 1997). Compared to traditional open surgery, MIS led to higher complication rates (Gallagher, Leonard, & Traynor, 2009).

For reasons of human and financial costs better training methods and assessment criteria for professionals were requested (Basdogan, Sedef, Harders, & Wesarg, 2007; Darzi & Munz, 2004; Gallagher & Smith, 2003; Gallagher, Leonard, & Traynor, 2009). The need for a paradigm shift away from traditional apprenticeship training to a more objective virtual-reality (VR) simulator training became more relevant (Basdogan, Sedef, Harders, & Wesarg, 2007). Research has demonstrated the usefulness of VR simulators to foster and evaluate endoscopic skills, as dexterity, speed or accuracy (Colt, Crawford, & Galbraith, 2001). Even more important for the current study, VR simulators were reported to improve performance in bronchoscopy (Medford, 2008; Ost, et al., 2001).

Much research has been published about the relevance of VR-simulator training and skill acquisition for performance enhancement in MIS, but professional predictive power has not been reached yet (Groenier, Schmettow, Huijser, & Gallagher, n.d.). One of the main problems of VR-simulator training in MIS is the one of high costs (Gallagher & Smith, 2003). On the contrary, test instruments which are less costly often show decreased resemblance of realistic MIS tasks, as for example demonstrated for laparoscopy (Arendt, 2017). Therefore, the main problem lies in the polarity between resemblance and predictive power, which is examined by an adequate framework: The resemblance spectrum (Schmettow & Groenier, 2017). This framework deals with the challenge of meeting the necessary degree of resemblance of low-cost, low-fidelity MIS training to reach predictive power. This MIS training can for example be supplemented through finding valid and reliable test instruments to generate abstractions of costly VR-simulator tasks or underlying dexterity skills. In the current study, we focus on this resemblance spectrum.

**Focus of the Study**

In accordance to the resemblance spectrum, a low-budget low-fidelity prototype for training endoscopic dexterity skills (EndoProto) was developed. With this prototype we face the challenge of (1) finding dexterity tasks which train necessary skills for VR-simulator task improvement and (2) enable the necessary degree of resemblance for (valid and reliable) predictive power. The EndoProto is thus not regarded as an alternative for VR-simulators, but as a pre-stage for training and developing surgical skills necessary for VR-simulator tasks and resulting real MIS. In this way we supplement previous research about box trainers for endoscopic performance training (Buser, Pouw, & Kusters, 2016-2017; Ohuchida, et al., 2009).

For this aim we wanted to focus on the exploration of learning, through the estimation of learning curves. Nevertheless, the estimation of learning curves failed. Thus, we made use of a multi-level mixed-effects model, where trials became exchangeable repeated measures, which enabled us to investigate several underlying aspects of learning – different performance parameters separately. Therefore, we focused on performance differences to investigate possible reasons for missing learning curves. According to this we focussed (1) on individual differences and (2) on differences between the developed basic scope insertion tasks. Our basic research setup is visible in Figure 1.

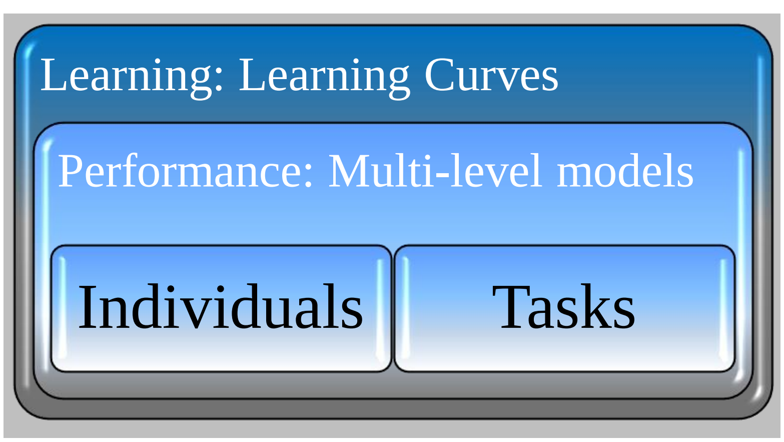



*Figure 1*. Research setup .

**Design requirements**. Based on the previous introduction we face several requirements which our prototype should fulfill. According to the mentioned problem statements we request the following design requirements, ranged in relevance: (1) Approaching the degree of resemblance (scientific predictive power), (2) Supplement VR-simulator training and assessment, (3) Foster learning and performance, (4) Define skill acquisition to surgical dexterity (visual and spatial), (5) economics (using small resources (a low-cost and low-fidelity construction) to reach the highest possible effect (learning and performance increase)).

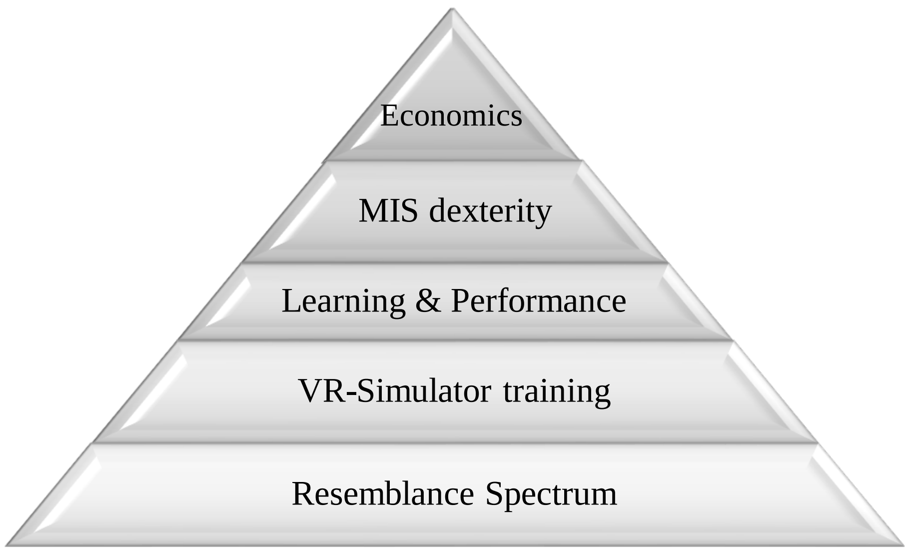



*Figure 2*. Schematic representation of design requirements

**Requirement for resemblance**

**Shortages in science**. Literature about the relevance of VR-training and assessment for skill acquisition in MIS hints at lacking predictive power for either scientific methodology (Medford, 2013; Kramp, et al., 2016) , skills (Wolbers & Hegarty, 2010), innate abilities (Groenier, Schmettow, Huijser, & Gallagher, n.d.; Anastakis, Hamstra, & Matsumoto, 2000) and inferences on individual differences (Stather, MacEachern, Rimmer, Hergott , & Tremblay, 2011). Further shortages related to perspective taking and the content of representations (Klatzky, 1998) or the insecurity about the causes for individual differences in navigational abilities (Wolbers & Hegarty, 2010) were discussed. Shortages in explanatory power reach from tests, objective evaluation criteria for the tested skills or capabilities and suitable tasks related to MIS procedures (Anastakis, Hamstra, & Matsumoto, 2000). To summarize, what still needs more scientific verification, is the transfer from what has been tested and trained (skills, tests, methodology) in the ‘laboratory’ to the real world MIS, related skills and resulting performance – the spectrum of resemblance.

**The resemblance spectrum**. The resemblance spectrum describes the continuum of validity and reliability from beginning trainable dexterity skills over more realistic simulator tasks until performance improvements in real MIS. The problem is that innate abilities often are imprecise abstractions of real surgical tasks and often request for simple responses, as for example button presses. In comparison to this, professional simulator tasks are complex and realistic imitations of realistic surgery tasks (Groenier, Schmettow, Huijser, & Gallagher, n.d.). Until now, research about simulator performance has provided the best predictions for real surgery performance, but ability and skill tests have rather resulted in unsatisfied conclusions and still prove the trade-off between resemblance and costs. Tasks with a low degree of resemblance of real surgery tasks, for example creating origami (Arendt, 2017), are low priced, but more resembling tasks as professional simulator tasks depend on high costs. This problem emphasizes that there is still a need for low-budget tasks, which contribute to training MIS skills and can support expensive simulator tasks. Another reason is the need to better understand the underlying mechanisms, which make virtual reality simulators for MIS so successful.

Originally, we aimed to approach the polarity between resemblance and predictive power, through the validation of our prototype with a VR-simulator performance comparison (see Figure 3). Now we approach the resemblance spectrum through the creation and evaluation of a new low-budget tool for training MIS related skills.

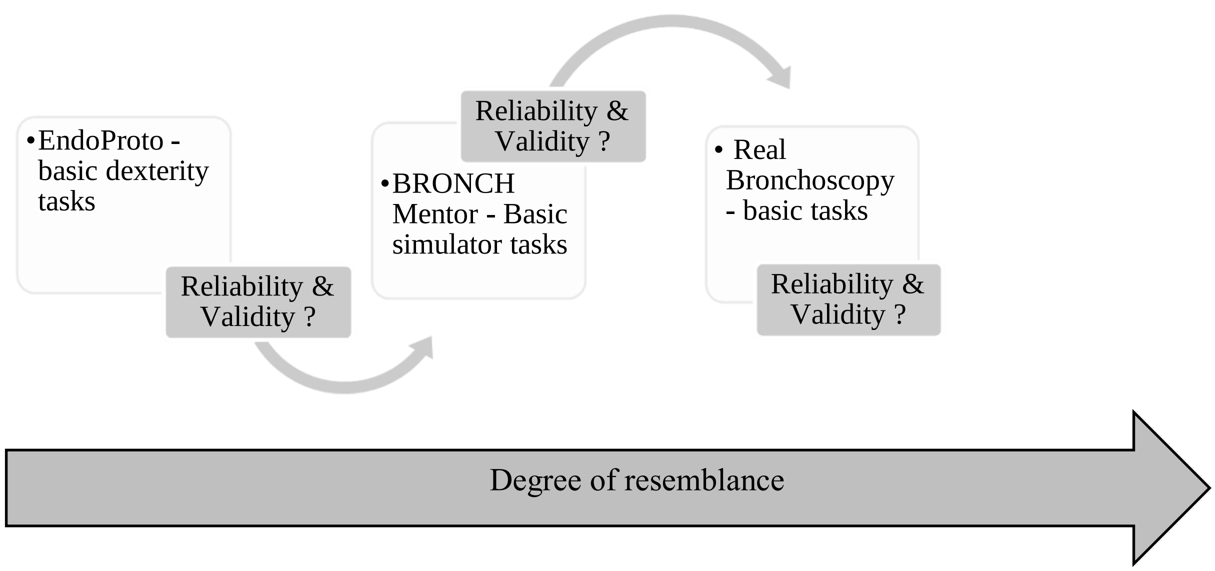



*Figure 3.* Resemblance Spectrum. The x-axis demonstrates the degree of resemblance. The y-axis shows the performance parameter. The linear relation of the test suits indicates an ideal (but unrealistic) training scenario with performance increase. The EndoProto shall be used as a first instrument for unexperienced trainees. The arrows indicate a relationship based on reliability and validity between the test suits. Own design, based on: (Schmettow & Groenier, 2017, p. 26; Arendt, 2017, p. 13).

**Requirement for VR-simulator training and assessment**

Medical training was long time only regarded as pure education. This means, it was less focused on the available research on learning, training and the development of skills (Gallagher & Smith, 2003). Traditional training methods in MIS and were mainly characterized by completing the curriculum as well as observing and assisting professional surgeons. In this context, practical training consisted of extracting the endoscope from the human body and finally exhibiting it for simple tasks. Complex procedures as biopsies were only admissible to highly experienced surgeons (Ost, et al., 2001).

**Virtual reality (VR) simulators**. The relevance and opportunities of technical devices as VR simulators were long time not recognized for skill acquisition in MIS (Gallagher & Smith, 2003). This verifies, why traditional educated surgeons were either rarely (Ost, et al., 2001) or not at all (Expert 1, personal communication, June 23, 2018) experienced with VR simulators. For example, from the 9 experts out Ost et al's study (2001) only two had experience with simulators (mean age: 41). Even before the 1990 VR-simulators for training MIS skills existed, but they were facing problems of validation (Gallagher & Smith, 2003). Finally, in 2002 VR training was argued to decrease error rates for endoscopic procedures in the operation room in a scientific justified way (Gallagher & Smith, 2003). The effectiveness of VR simulator training has also been reported to improve performance for endoscopic (Ohuchida, et al., 2009), laparoscopic (Sangani, Bradley, & Fidopiastis, n.d.; Seymour, et al., 2002) and most important here, flexible bronchoscopy (Medford, 2008; Ost, et al., 2001; Stather, MacEachern, Rimmer, Hergott , & Tremblay, 2011). On the basis of expert levels the skill acquisition of novices was confirmed via learning curves and parameters as speed or wall collisions (Ost, et al., 2001).

These results show on the one hand the advantage of skill acquisition for novices until levels of expertise (Ost, et al., 2001) via simulator training against conventional methods. Furthermore, a background for an additional training and evaluation method was set, which offers more objective assessment criteria than traditional training methods (Ohuchida, et al., 2009; Ost, et al., 2001; Seymour, et al., 2002). These arguments build another aspect where the current study ties in – the constant adjustment of training and assessment for surgeons. Therefore it shall be stated here, that the scientific relevance of stimulator training for skill acquisition of real MIS is still questionable and insecure (Kennedy, Maldonado, & Cook, 2013), because of the shortage of studies (Medford, 2013) and the broad diversity in methodology. This diversity is mainly grounded on p-values, as the study of Stather et al. (2011) and leaves fewer possibilities to evaluate individual differences.

We use another approach instead, as we base our results on Bayesian statistics, in terms of building inferences on the quantification of uncertainty (Schmettow, in prep., p. 4 ch 3). Moreover, we offer a contribution through including individual differences in our statistical and argumentative analyses. Altough we were not able to integrate VR-simulator training in the current study, we still were able to compare the developed prototype with a professional and validated (Konge, 2015) VR-simulator – the BRONCH Mentor.

**Requirement for learning and performance**

**Skill acquisition**. Skill acquisition for MIS is of high relevance for the success and quality of a surgery. The acquirement of skills is defined in four phases of learning: First, the unconscious incompetence, second the conscious incompetence, third the conscious competence and fourth the unconscious competence. Stage one and two are characterized by low success rates, high need cognitive effort and concentration, leading to phases of exhaustion. The latter two stages show decreased cognitive efforts and increased amounts of automated movements. Professional skills most likely can be achieved in the last phase (Mohamed, Raman, Mclaughlin, Anderson, & Coderr, 2010). Other authors have categorized skill acquisition in the following three phases: cognitive, associative and autonomous stage, at which the last phase symbolizes the one in which expertise can be established (Fitts & Posner, 1967).

In order to master certain skills learning is essential and learning is consolidated through repetitions. Research about the amount of necessary repetitions for training MIS procedures, vary according to domains. It was claimed, that the first 10 repetitions show the steepest part of learning in laparoscopy and can reach until 50. Based on estimated learning curves, the number of repetitions ranged from 25-100 for cystoscopies, over 45 for procedures in the medical field urology and until 300 for gastrointestinal endoscopies (Gallagher & Smith, 2003). The schedule for medical training in bronchoscopy demands the execution of up to 250 procedures in a training-period of 5 years. The recommendation for repeating bronchoscopies involves at least 100 procedures (Medford, 2008). Other authors have reported performance increases at up to 120 specific bronchoscopic procedures (Lin, Lai, Chang , Wen, & Ho, 2018). Finally, surgeons who are described as experts in flexible bronchoscopy have conducted more than 1200 examinations and observed even more than 2000 surgeries (Ost, et al., 2001). Nevertheless, there is no consensus about the minimum of necessary training intervals in bronchoscopy and reported results are lacking scientific relevance (Medford, 2013; Ost, et al., 2001).

**Learning curves.** It is obvious, that training and experience are necessary to reach expertise. This expertise is characterized by decreased time and effort and an increase in repetitions of a certain task or procedure, which finally is defined as learning effect. An analytical representation of such an effect is a learning curve (Srour, Kiomjian, & Srour, 2015). A learning curve demonstrates an exponential function, expressing a skill which develops over time (see Figure 4).

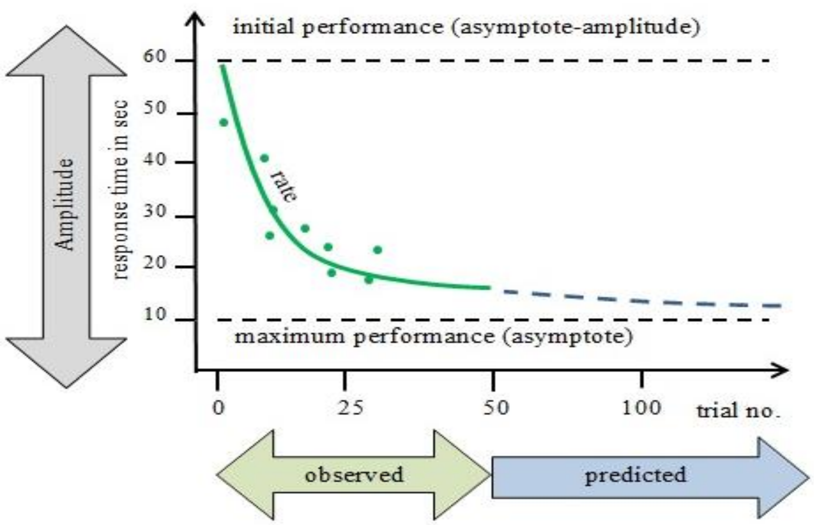



*Figure 4*. Anatomy of a learning curve. Own design, based on (Schmettow & Groenier, 2017, p. 6). For illustrational reasons trial number and response time intervals can be broadly regarded as examples for the current study.

Every skill is starting with initial performance (asymptote – amplitude) and typically is reinforced with repeating a task (Estes, 1956). The amplitude represents the initial performance or state of a certain skill (Groenier, Schmettow, Huijser, & Gallagher, n.d.). It shows the change of learning expressed through time. The slope parameter indicates if a performance parameter as for example time-on-task is dependent on repetitions. In other words, the acquisition of a certain skill can be indicated by decreased ToT and thus show the proposition of learning through the slope. In Figure 4 an ideally case is demonstrated, where a person continuously is getting closer to the maximum performance. Learning generally is faster at the beginning and progress gets smaller over time. Professional skills can be regarded as a grade nearest to the maximum performance, the asymptote. At this maximum stage theoretically no learning is possible anymore (Heathcote, Brown, Mewhort, 2000).

Learning curves can help to identify individual differences, because they give a detailed picture on procedural learning effects. In comparison to that, mean learning curves should not be averaged, because they are not adequate for drawing conclusions about individual learning effects (Ackerman, 1987; Estes, 1956; Heathcote, Brown, & Mewhort, 2000). In addition to that, career recommendations should be specified to individuals, due to the fact of the widespread margins and suggestions of necessary training units reported before by Gallagher and Smith (2003) and Ost et al. (2001).

Even more important, the low success in finding reliable learning curves in literature about bronchoscopy show the limits in research for this domain. Some authors even state, that learning curves for conventional bronchoscopy are *very variable, hard to predict and may be longer than thought, even among experienced bronchoscopists* (Medford, 2013, p. 418). Supplementary, other authors proved strong differences in learning speed among experienced bronchoscopists in endobronchial ultrasound examinations, even showing learning progress above a margin of 100 repetitions (Kemp, et al., 2010). Related to these outcomes we would expect that novices would differ in learning basic bronchoscopic procedures. Based on the reported findings, the current study focuses not only on the exploration of procedural skill acquisition through learning curves. We also emphasize the relevance of individual differences for novices in bronchoscopy.

**Prototype Design**

Until now we argued for the first three design requirements (1) degree of resemblance, (2) VR-simulator training and (3) learning and performance. These requirements correspond to a fundament of science (resemblance), technology domain (MIS and VR-simulator training and assessment) and related approach (learning curves) and outcome (learning and performance enhancement). What is still missing is the content of the training and learning we aim to achieve (see Figure 5). Our contribution in the first place was the construction of a prototype for flexible endoscopy and bronchoscopy. This prototype (EndoProto) first had to fulfill two more requirements: First, to foster MIS dexterity skills in order to establish domain related training and second, to be practical and cost efficient in order to be simply reproducible and easy to use. These two requirements are discussed in the following. Finally we present the constructed prototype and finish the chapter with our research questions.

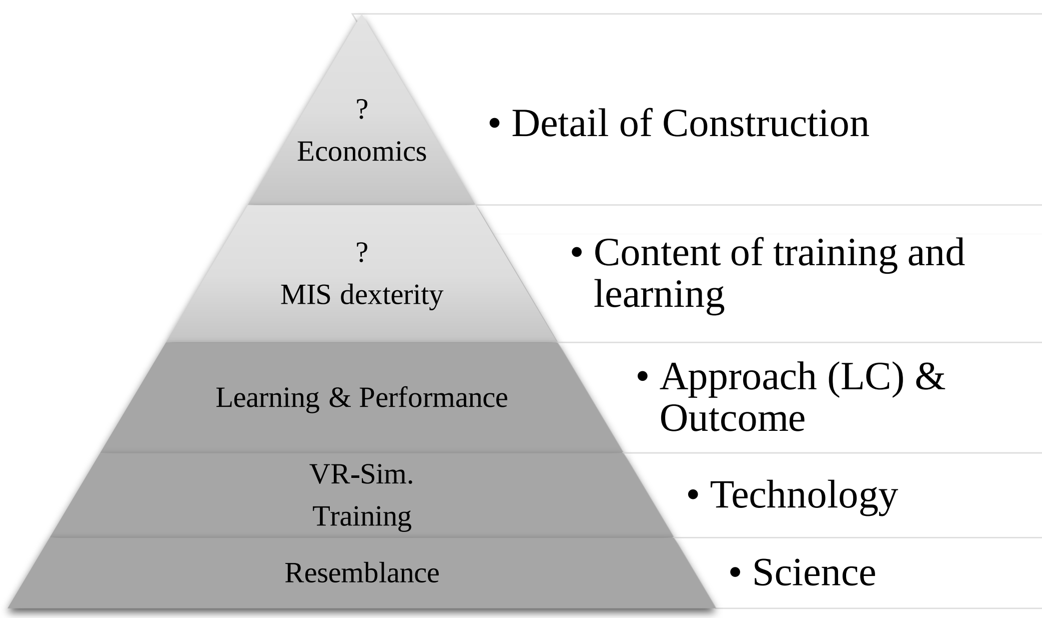



*Figure 5.* Progress in approaching requirements.

**Requirement for MIS dexterity**

The present thesis focuses on surgical dexterity, because aspects of hand-eye coordination were integrated in the developed prototype tasks (related to the original interest in investigating hand movements with a sensor-glove). Although we face the risk of not being able to train these abilities, we still hope to gain more insights in individual differences and in the aspects which make the VR-simulator training successful in training MIS skills.

**Surgical dexterity**. The importance of psychomotor and dexterity skills is widely discussed in literature about MIS, especially for the purpose of training (Gallagher, Leonard, & Traynor, 2009). Professional surgeons themselves evaluated cognitive factors, personality traits and innate dexterity as relevant attributes for surgical competence (Cuschieri, Francis, Crosby, & Hanna, 2001; Gallagher, Leonard, & Traynor, 2009). Surgical dexterity defines the following abilities: *Spatial perception, hand eye coordination, aiming, multilimb coordination, hand-arm steadiness* (Cuschieri, Francis, Crosby, & Hanna, 2001, p. 112). Abilities were valued as relevant for open and MIS. Visual-spatial abilities are defined as fundamental abilities. Due to their assumed genetically predetermination they are differentiated from technical and thus trainable skills. Fundamental abilities as dexterity or hand-eye coordination are regarded to be necessary to execute some MIS related skills as for example suturing (Gallagher, Leonard, & Traynor, 2009).

***Visual-spatial abilities.*** In addition to the above named abilities, in MIS it is necessary to draw adequate conclusions form images (Cuschieri, Francis, Crosby, & Hanna, 2001). In order to be able to draw such conclusions, surgeons must transform and coordinate visually perceived information and spatial relations. This is referred to as visual-spatial ability. Results of previous research add up to the following categories of visual-spatial perception: (1) perceptual recognition of objects, (2) visual imagery involving 2D representations of the reconstruction of objects from their parts and (3) visual imagery involving 2D and 3D whole object rotations and translations (Anastakis, Hamstra, & Matsumoto, 2000, p. 470). Related to MIS, one concrete example would be the transfer of 3-dimensional (3D) bodies into 2-dimensional (2D) pictures. This ability is highly relevant in flexible endoscopy, because surgeons have to transform the bodies or related parts they imagine onto the screen on which they see the inspected area of the endoscope. Vice versa, they have to encode the 2D video information they see on the screen into related motor movements to operate the endoscope inside a 3D human body (Anastakis, Hamstra, & Matsumoto, 2000).

Summarizing, visual-spatial abilities are strongly related to surgical performance (Gallagher, Leonard, & Traynor, 2009; Wanzel, et al., 2003) and reported to be highly relevant for gaining professional performance in MIS (Anastakis, Hamstra, & Matsumoto, 2000). Important to mention is also, that newest scientific findings prove, that innate abilities have no predictive value skill acquisition in MIS (Groenier, Schmettow, Huijser, & Gallagher, n.d.). The according authors warn to rely on innate abilities in accordance to surgical performance assessment and career recommendations. The aim with the current study was not to try to encounter these findings, but to investigate learning and performance with a self-created prototype. We still integrated surgical dexterity in our research setup, because we regarded even the replication of previous findings as valuable.

***Visual ability - perspective taking*.** As the functionality, comprehension and the transformation of visually perceived information are so important for surgical dexterity, the topic of perspective taking should be covered as well. Prior to this, the meta-level of a reference frame needs to be defined. In simplified terms, a reference frame is a *means of representing the locations of entities in space* (Klatzky, 1998, p. 1). Related to reference frames, two broad categories are mentioned in literature - the allocentric and the egocentric perspective. Broadly spoken, these two are dedicated to *spatial perception, spatial cognition and spatially directed acting* (Klatzky, 1998, p. 2). The egocentric reference frame is always related to the perspective of a certain perceiver and to his/her localization. Hence, in the egocentric perspective the perceiver defines the axis of orientation (Klatzky, 1998). In other words, it represents a distance from the self to an object (Wolbers & Hegarty, 2010). In comparison to that, the allocentric or exocentric reference frame is related to the positioning of the surrounding, external to the perceiver. Simplified, allocentric describes the distance form an object to another object (Wolbers & Hegarty, 2010).

In order to relate these perspectives to the context of MIS, surgeons now constitute the primary perceivers. As the perspective of a perceiver defines the reference frame, it is the perception of a surgeon which defines the perspective he/she takes. Further, our context includes VR-simulators for training purposes, which indicates that the environment in a simulator is not real, but virtual. As stated in literature, the perception in a virtual environment is *centered on the position in virtual space of the virtual body* (Slater & Wilbur, 1997, p. 3), or more exactly to the eyes of the body. In bronchoscopic surgery, there is no virtual body displayed as for example in a video game. Instead surgeons see the innards of a human body and maybe instruments of a biopsy. Bronchoscopists maneuver through the human bronchia form an egocentric perspective, because they autonomously maneuver the movements of the endoscope, which can be regarded as a mechanical abstraction of their virtual body. In this context, the camera of the endoscope represents the eyes of a surgeon which demonstrates the axis of orientation and thus can be regarded as an egocentric perspective.

Including egocentric perspective taking in the current study offers a possibility to reach a higher resemblance of real simulator task, which also shows the camera-view of the endoscope.

Even more relevant is the argument, that the application of an egocentric perspective in a virtual environment is reported to increase task performance, even when controlled for the same ability level of participants. Furthermore, egocentric perspective taking seems to give a higher impression of presence, which is defined as a sense of being in a place (Slater & Wilbur, 1997). Although this presence seems less relevant for the current study, it is relevant to mention that presence seems to influence the autonomous response and behavior of a person in a virtual surrounding, which might have an effect on the performance of a surgeon.

***Spatial ability – Navigation and cognitive maps.*** The above described ‘mechanical abstraction of their virtual body’ sounds a bit odd, and this egocentric perspective surely cannot be generalized to every surgeon and every situation during surgery. However, the paragraph above shows that a reference frame is connected to movement, especially in a virtual environment. Furthermore, this implies that perspective taking also is related to spatial abilities, as for example navigation abilities and the ability to generate cognitive maps.

The concept of a cognitive map describes the comprehension of an environment in its dimensional properties. The labelling cognitive map is derived from the idea that the human mind creates something similar as a city map, with local representations of attributes. This implies that such a map is ‘drawn’ from the perspective of a bird, or differently said in an allocentric perspective. Differently, in the context of spatial navigation abilities researchers define that the perceivers of such maps can view through its local points of orientation. This means, that the cognitively generated representations of the surrounding are internally visible and not related to a specific orientation or perspective. This adds up to the point that any perspective can serve in a cognitive map to navigate between objects or to infer distances between them (Wolbers & Hegarty, 2010).

Literature about MIS also covers the relevance of spatial navigation abilities for bronchoscopic training (Naur, Konge, Nayahangan, & Clementsen, 2017) and cognitive maps also for endoscopy (Taylor, Brunyé, & Taylor, 2008). Concluding, as MIS demands for visual-spatial abilities, cognitive maps can be relevant in order to orientate inside the human bronchia and to give a correct localization of a possible tumor. Further, orientation in navigation is fixed to the scope end, which defines the point of view, enabling illumination and symbolizing the central point of movement (Taylor, Brunyé, & Taylor, 2008). Grounded on these facts, we can as assume that a surgeon primarily maneuvers the endoscope via an egocentric perspective, but probably also make use of perspective switching.

On the other hand, a cognitive map of the anatomy of a human lung helps to orientate inside the human bronchia. The skill to shift between the perspectives is thus regarded as essential for mastering expertise in surgery. Moreover, perceptual and visual-spatial abilities are strongly related to surgical performance (Gallagher, Leonard, & Traynor, 2009), but surgeons also differ clearly in their navigation skills and preferences for perspectives (Taylor, Brunyé, & Taylor, 2008). We want to supplement these findings through the integration of different perspectives in the prototype tasks. Additionally, disorientation is reported to be one of the most occurring causes for errors and restriction for success in MIS. Disorientation is defined as a surgeon's *uncertainty of the exact location of the scope end* (Taylor, Brunyé, & Taylor, 2008, p. 27). Although, we do not primarily focus on disorientation in this study, we still measure the task precision of a participant, which can serve as an indicator for orientation.

**Surgical dexterity (skills) vs innate ability.** The reference point of the described surgical dexterity skills is the effect of training. If training dexterity skills relevant for MIS or bronchoscopy could improve surgical performance, this could contribute to training and assessment criteria for surgeons. Until now, it is not clear if the relation between spatial abilities and surgical skills is rather transient or enduring. Transient means that the association exists at early levels of learning, whereas enduring means that the relations is persistently affecting performance, not only in the early phases of learning (Keehner, Lippa, Montello, Tendick, & Hegarty, 2006). In this context surgical dexterity skills may be grounded to some extend in innate visual-spatial abilities and thus are reported to be genetically predetermined (Gallagher, Leonard, & Traynor, 2009). On the contrary, there are authors arguing that experience can positively influence the construction spatial mental models and resulting navigational skills for endoscopy (Taylor, Brunyé, & Taylor, 2008). Virtual-reality trainers were constructed for training endoscopic navigation skills and construct validity was reported (Haluck, et al., 2001).

The findings above indicate that the boundaries surgical skills and innate abilities may be not that clear. Mentioned aspects of MIS related visual-spatial abilities, skill acquisition and performance processes emphasize the need for a better understanding of the underlying relations for training. The evaluation of skills can have severe consequences for careers regarding recruitment and assessment. Therefore it is necessary to know the underlying reasons for drawbacks in surgical skills and if training can compensate for them. If individual differences in surgical dexterity are purely based on genetics and do not positively affect training, this would have severe consequences for applications and job selection criteria. The current study contributes to get a better insight about learning effects of basic scope insertion tasks with a low-fidelity prototype. We regard even the replication of previous research findings as relevant.

**Requirement for economics**

**Low-budget & low-fidelity**. One of the main problems of VR-simulator training in MIS is the one of high costs (Gallagher & Smith, 2003). These costs are not only related to the training tool (VR-simulator), but also to the training sessions, supervision and evaluation. For this reason, one of the requirements for our prototype was summarized as economics. We aimed to create a low-budget prototype with which we hoped to prove performance improvements for novices. Further, this economics included that the prototype should be easily transportable, easily usable for a possible wide range of people and less dependent on technical equipment as costly VR-simulators.

**Endoscopy prototype (EndoProto)** [1]

Based on the discussed requirements for (1) resemblance, (2) for VR-simulator training and assessment, (3) learning and performance, (4) MIS dexterity and (5) economics, the following prototype for flexible endoscopy (EndoProto) was created (see Figure 6). In the following we supplement the demonstration and explanation of the prototype with our motivations and decisions for the task design and end this chapter with a summarizing contribution of the current study.

*Figure 6.* EndoProto from frontal perspective with dividing wall. Own photo.

[1]The basic design of the endoscopy prototype was inspired by the training box for pre-oral endoscopic myotomy, produced by students from the University of Twente and the department of technical medicine (Buser, Pouw, & Kusters, 2016-2017)

The prototype design was primarily based on the domain of flexible endoscopy, with the aim to enable to train several related domains as the flexible bronchoscopy. The EndoProto is an open polymer box[2]. This box has uniformly distributed holes in its outer frame. These holes can be used to place dividing walls inside the box. Such a dividing wall has the same hole-pattern as the broadside of the box. The hole-pattern of a dividing wall demonstrates different routes, which can be crossed with a provisional endoscope (low-budget wire camera, see Figure 7). The insertion of the endoscope inside the box and through a route was the fundamental scope insertion task in the current study. In total we defined two main tasks, which based on the same insertion process but differed in the induced perspective. One task was defined as allocentric task (without camera perspective) and one was defined as egocentric task (based on the direct view ‘through’ the endoscopic camera visible on a laptop screen. For further technical details on the EndoProto and reasons of reproducibility we refer to the technical drawing (see Figure E1) in Appendix E.

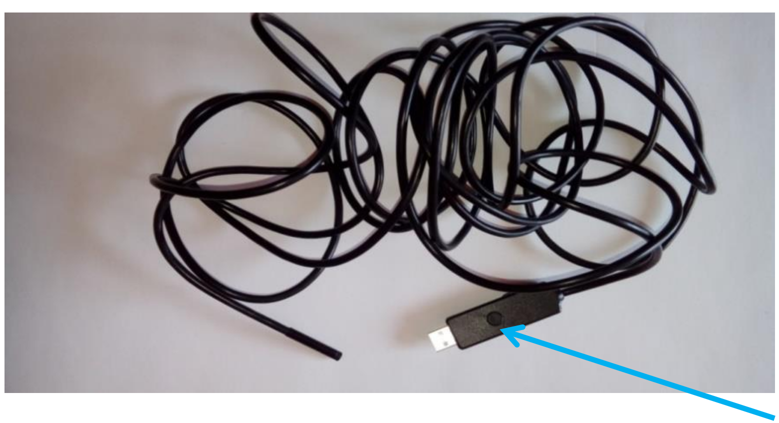


*Figure 7.* Endoscope (low-budget Android wire camera) with light control button. Own foto.

[2] For describing the construction of the EndoProto in more technical terms, the labelling (box) is usede here exchangeable with EndoProto

**Dividing walls**. For the allocentric and the egocentric task each a fitting dividing wall was created. Every of these two dividing walls had the same number of similarly distributed holes. These holes are defined as routes. The routes are numbered based on their levels of difficulty (see Figure 8). The levels of difficulty rise from 1-20 in circuits around the center (route 1). A difficulty level based on circuits means, that the difficulty is not rising linear from 1-20. The reason for this non-linear difficulty levels is, that they are dependent on the technical and physical possibilities of the used endoscope. This means, that the complexity rises with greater distance and angles to the starting point parallel to route 1 (yellow). In addition to the numbering a coloring pattern was created, which also visualized the circular levels of difficulty. This visualization was chosen for reasons of supporting orientation, instruction simplicity and comprehension. The coloring scheme was chosen based on observations and experience with the pilot test.

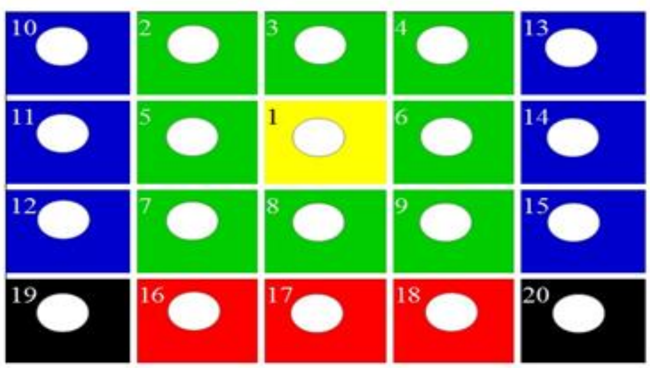



*Figure 8*. Numbering and color scheme for the routes (yellow = very simple, green = simple, blue = moderate, red = difficult, black = very difficult). For reasons of clarity routes were numbered in colors with the highest contrast possible, in white and black. Own design.

**Manipulation of the EndoProto for the egocentric task.**

***Starting Point***. We created two main tasks for the current study, according to the induced allocentric or egocentric task. One broadside of the box was marked with numbers (see Figure 6), which indicated the starting point for the allocentric (1) or the egocentric condition (2). Two different starting points were necessary, in order to start in both conditions at the same position, parallel to the first route (1 = yellow, see Figure 6). In the egocentric task the box was turned vertically and the dividing wall was fixed upside down, which led to a shift of the first route. For this reason we had to adjust the starting point (2) to the position of the first route.

***Dividing wall***. For the different allocentric and egocentric tasks the dividing walls had to be adjusted. For the allocentric task, the hole-pattern of the dividing wall was similar in number and diameter to the outer frame of the box. For the egocentric task, the hole-pattern was just similar in number, but had a bigger diameter (see Figure 9, dividing wall 3 vs. 2). This difference in diameter was designed because of reasons of feasibility. Own testing's and the pilot test revealed a higher complexity of the egocentric task. This led to an adjustment in the diameter, in order to prevent disproportionate rates of wall contacts. The dividing walls for both conditions had the same amount and pattern of routes, visualized in numbering and color.

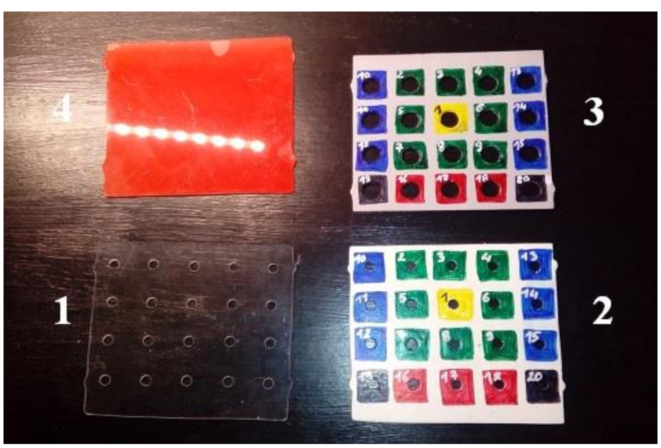



*Figure 9.* Dividing walls (right photo). The transparent dividing wall (1) shows an unpainted construction of a dividing wall, which served as a contingency replacement. The dividing walls with the coloring pattern were used for the allocentric (2) and the egocentric (3) task. We used the red wall (4) for haptic feedback and as a more realistic background for the egocentric task. Own photo.

***Feedback and measurement security.*** The EndoProto was manipulated for the egocentric task. This manipulation is defined by the creation of a second dividing wall, which was inserted one row beyond the one with the coloring scheme (see Figure 9, dividing wall no. 4). The main reason for this manipulation was to enable a saver registration of the performance of participants for us researchers. First, through the coloring a better visual contrast was reached in the black box, which helped to observe the success of crossing a route. Second, hitting the red wall with the endoscope generated acoustic feedback. This feedback helped us researchers to observe and calculate the success for passing a route and the amount of wall contacts of a participant. Third, hitting the red wall served as haptic and acoustic feedback for participants. The last reason was to reach a better resemblance of the inside of a human body, through a red-colored human-tissue-like background.

**Motivation for Task Design**

**Endoscope (Scope) insertion task**. The insertion of the endoscope inside the human body is a standard procedure in MIS (Lechtzin, Rubin, White, Jenckes, & Diette, 2002), which served as a model for the current task construction. Insertions are important to train, because of their risk of infecting a patient with microorganisms (Kovaleva, Peters, van der Mei, & Degener, 2013). Moreover, some surgeries for example the treatment of gastrointestinal fistulas are of high risk for a patient's life (Truong, Böhm, Klinge, Stumpf, & Schumelick, 2004). A certain amount of dexterity (hand-eye-coordination) is necessary to handle the tiny and sensitive parts of an endoscope, especially when inserting it in the human body. Therefore, this task and the related handling are recommended to be trained by surgeons (Mason, 1992). Based on these findings, we developed a basic scope insertion task, in order to try to simulate simplified bronchoscopic insertion procedures (which originally would have been trained in an even more specified way with the VR-simulator). Due to the simplistic prototype design and the missing operation system for the endoscope, we just let participants insert the endoscope from a starting point of the EndoProt into a directed opening (route) and extract it again.

**Perspectives**. Referring to literature, box trainers were reported as efficient training tool for MIS skill acquisition (Ohuchida, et al., 2009). Such skills were defined as hand-eye coordination in accordance to the comprehension and related execution of visual information (as transmitted from the endoscope and presented on monitors). This finding was one of our motivational reasons for including perspectives in our training tasks. Another fundamental reason was to approach the resemblance of a real VR-simulator (BRONCH Mentor) tasks. Therefore, the EndoProto training was designed to consist of two simple endoscope inserting tasks which were defined by their induced perspectives, as allocentric and egocentric task. We hoped to simulate and train the shifting between allocentric and egocentric (camera-view) perspectives, through the use of a provisional endoscope. The first task (allocentric) was defined by an open box (EndoProto) and a direct view on the own hands and related movements. This task was thus expected to foster primarily an allocentric perspective. The second scope insertion task was defined by a closed box and a view through the endoscope inside the box via a camera-perspective. This task was thus expected to primarily foster an egocentric perspective. It is important to highlight, that these conditions cannot exclude that participants made use of different or more than one perspective in the described tasks

**Posture.** The Endoscopic task was regarded as a second training task. This means, first that we expected participants would have gained some experience with the training tool and related skills. Second, we wanted to adjust the next task more to a basic VR-simulator task. Therefore, we changed the position of the EndoProto through turning it upside down. In accordance to this change in posture we instructed participants to execute the egocentric task in a standing position, which led to a higher resemblance to the standing posture for the VR-simulator. Related to our original research focus, this change in posture was also expected to maybe change hand and finger movements (measurable with the sensor glove).

**Distance**. We decided to integrate distance changes in the sub-tasks in order to keep the attention of the participants high. Task 1.1 and task 2.1 were executed in the same distance (1). From the three chosen distances of the allocentric condition, the easiest was chosen for the egocentric task, because this task was expected to be more difficult and time intense than the allocentric task. All mentioned expectations and evaluations were based on trials of another researcher; own trials and the pilot test (Appendix C).

***Distance in allocentric condition***. The decisions for the chosen distances of the allocentric conditions are based on the following evaluations. Previous tests showed that distance 1 was most fitting for first encounters with the EndoProto. Therefore, the closer distance (2) was chosen as a second sub-task for the allocentric condition instead of being the first task. The task with distance 2 was valuable in addition to distance 1, because of the expected simplicity and resulting learning effect for the handling of the endoscope. The third distance appeared to make task 1.3 nearly as difficult as the first task, maybe even simpler. Task 1.3 was thus regarded as an additional task, if time permitted it. All mentioned expectations and evaluations were based on trials of another researcher; own trials and the pilot test (Appendix C).

***Distance in egocentric condition.*** The egocentric task was not varied in distance. The reasons for the chosen distance are the pilot test and our own trials, which proved this one to be the most realizable in accordance to time and performance measures. In addition to that, we decided against the nearest distance of task 1.1, because it was nearly not executable in the camera perspective, because of the bad view on the display. We also decided against the widest distance of task 1.3, because of the expected lack in time

**Performance parameters**. Related to the developed tasks, we accounted for relevant performance parameters discussed in literature about MIS training and assessment. We integrated three basic performance parameters discussed in relation to MIS and bronchoscopy (Ost, et al., 2001): First, one parameter for the task duration (ToT) second, one parameter for errors (wall contacts) and third, one parameter accounting for problems with disorientation (precision), which is discussed as a main problem in MIS (Taylor, Brunyé, & Taylor, 2008).

**Contribution of the current study**

With the current study, we wanted to contribute to lacks of pulmonary (Stather, MacEachern, Rimmer, Hergott , & Tremblay, 2011) or endoscopic training and assessment, through approaching the resemblance spectrum with a new test instrument fulfilling our five design requirements. We tried to slightly approach the resemblance of a real VR-simulator (BRONCH Mentor) and develop our basic scope insertion tasks grounded on previous research. Our aim was to hopefully foster the learning and performance of MIS relevant skills with the low-budget and low-fidelty EndoProto. We contribute to previous research through investigating individual learning curves, instead of just focusing on inaccurate group means. Further, our study contributes with the statistical analysis, which is based on magnitudes and measures of uncertainty, instead of conventional p-values. Summarizing, our extended research setup is visualized in Figure 10.

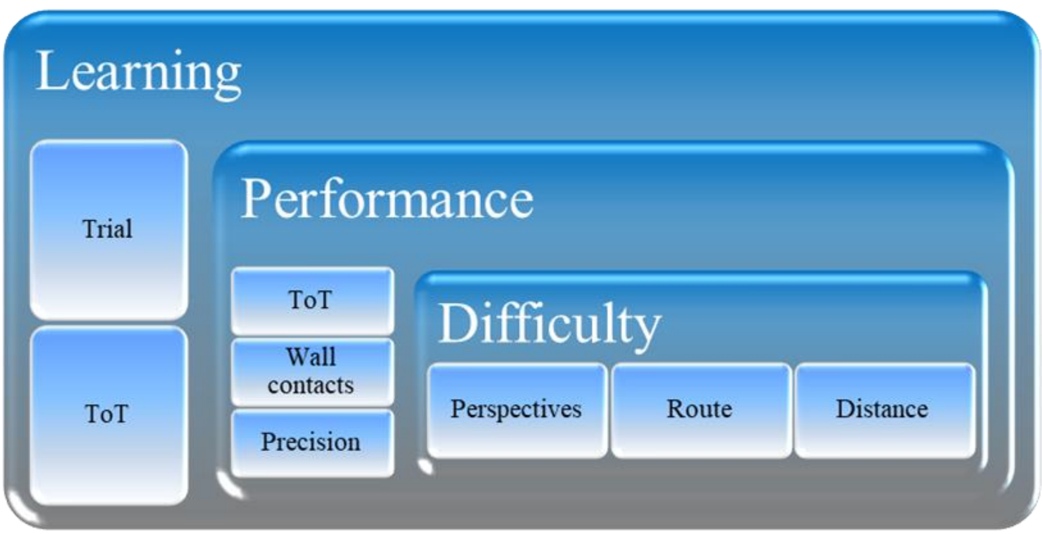



Figure 10. Advanced research setup with according parameters.

**Research question**

For the current paper, no hypotheses have been generated. Instead, an explorative approach was applied, in order to get insights in the scientific contribution of the EndoProto for the resemblance spectrum of MIS and especially bronchoscopy. Our main focus was the investigation of learning. For this purpose, First wanted to estimate learning curves with the parameters asymptote, rate and amplitude with our parameters for ToT and trials. Our second aim was to focus on individual differences in learning. And our third aim was to investigate differences of our two main tasks due to expected perspectives. Therefore, our research questions were:

1. Is it possible to estimate learning curves with the amplitude ToT and the rate parameter trial?
2. How do learning curves differ between individuals?
3. How do the learning curves differ under the allocentric and the egocentric condition?

## Experimental Design Method

### Recap Focus of the study

The estimation of learning curves failed. Therefore, we shifted the focus of the study from the exploration of learning towards an exploration of the underlying performance parameters (design requirement 3: Learning and performance). Our analyses simplified to a multi-level mixed effects model in which trials became exchangeable repeated measures. This enabled us to investigate different aspects of performance and the underlying parameters of difficulty. Accordingly we explored possible reasons why no learning took place. Finally, we interviewed two experts, in order to evaluate and adjust the EndoProto for possible future investigations. We were not able to answer our research questions without learning curves. For this reason, we generated two quantitative and one qualitative research questions according to the adjusted research setup (see Figure 11).

1. How do individuals differ in the performance parameters?
2. How do participants differ under the circumstance of perspective (allocentric and egocentric)?
3. What can the opinions of experts add to the current study?

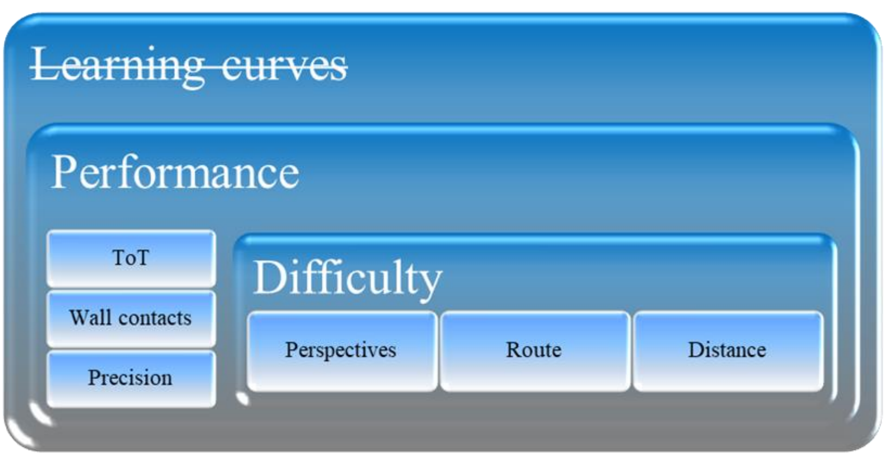



Figure 11. Adjusted research setup

**Design[3]**

For the current study a within-subjects design was applied. Every participant executed two main tasks in a training session of 1 hour. The tasks were differentiated by conditions. Every task was expected to induce the use of a different perspective. Therefore, we defined the tasks by their perspectives as allocentric and egocentric task (see Figure 12).

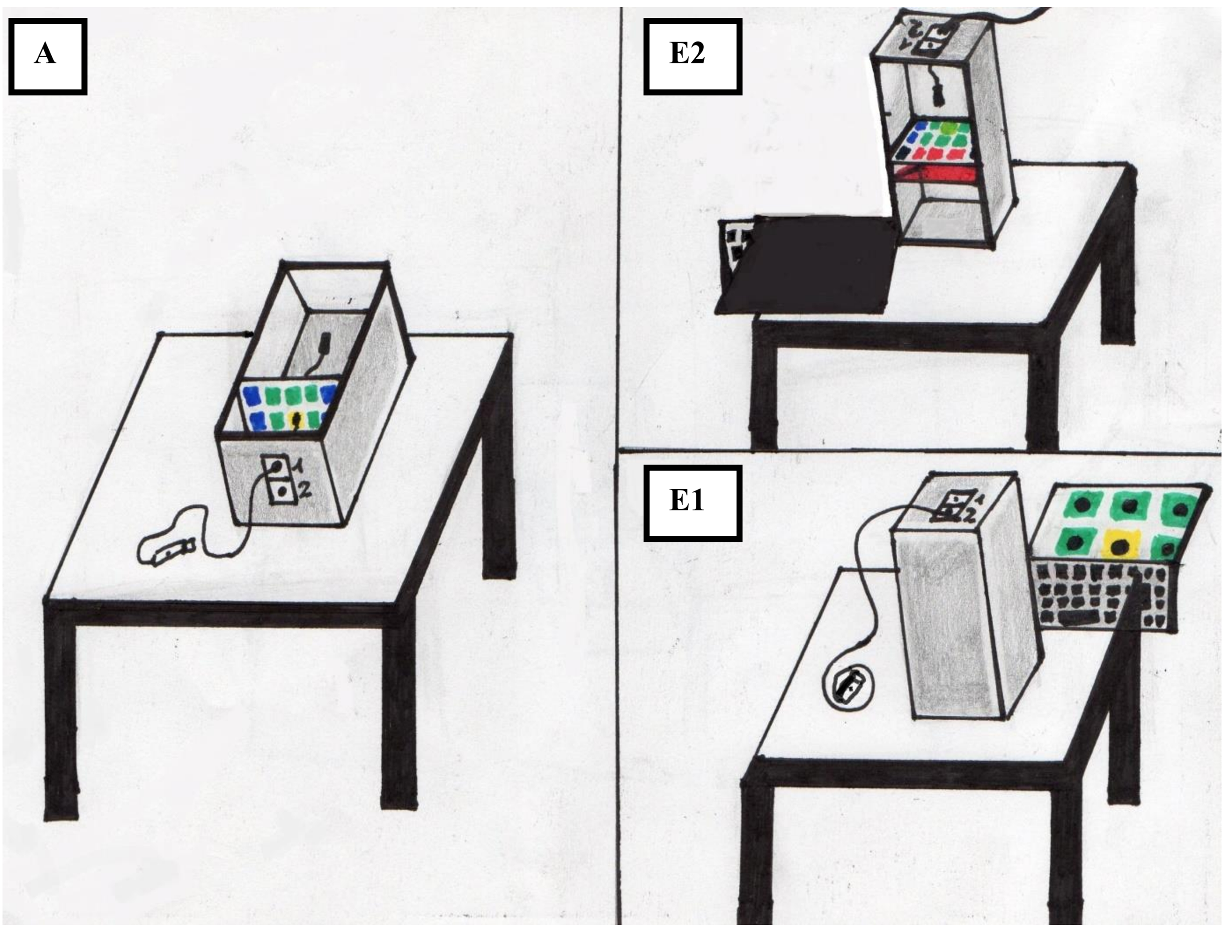



*Figure 12*. Schematic illustration of the allocentric (A) and egocentric (E1 and E2) condition. In (A) researchers and participants viewed the EndoProto from the same perspective. The egocentric condition is illustrated from the viewpoint of the participant (E1) and from the viewpoint of the researcher (E2), sitting on the opposite side of the table in order to be able to count wall contacts. Own drawing.

[3] The word *design* has been used before in this paper, but with a different meaning. Under the heading design just the experimental design is defined. All other references to design are related to the creation of the EndoProto.

The allocentric task was separated in three sub-tasks (1.1-1.3), which were characterized by distance (see Table 1). This distance was defined in accordance with the numbering of the sub-tasks and their executed sequence. The distance of a sub-task was read off from the rows on the longitudinal side of the box. Rows were technical indicators for the placing of the dividing walls (see technical drawing E1, Appendix E). A dividing wall included possible routes to take. The numbers of routes ranged from 1-20 and indicated their levels of difficulty (see Materials). To summarize, difficulty was expected to be induced first by the condition of perspective, second by the distance of the dividing wall and third by the dividing wall's routes.

Table 1

*EndoProto Task Overview*

| | Level of Difficulty | | | | | |
|---|---|---|---|---|---|---|
| *Condition* | Sub-task | *Distance* | Row* | *Routes* | Time | Trials |
| Allocentric | 1.1 | 1 | 3 | 1-20 | 5-10 | 20 |
| Allocentric | 1.2 | 2 | 2 | 1-9 | 1-5 | 10 |
| Allocentric | 1.3 | 3 | 4 | 1-20 | 5-10 | ≥ 20 |
| Egocentric | 2.1 | 1 | 3 | 1-9, 16-18 | 25 | ≥ 30 |

*Note*: Time shows the expected duration in minutes and trials the expected number of repetitions for a sub-task. We generated random routes for every trial.
*The row is a technical indicator for the placing of the dividing wall. A row is no variable and does not signalize a level of difficulty. The indicated numbers stands for the holes on the longitudinal side of the EndoProto (See Figure E1, Appendix E). Own design.

**Task restriction**. We restricted both tasks in a time-frame of 25 minutes for a session of one hour. This time restriction was necessary, in order to enable the same amount of training for both conditions. For this reason sub-task 1.1 and 1.2 had to be also restricted in their number of trials. We expected that participants would reach at least 30 trials in every condition. Therefore, task 1.3 was regarded as an additional task, which was only possible if the time-frame permitted it. The relation between the expected time and related number of trials in Table 1 was grounded on own trials and results of the pilot test (Appendix C). We chose for a guideline of approximately 30 repetitions. This decision was based on scientific results about bronchoscopic learning curves showing a steady performance increase even after 20 repetitions (Ost, et al., 2001).

**Materials**

**Endoscopy prototype (EndoProto).** A self-created prototype for testing skills in flexible endoscopy and related domains as flexible bronchoscopy was created and used for the current study (see Prototype Design, Figure 6).

**Endoscope[4].** We used a USB Android wire camera (see Prototype Design, Figure 6) as endoscope. It served as a provisional and simplified version of a real endoscope. Compared to a real endoscope, the tip of the wire-camera was not movable mechanically. The camera was just movable via the movement of the wire itself. The dimensions of the endoscope are a 5.5mm diameter and a length of 5m. The device is equipped with LEDs and offers the possibility of HD video recordings. By connecting the device to Android or Windows XP/VISTA/7/8/10 it enables to take snapshots and real-time-camera-perspectives. For the last mentioned action we made use of the software "ViewPlayCap".

**Demographic questionnaire**. On the online platform Survey Monkey we created a demographic questionnaire (Appendix D). The questionnaire was a precondition for taking part in our study. Therefore, we made the link available with our study via SONA systems (https://www.surveymonkey.de/r/NN383DD, accessed 2018-06-15). The survey contained questions about demographics as age, gender, nationality, study, handedness, limitations of visual strength or color vision, dyslexia, gaming behavior, prior knowledge and experience in the field of simulated endoscopy.

[4] The used endoscope is a provisional low-budget and low-fidelity version of a real endoscope. Due to its basic feature of the wire camera and its usability for flexible endoscopy, the terminology 'endoscope' will be used here adequate to the EndoProto. The term 'bronchoscope' is not used, because this would just include the context of the current paper, but not the assumed range of possibilities for using and exploring this tool.

**Data collection**

**Qualitative and quantitative data.** We collected and measured quantitative and qualitative data for the current study. We noted the quantitative data for every participant directly in a research protocol (Appendix E). The quantitative data was transferred and saved in anonymized form in Excel tables on Google Drive. Additionally, we collected qualitative data (1) during the training sessions with the EndoProto and (2) after the completion of participant training in form of expert interviews. The first mentioned qualitative data were observations and personal comments of participants, which we tabulated in an observation protocol (Appendix G). The second mentioned qualitative data was extracted and tabulated in Appendix I (Expert Talks).

**Quantitative: Performance measurements**. In total we measured five quantitative performance parameters: (1) Trials, (2) ToT, (3) wall contacts, (4) precision and (5) skips. The most relevant performance parameters for the current study were the number of repetitions (trials), task duration (time-on-task: ToT) and the number of errors (wall contacts). The last two parameters have been validated in a study about the assessment of a bronchoscopy simulator for the skill acquisition of novice bronchoscopists (Ost, et al., 2001). We defined our performance parameters as follows:

(1) We calculated the trials of every participant in every task. Every single trial started at the related starting point of the allocentric or the egocentric task (see Figure 5) and ended with visibly crossing a route. After a trial was finished, the participant visibly had to extract the endoscope out of the box and insert it again for the start of a new trial. Every trial was counted as 1.0.

(2) We measured the time for every trial (ToT) with the stopwatch, which was integrated in the mobile phone of the related researcher. The limits of ToT were equivalent to the

defined begin and ending of a trial. We noted the time in seconds into the research protocol of the related participant.

(3) One aim of every task was to work as *precisely as possible* (Appendix F, 5. Guidelines). This means, that we explained participants to try to not touch any wall during the execution of a trial. Every caused contact with the walls (if touching or scratching) during a trial was counted as one wall contact (1.0). Wall contacts were only calculated in relation to the definition of a trial. This means, that wall contacts which were caused after crossing a route, for example during the extraction of the endoscope, were not counted. When participants touched a route directly – this means, when they already moved the tip of the endoscope inside a hole of a dividing wall - this action was not counted as a wall contact. We decided for this exclusion criteria, because the diameter of the routes was too small to not touch it at all. Conversely, when participants already touched a route, but slipped off again, their action was counted as one wall contact. We decided to include this action as a wall contact, because participants were not successful at finishing their trial.

(4) The precision indicated if a participant was able to insert the endoscope through the right route. A correct precision was labelled in the Boolean programming variable TRUE. If the participant went through a wrong route, this action was noted down as false precision (FALSE). Additionally, the number of the wrong route was noted down in the research protocol. Then the current trial was automatically skipped and a new route was instructed to the participant.

(5) Participants who took longer than two minutes to complete a route were granted the option to skip. When the skip option was approved by us researchers, we instructed the participant with a new route. When approved, we labelled this action in the Boolean

programming variable TRUE, otherwise we labelled it as FALSE. In contrast to this option, participants could also complete their present route in a longer time than two minutes. The option to skip a route was introduced in the sessions, because previous tests of us researchers and the pilot test showed strong individual differences in ToT. Due to the reason that ToT and the number of trials are relevant for the estimation of learning curves, we tried to foster reasonable ToTs, which were not only dependent on the difficulty of routes. Finally, we also wanted to prevent people from frustration. For more detail on the measured performance parameters we refer to the glossary of the research protocol in Appendix F.

**Procedure**

Two researchers supervised the participants in the procedure. We informed all participants about the nature of the research, the requested tasks, related rules, answered open questions and thanked them for their participation. Explicit verbal instructions can be found in the research protocol (Appendix F). Generally participants took part in one session of one hour. However, two participants had to split the session, because of personal circumstances. In these cases we decided for a splitting of the session based on the two conditions, which resulted in two sessions of 30 minutes each. The session was split in half, in order to not interrupt the training of one condition.

**Allocentric task.** Participants started to perform simple scope insertion tasks on the EndoProto. This allocentric tasks took about half an hour. Depending on the performance of participants they executed either two (1.1-1.2) or three (1.1-1.3) sub-tasks (see Table 1). All sub-tasks were executed in a sitting posture. The EndoProto was placed horizontal in front of the participant and the coloring scheme as well as the inner space of the prototype was visible to them (see Figure 13). We fixed the prototype on the table with a rubber map to prevent it from moving. We generated random routes for every participant and instructed them one by one until the limited number of repetitions or time was reached (see Table 1).

*Figure 13.* Demonstration of a training-like scenario. The box is placed in a horizontal position (for task 1.1.-1.3.) with inserted endoscope through translucent wall (for visibility, not part of the study) and with illumination. Own photo.

***Rules and instructions***. The task of participants was to insert the endoscope in the right starting point and cross the route given by us researchers. Our instruction before every condition was to move *as fast and precise as possible* (Appendix F, 5.). Then the endoscope was extracted out of the box and a new trial was started. Participants had one rule: Hands were only allowed outside the box. In view of this rule, we assumed to foster relevant hand and finger movements for basic scope insertion tasks.

**Egocentric task.** We granted the participants a break of 5 minutes between the allocentric and the egocentric task. The egocentric task is defined by the assumed perspective participants would take to execute it. For this reason, we adjusted the setting of the condition through placing the EndoProto in a vertical position (see Figure 14) and connecting the endoscope with a computer, which resulted in a real-time transfer of the camera view of the endoscope (see Figure 15). We hoped to induce an egocentric perspective with the direct view through the camera (read *visual ability* - *perspective taking*).


*Figure 14.* EndoProto in vertical positioning for the egocentric task with red wall for haptic feedback from the perspective of the researcher (left photo). The white arrow shows the visible side for participants, looking from above on the EndoProto. Own photo.
*Figure 15*. Camera perspective shown on a laptop screen (right screenshot). Own photo.

We asked the participants to stand up and hold this position for the remaining 30 minutes.. The ground of the EndoProto was turned to the participant, so that they could not see their movements inside. One of us researchers took place on the other side of the table, where the open sided of the EndoProto was viewable. This positioning was necessary to be able to see the counted performance parameters (wall contacts, precision and trial).

***Rules and instructions***. The egocentric task was the same as the task of the task of the allocentric condition: To insert the endoscope in the right starting point and cross the route given by us researchers. We gave the same instruction as for the allocentric task and explained the same rule. This rule was enlarged by the restriction to only follow the movements of the endoscope via the laptop screen. We allowed participants to extract the endoscope in order to re-orientate at any time. Such an extraction was not counted as a new trial, when the instructed route was not already crossed.

***Difference to allocentric task.*** The egocentric task differed in two aspects to the allocentric task: Frist in the fixed distance (see Design: Motivation for distance) and second in the time restriction. Our guideline was that participants could reach at least 30 repetitions and our aim was that they would reach as many as possible in the 25 minutes of the session. Nevertheless, we expected participants to reach fewer repetitions as in the allocentric condition, because of the difficulty in perspective taking and orientation. Our experience with the pilot test and our own trials demonstrated the difficulty to estimate the range of possible repetitions in a given time. We discovered that the repetitions in the egocentric condition took more time than in the allocentric condition. In order to show possible learning effects, we decided for a time restriction and aimed for a high amount of repetitions.

**Location.** In most cases the EndoProto training took place at the University of Twente in the ECTM's (Experimental Center for Technical Medicine) simulator room 2. This simulator room was our initial place for all planned EndoProto and simulator sessions. However, acoustic disturbances and unavailability of the room induced us to book separated conference rooms in the University. We expected room changes to not negatively influence the performance of participants first, because silence would rather contribute to the concentration of participants and second, because the experimental execution of the EndoProto tasks just required a table and a chair.

**Debriefing.** We thanked the participants and offered to send them their results via email. This email would have contained the error rates (wall contacts) task duration (ToT), success or failure of reaching a destination (precision) and their skip rates.

**Participants**

**Formality**. The University of Twente's faculty of Behavioural, Management and Social Sciences (BMS) has an Ethics Committee which approved our study (approval No.: BCE18145). All participants signed an informed consent before participating in our study (Appendix B). If requested, we handed out a copy of the informed consent to the participants. The University provided a SONA test subject pool where we recruited our convenience sample of ($N = 25$) participants. The first participant served as a pilot test (Appendix C) and was therefore extracted from the dataset of the analysis. If not stated differently, the remaining ($n = 24$) participants represent the sample we talk about in the following.

**Demographics**. Six participants of our sample were male. The average age was 21 years (range: 18-28 years, $SD = 2.0$ years). The majority of the participants studied psychology ($n = 21$), followed by communication studies ($n = 3$). Twenty participants were German, two were Dutch, one was Bulgarian and one was Iranian. One participant was ambidextrous. The remaining participants were right-handed. One of the participants was colour-blind and another one had dyslexia. Four participants were wearing glasses. Twelve participants had experience in video/computer gaming, ranging from 1 until 10 hours a week. No participant had experience in the field of simulated endoscopy.

**Exclusion criteria**. Physical disabilities regarding hand, arm and leg movements were exclusion criteria for the study, because these abilities were relevant to conduct the tasks sufficiently. As an example, for the second task it was necessary to stand upright for about 30 minutes. Corrected vision was no exclusion criterion, such as wearing glasses or contact lenses. Colour blindness was also no exclusion criterion, because the comprehension of the colouring scheme was not essential to complete a trial.

**Data analysis**

The original focus of the study was to investigate learning effects through the estimation of exponential learning curves with a non-linear learning curve model. The estimation of learning curves failed. Instead, we based the data analysis on a multi-level mixed-effects model in which trials became exchangeable repeated measures. Regarding our research questions, we focused our analyses (1) individual differences in performance based on the parameters ToT and wall contacts and (2) the exploration of differences between the allocentric and the egocentric perspective on individual and group levels. The performance variable precision is a logic dichotomous variable and was therefore not usable for a regression analysis. The difficulty parameter distance was integrated in the parameter task, in order to analyze either the sub-tasks or the different perspectives. As a result we only had to deal with the two difficulty parameters perspective and route. Finally we found an interaction effect of route by task for wall contacts.

For the sake of formality, first the learning curves model and then the multi-level model will be described. For all analyses we used the program Rstudio and the programming language R (version R 3.4.4.), in which we integrated the packages ‘brms’ (version 2.3.0). Figure 16 shows the extended research setup for the data exploration and result section.

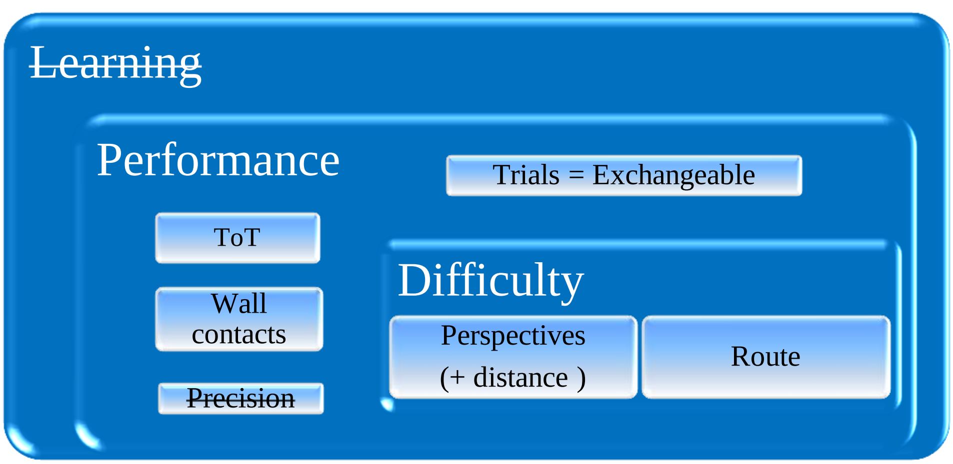



Figure 16. Research setup for data exploration.

**Learning curve model**. Learning curves are used as likelihood functions of a statistical model. The mathematical formula of a learning curve is as follows:

$\mu_i = \beta_{asym}(1+\exp(-\beta_{rate}(x_{training}+\beta_{pexp})))$. (Schmettow, in prep.)

For our estimation of learning curves, we created the LARY model, based on the parameters (amplitude, rate, asymptote). In our LARY model the parameter ToT was the amplitude, trials expressed the rate parameter and the asymptote was the maximum performance. A learning curve demonstrates, ththat the rate of learning is related to the amount of training. Learning typically rises with every unit of (x) and experience ($\beta_{pexp}$), which together are signalized as the slope. $\beta_{asym}$ stands for the asymptote.

***Manipulation check.*** Through link functions, linearity was generated ranging from a log-scale of [-∞, +∞]. This means, that all possible outcomes were covered by this function, which may also include unrealistic ones. For example, the experience of a participant expressed in trials could not be lower than 1, because every participant started with his/her first trial on the EndoProto training. As a reverse conclusion, a negative amount of trials would also not be realistic. Moreover, the variance in a sample due to individual differences can be huge. Therefore, we used link functions to account for these differences through generating random effects. This means, that individual learning curves were bundled and included in the learning curve parameters (asymptote, amplitude, rate) per task, so that an analysis could be executed on an individual and on a population-level. This population level is indicated by averages (fixed effects). Moreover, variance residuals typically get smaller by closing to the asymptote. For that reason, we chose Gaussian over Gamma distribution. We dealt with frequently left-skewed reaction times.

**Generalized linear model (GLM).** After the estimation of learning curves failed, we analyzed the effects of the performance parameters ToT and wall contacts. For this purpose, we made use of generalized linear model (GLM), in which trials became exchangeable repeated measures. We created a multi-level mixed-effects model with a negative binominal distribution and executed regressions. Our model accounted for overdispersion, which is the linear increase of variance proportionally to the mean (Schmettow, in prep.). In order to prevent this risk we made use of a two parameter distribution – the negative binominal distribution. Through link functions, we reached linearity. Most importantly, the variance of every measurement could be used as a function of the predicted values. This was important, to investigate performance on the population-level and to include random effects in the model. (1) In order to present results about individual differences, we had to compare population-level effects to the standard deviations (SD) on the participant level on a linear predictor scale. Because SD cannot be retransformed to their original scales, we compared on linear predictor scales and found fixed effects on the population-level. (2) We created plots to visualize the differences in performance between the perspectives. (3) We accounted for the possibility that route effects differ by tasks and accordingly we found an interaction effect of routes by tasks.

Calculated coefficients were based on 95% credibility intervals. 95% CI is an indicator for the estimation that the true value has a chance of 95% to be in the range between the upper and lower boundaries given by the coefficient table. Values which are closer to each other are an indicator for good certainty. Regarding the calculated estimates certainty is measured on a probability scale between 0 (not possible at all) and 1 (certain) (Schmettow, in prep.).

# Results

In the current study we analyzed quantitative and qualitative data. First, data exploration of exponential learning curves will be presented. Second, we will present an explorative post hoc analysis on individual differences and the differences between the allocentric and the egocentric condition. Finally, we present the qualitative results of the expert interviews.

## Quantitative analyses

**Exploration of Learning.** For the exploration of learning we made use of a non-linear multilevel mixed-effects model to estimate learning curves. We analyzed the effect of the performance variable ToT on the asymptote. It was not possible to estimate learning curves with the available dataset. This means, no learning took place for the performance variable ToT. In more detail, the performance of the participants varied but did not improve after repeated trials, neither on a population (see Figure 17) nor on a participant level (see Figure 18 and 19). Consequently, predicting skill acquisition of participants through the BroProto was not possible. Out of these reasons, we evaluated the predictor trials as not fitting independent variable for the non-linear multilevel mixed-effect model. Finally, we concluded the learning curves model to be not beneficial for the current study.

***Raw learning curves: Population level.***

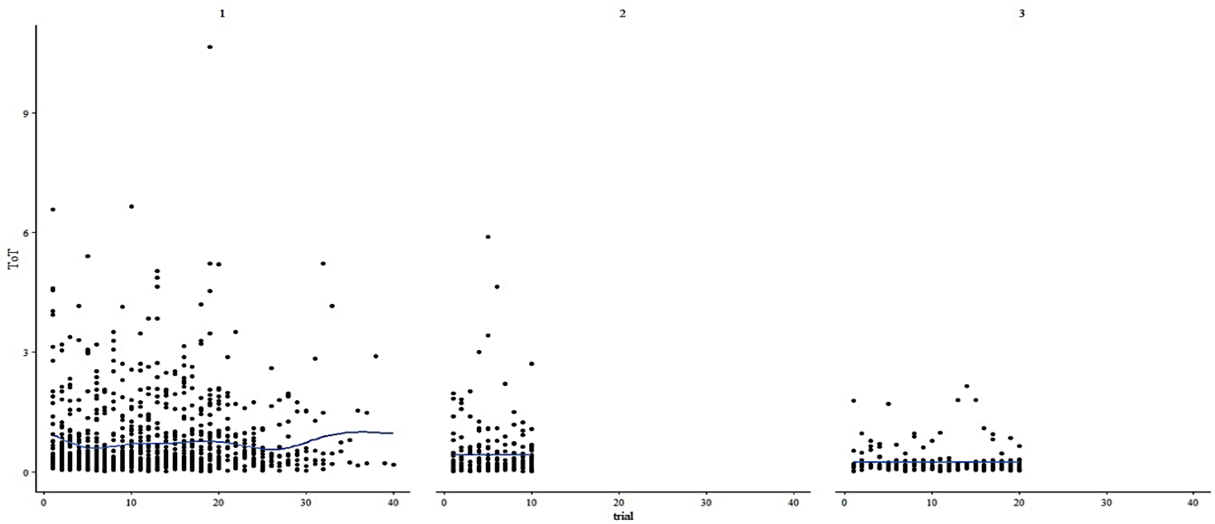



*Figure 17.* Raw learning curves – population level. The learning curves are estimated for allocentric, task 1.3 (1), 1.2 (2) and 1.1 (3), showing the number of trials on the horizontal axis and the task duration in minutes (ToT) on the vertical axis. Points indicate separate trials and colored smoothers show the gradients. Learning curves show the relation between learning (a progress would be indicated by a decrease of task completion time) and experience (demonstrated through an increase in repeated trials).

***Individual learning curves: Allocentric task.***

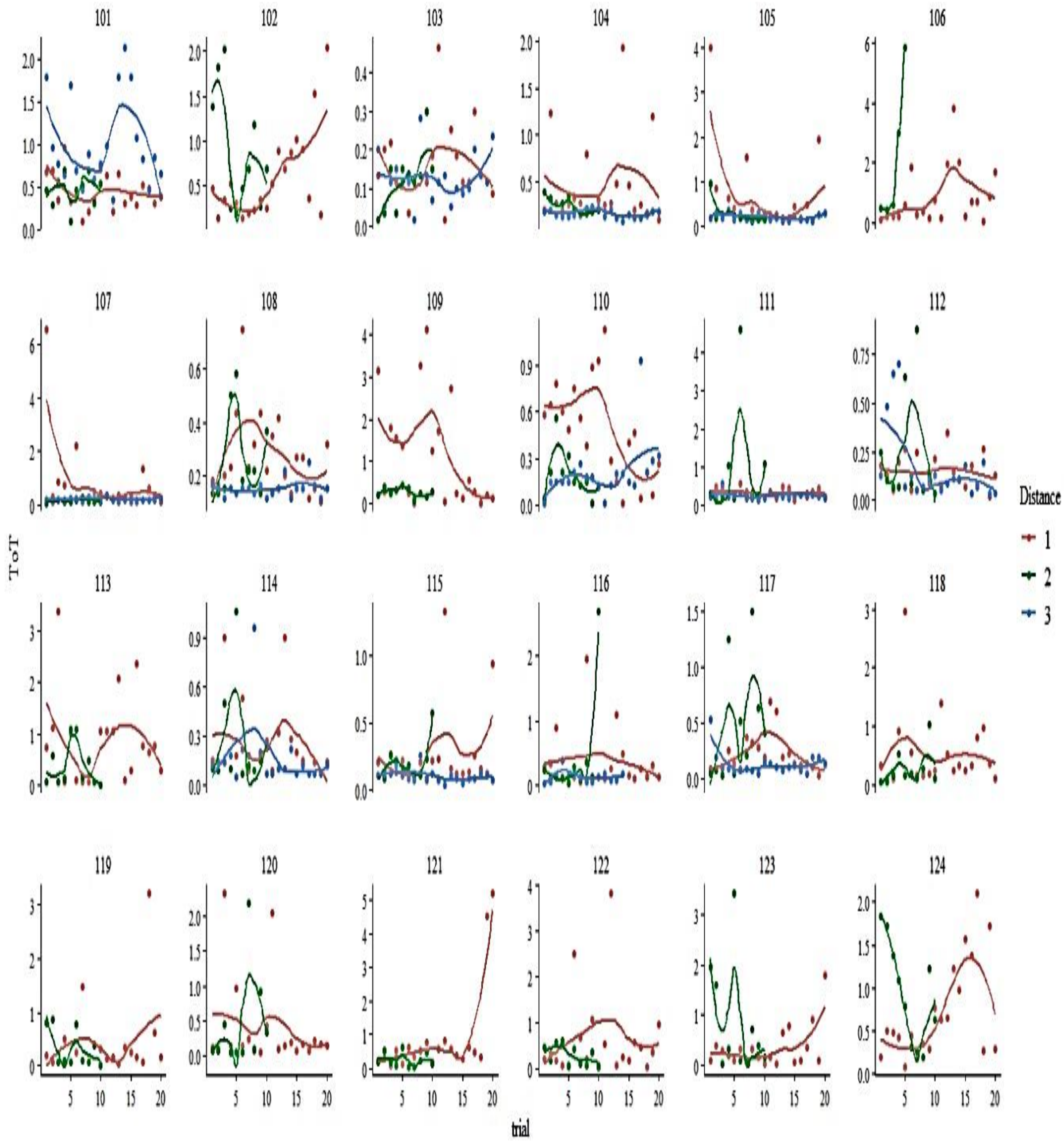



*Figure 18.* Raw learning curves – participant level – allocentric. The learning curves are estimated for task 1.3 (1, blue), 1.2 (2, green) and 1.1 (3, blue), showing the number of trials on the horizontal axis and the task duration in minutes (ToT) on the vertical axis for every participant. Points indicate separate trials and colored smoothers show the gradients.

***Individual learning curves – Egocentric task.***

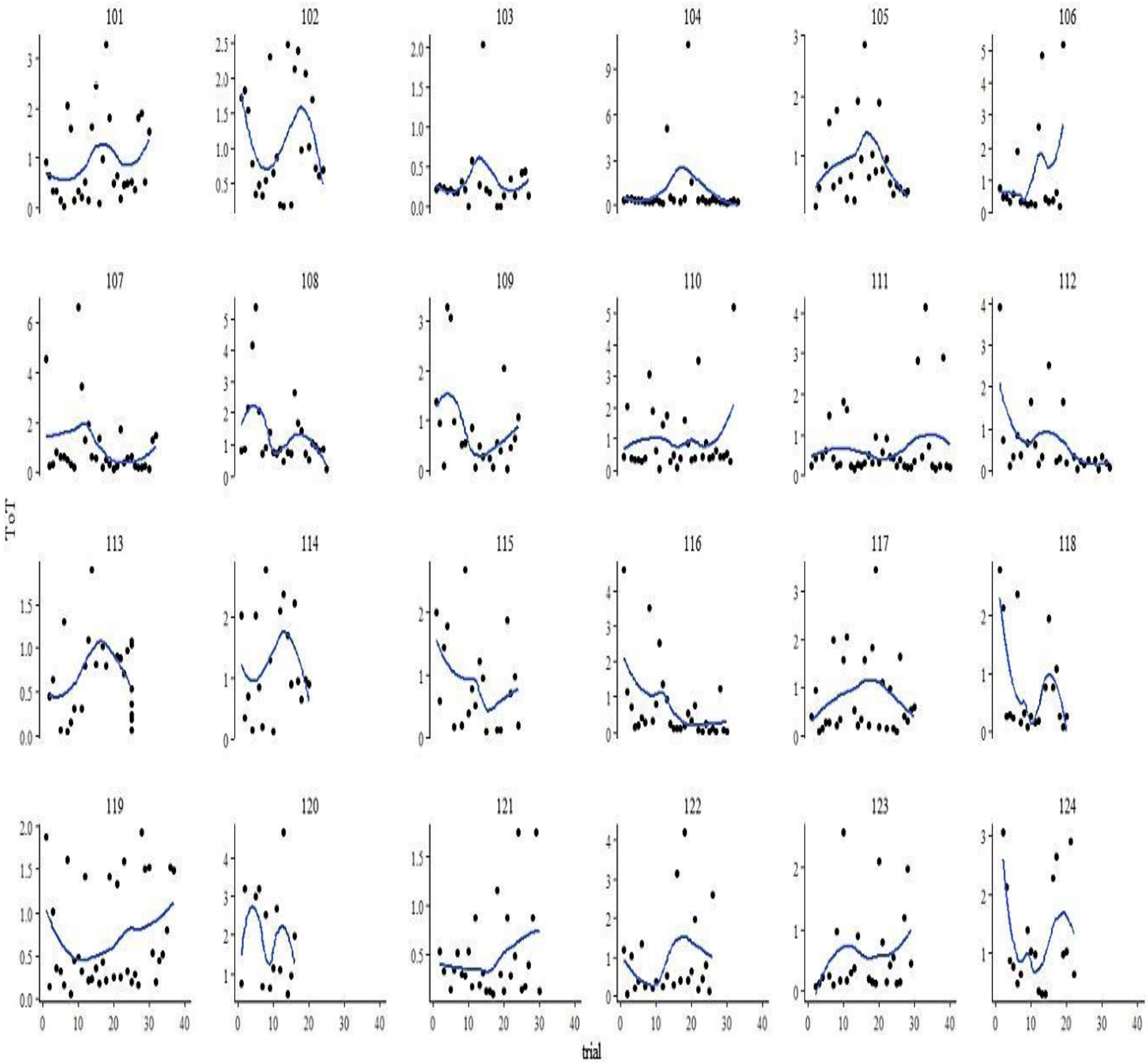



*Figure 19.* Raw learning curves – participant level – egocentric. The learning curves are estimated for task 2.1. The x-axis shows the number of trials and the y-axis shows the ToT in minutes. Points indicate separate trials and colored smoothers show the gradients.

**Exploration of performance**

The estimation of learning curves failed. Instead, we focused our data analysis on the exploration of performance through the parameters ToT and wall contacts. We applied a mixed effects model with a Poisson distribution. Through link functions linearity was established and the variance of the measurements was integrated as a function of the predicted value. We used the predictor trial as exchangeable repeated measure. We centered ToT on the population average and shifted the variable trial by one so that the intercept represent the first trial. We executed regressions with Participants, Routes and Tasks (including distance and perspectives) as our main predictors for ToT and wall contacts. This led to the following aspects of analysis: (1) The differences in performance between the two conditions (allocentric and egocentric) and (2) individual differences and (3) the influence of the routes on performance.

**Individual differences**. Population level effects for the association between performance variables ToT and wall contacts were compared to individual differences as standard deviation of by the parameters (see Table 2). Effects were predicted on a linear predictor scale, because standard deviations cannot be transformed back to original scales. The average effect of causing wall contacts at the first trial is 0.22 [-0.19, 0.60] *95CI* on a population level, but the difference between individuals is more than twice 0.55 [0.39, 0.80] *95CI*. Security about these effects is not high, because confidence intervals are not dense. The effect of ToT is 0.57 [0.44, 0.72] *95CI* on the population level and the individual variation is nearly half of it 0.24 [0.11, 0.42] *95CI*. Security about these effects is not high, because confidence intervals are not dense. Most participants have a positive relation between ToT and wall contacts (see Appendix H – R syntax). Individual variation is visible, but not overwhelming, which is visible in Figure 20.

*Table 2.* Association of ToT and wall contacts on population and individual levels

| | Population | | Individual SD | |
|---|---|---|---|---|
| Performance Parameter | Effect estimate | 95% CI | Effect estimate | 95% CI |
| Wall contacts | 0.22 | [-0.19, 0.60] | 0.55 | [0.39, 0.80] |
| ToT | 0.57 | [ 0.44, 0.72] | 0.24 | [0.11, 0.42] |

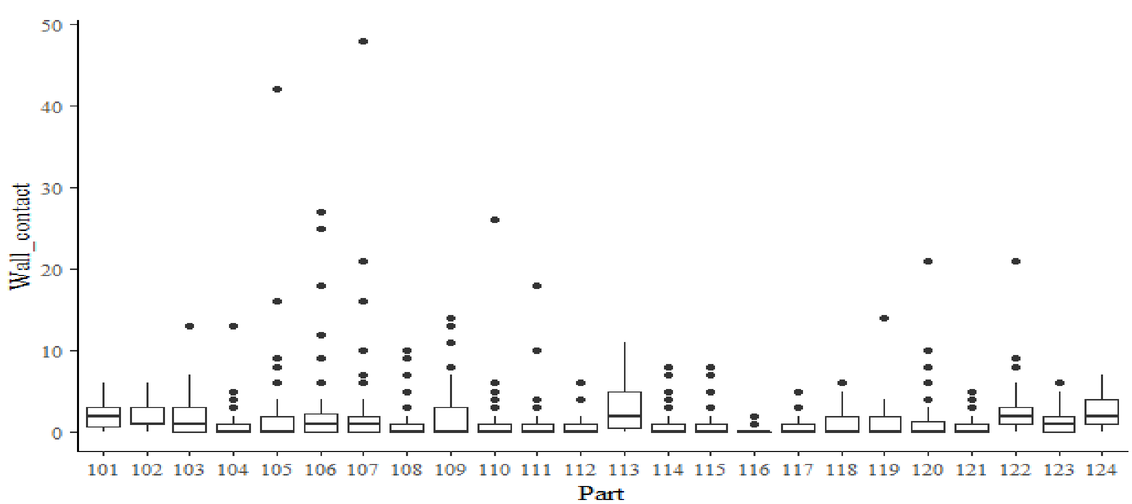



Figure 20. Individual number of wall contacts with outliers.

**Differences between conditions: Population level.** The effect of ToT is 1.77 [1.55, 2.06] *95CI* (see Table 3). With every minute longer the number of wall contacts multiplies by 1.77, which is an increase by 77%. Estimates are based on pretty good certainty. The effect of causing wall contacts differs strongly between sub-tasks and conditions. The effect estimate of every task indicates the difference (values are multiplicative) in causing wall contacts to the average at the first trial in the allocentric task. On a population level participants were estimated to cause most wall contacts on the first allocentric task 1.25 [0.83, 1.82] *95CI*. Participants are estimated to cause 1.22 [0.65, 2.12] *95CI* more wall contacts in the second task compared to the first, followed by 0.51 [0.36, 0.70] *95CI* for the third and 0.74 [0.53, 1.01] *95CI* for the egocentric task. Considering the upper and lower boundaries of the confidence intervals, estimates are rather uncertain.

These estimations on the population level make two points visible. First, the estimation for causing wall contacts is huge and predicted with good certainty. Second, the estimation for causing wall contacts decreases from the first until the third allocentric sub-task. Third, the difference in estimated wall contacts is lower for the egocentric, compared to the allocentric task.

*Table 3*. *F*ixed-effects for wall contacts on population-level for the different tasks.

| | Population | |
|---|---|---|
| Wall Contacts | Effect Estimate | 95% CI |
| ToT | 1.77 | [1.55, 2.06] |
| Condition: | | |
| Allocentric_1 | 1.25 | [0.83, 1.82] |
| Allocentric_2 | 1.22 | [0.65, 2.12] |
| Allocentric_3 | 0.51 | [0.36, 0.70] |
| Egocentric | 0.74 | [0.53, 1.01] |

***Visualization of conditions: Population level.***

For the analysis of the differences between the conditions we plotted the wall contacts differentiated by perspective and by sub-task and the ToT scores differentiated by perspective and sub-task on population levels. Participants caused less wall contacts in the egocentric compared to the allocentric condition (Figure 21). The number of wall contacts decreased from the first until the last sub-task in the allocentric condition (Figure 22). Participants were faster in the allocentric compared to the egocentric condition (ToT) (Figure 23). ToT decreased from the first until the third sub-task of the allocentric condition (Figure 24).

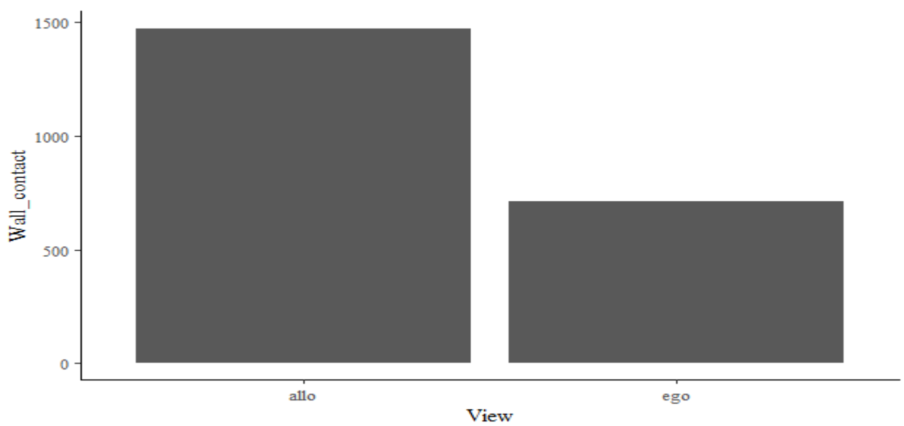



*Figure 21*. Comparison of wall contacts between conditions (view) on population level.

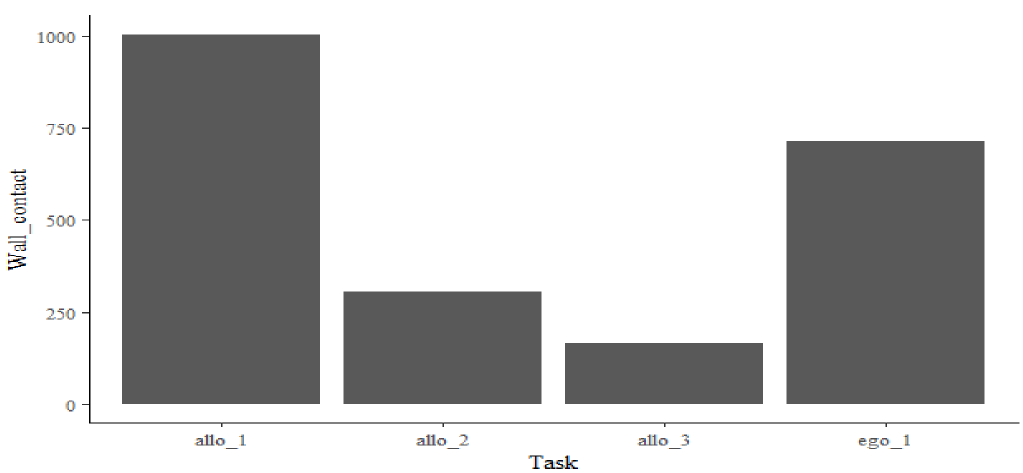



*Figure 22*. Comparison of wall contacts between sub-tasks of conditions on population level.

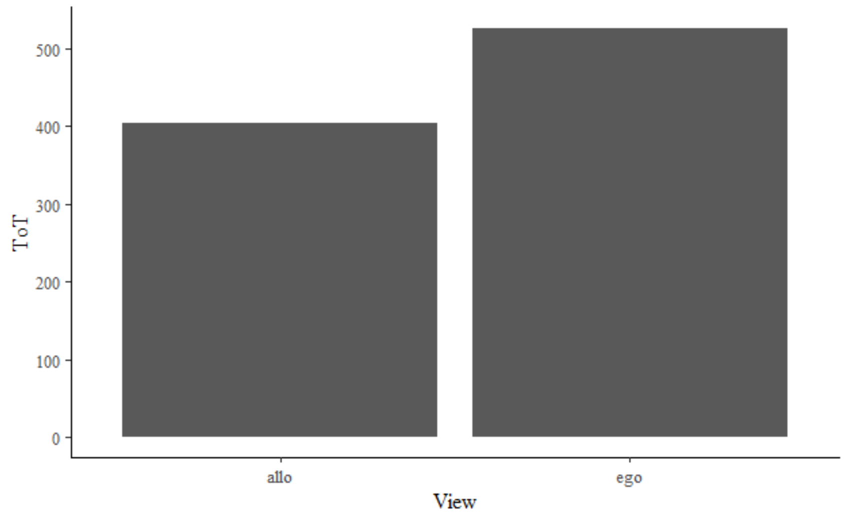



*Figure 23*. Comparison of ToT between conditions (view) on population level.

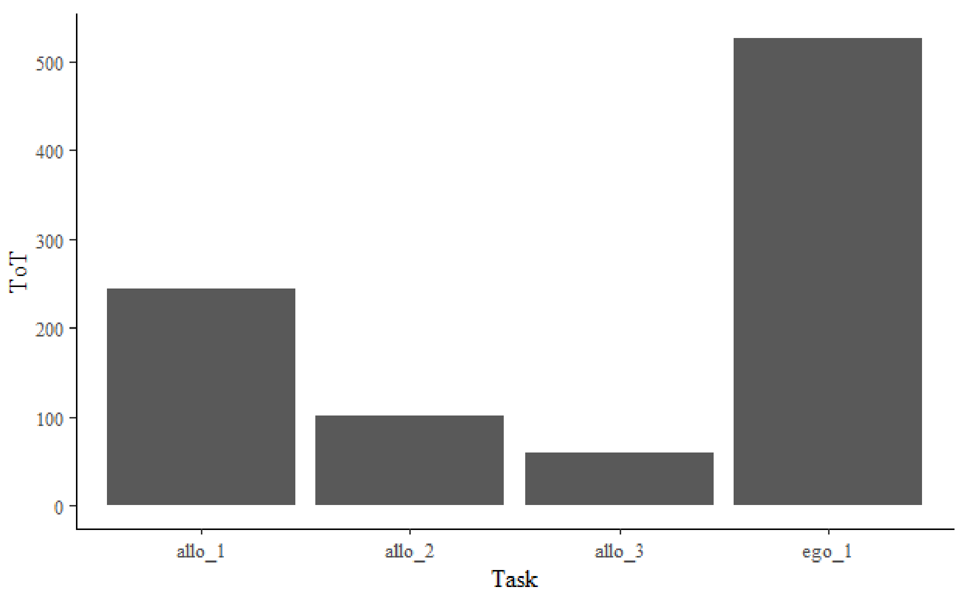



*Figure 24*. Comparison of ToT between sub-tasks of conditions on population level.

***Visualization of conditions: Individual level***. Individual differences in wall contacts and ToT were visualized for the allocentric and the egocentric condition (view). Most participants caused visibly less wall contacts in the egocentric compared to the allocentric condition (Figure 25-26). Most participants were faster in the allocentric compared to the egocentric condition (Figure 27).


*Figure 25*. Comparison of individual wall contacts between conditions (views).

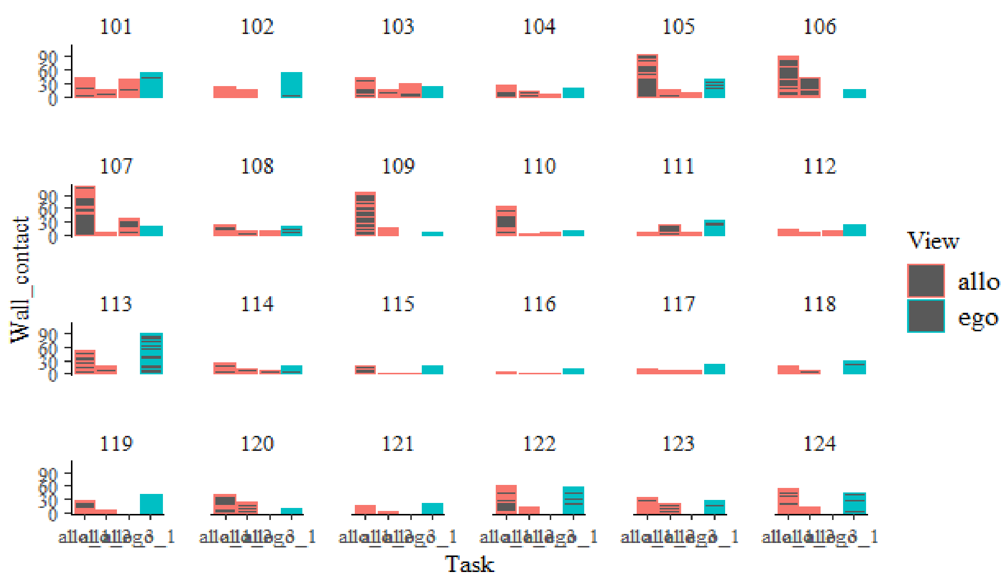



*Figure 26*. Comparison of individual wall contacts between sub-tasks in conditions (views).


*Figure 27*. Comparison of individual ToT between conditions (views).

**Interaction effect: Route by tasks**. In order to investigate the performance difference between the tasks, we included the difficulty parameter routes and participants as multiple factors to the model. We found an interaction effect of route by task.. Coefficients are provided in Table 4. The coefficients show variation in how the routes influence wall contacts conditional on tasks.

The mean estimate for the effect of the difficulty of the routes on the number of wall contacts is 0.65, ranging from 0.46 on the lower until 0.94 on the upper bound, indicating no good certainty. We see that the effect of the routes on the wall contacts is highest for the first task of the allocentric condition (0.65) second for second allocentric task (0.52), followed by the egocentric task (0.46) and the third allocentric task (0.41). The interval boundaries for the second and the third allocentric task are very big, which indicates great uncertainty of the effects. The certainty for the effect of route by the egocentric task is better, but not convincing.

Comparing the association between numbers of wall contacts by ToT with the estimates of Table 3 without the interaction effect proves strong variance, but predictions are not convincing in certainty. We visualized the interaction effect in a summarized form for the condition perspective instead of sub-tasks (see Figure 28).

*Table 4.* Interaction effect route by task for wall contacts.

| | Interaction: Route | |
|---|---|---|
| Wall Contacts | Effect Estimate | 95% CI |
| Condition: | | |
| Allocentric_1 | 0.65 | [0.45, 0.94] |
| Allocentric_2 | 0.52 | [0.09, 1.20] |
| Allocentric_3 | 0.41 | [0.05, 0.89] |
| Egocentric | 0.46 | [0.20, 0.79] |

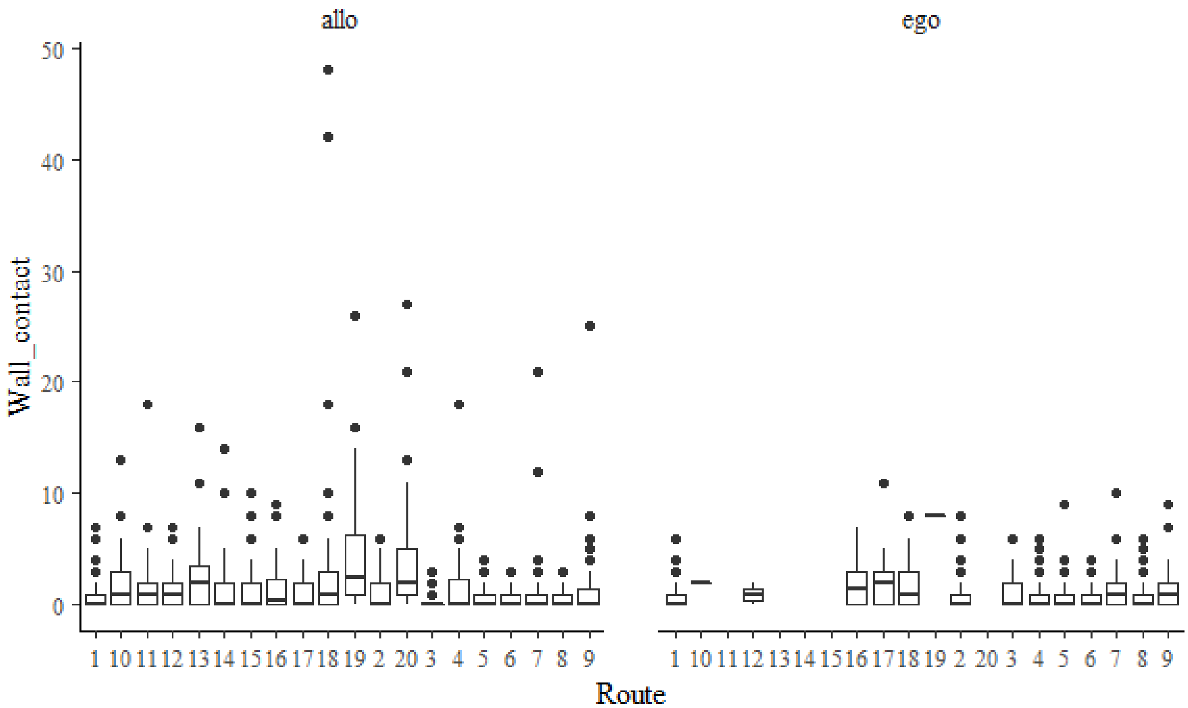



*Figure 28.* *Association of routes and wall contacts conditional on perspective.*

**Qualitative analyses: EndoProto Expert Evaluation**

The following qualitative results are based on personal communication with two experts in bronchoscopy. Tabulated excerpts of interview can be found in Appendix I. The selected qualitative insights shall (1) help to evaluate the possibilities and limitations of the EndoProto in relation to the resemblance spectrum and (2) help to get insights why based on our results no learning took place.

**Composition and proportion of material.** The interviewed experts evaluated the EndoProto as and *imaginable* as basic instrument for introducing the *simplest primary stage of a real endoscopy* (Expert 1, personal communication, 23th of June 2018), but not as a realistic endoscopy. Based on the features of the material, the box was evaluated as *fitting* for being a rather *rigid system* (Expert 1, personal communication, 23th of June 2018; Expert 2, personal communication, 2nd of July 2018), which would be comparable to real bronchia. The dimensions of the openings (diameter of the routes) were not perfectly, but nearly realistic. Expert 1 explained that the *real size in a bronchoscopy lies between the two of the walls*. The size of the provisional endoscope was evaluated by both experts as broadly fitting to small realistic endoscopes, slightly changing in size. Expert 2 explained that our endoscope would be comparable in size to the one *for children.*

**Endoscopic tasks**. Both experts interpreted the EndoProto training as imaginable as a first hands-on training for endoscopic skills. In relation to the proportions of the routes, expert 2 uttered, that *first training with the broad ones and then with the thin ones could be appropriate*. Expert 1 emphasized the relevance of the endoscopic camera-view, because to *look at a scree from the start, is was has to be done finally.*

**Performance parameters**. Both experts evaluated the performance parameters ToT and wall contacts as understandable for measuring performance, but less necessary in a realistic bronchoscopy. Time was regarded as less relevant, because the only factors limiting the time in a real bronchoscopy would be the operating team and the anesthesia (Expert 1). Wall contacts were regarded as less relevant in a real bronchoscopy, because it would be *nearly not possible to harm a patient in a bronchoscopy, because you do not pierce through the human tissue as for example in a coloscopy (Expert 2*). A similar utterance according to the compared risk in a coloscopy was made by the Expert 1, as he said that *bleedings are possible, but not dramatic. Not as for example the piercing through in coloscopy*.

**Main findings**. Expert 1 (personal communication, 23th of June 2018) summarized his viewpoint in the following citation: *For a first hands-on, such a thing would be not as bad, if it would be better operable; if it would at least offer approximately realistic conditions in relation to handling and the sharpness of the picture, of course*. His main critique on the EndoProto was the lack in sharpness of the camera. He concluded first, that this would not be the case with a *normal endoscope* and second, that this limited sharpness would lead to the problem that the routes in the EndoProto tasks were nearly not *accessible*.

Expert 2 (personal communication, 2nd of July 2018) emphasized two aspects, which would restrict the possibilities to learn with the EndoProto. First, the missing overview of the routes on the dividing walls, based on the space between them. The expert recommended to *make the openings thinner, more dense to each other in order to be able to follow plenty of directions; some bigger and some smaller openings.* Second, the expert evaluated the missing suturing system of the endoscope as *cumbersome,* because of the restriction in possible movements. On the contrary he also mentioned the strength of the endoscope, to be *good usable for insertions.*

# Discussion

## Answering research questions

Due to the fact, that we were not able to estimate learning curves with the rate parameter trial and the performance parameter ToT, we cannot answer our original research questions adequately. Nevertheless, we adjusted our focus to this situation and generated post hoc research questions, which we answer in the following.

1. How do individuals differ in the performance parameters? We can conclude that participants did not improve their performance levels, not regarding ToT and not for wall contacts. In other words, practically nobody proved to be a visible learner. Although we found individual variation for the performance parameters, this was not overwhelming.
2. How do participants differ under the circumstance of perspective (allocentric and egocentric)? On group and individual levels participants performed better in the allocentric compared to the egocentric condition. This result first proves our expectations and experience with pilot tests, that the egocentric task is more demanding than the allocentric task. More important, this finding replicates previous research about performance in MIS. The reason for achieving better task performance in an allocentric condition is related to the fact that the view of the own hands leads to a better hand-eye coordination (Rappel, Lahiri, & Teo, 2013).
3. What can the opinions of experts add to the current study? The personal communication with two specialists revealed some drawbacks of the EndoProto, which might have restricted participants in showing learning effects: (1) the missing control system for the endoscope (2) and the missing adjustment of camera-sharpness in close distances (Expert 1) which can be related to the *missing overview* due to the distance between the routes (Expert 2).

**Discussion of learning**

In addition to the presented results, we found an interaction effect of routes, which influenced wall contacts conditional on tasks.  We interpret this interaction effect of the difficulty parameter routes as one possible reason why no learning curves could be estimated. The routes 16-18 caused the highest variance in the egocentric condition. These routes had the greatest distance to the starting point (yellow route =1) and thus led to very small angle for the camera view. Combined to the missing control system for the endoscope these small angles also caused a visibly higher ToT (see Appendix H – R Syntax). Moreover, exchangeable trials helped us to investigate this assumption in mixed-effects multi-levels, which made the investigation of individuals more detailed. Exchangeable trials also resulted in the named interaction effect and investigate several independent questions as for example the correlation between the performance parameters ToT and wall contacts, which hint at possible speed-accuracy tradeoffs. Further analyses (see Appendix H – R Syntax) reveled that participants in our study also showed this effect – to make more wall contacts when having shorter ToTs and the reversed case.

**Requirements**. Regarding our five design requirements, the EndoProto proved to be low-budget and practically usable (requirement 5. economics). According to MIS dexterity we were not able to show that participants constant and visible learning (requirement 3. Learning and performance) (based on learning curves or the association between ToT and wall contacts, which might still replicate that the trained basic scope insertion task might be based on innate abilities (requirement 2. Surgical dexterity). But, we were able to show some possible tendencies towards learning (see skill acquisition). The final validation EndoProto still needs to be proven, especially in accordance to VR-simulator training (requirement 2) and the lacks in resemblance (requirement 1) as discussed further in the following.

**Latent variables**

We proved that participants made no learning progress based on learning curves (with the amplitude ToT) and also not for wall contacts, which were positively association to ToT. Nevertheless, these results do not prove that participants did not learn at all, because a complex phenomenon as learning cannot be covered in probably any study. Therefore we will now discuss possible reasons and latent influences for the lack of learning effects based on ToT and wall contacts.

**Error recovery.** The data analysis revealed a positive relation between ToT and wall contacts, based on good certainty. This result raises the question, if the correlation is the result of error recovery through time. Based on observations and pilot testing, wall contacts should not have caused a delay in task completion. Participants were not restricted in completing their trial, if they touched the walls. Nevertheless, they had the opportunity to go back to the starting point to re-orientate. In an overall impression, this ‘complete extraction’ until the starting point happened very rarely.

In comparison to this, nearly everybody took additional time to pull back the endoscope to re-orientate in egocentric. Reasons for this action were for example that the focus of the camera became too diffuse in a close distance beneath 2-3 cm to a wall. Another example is the loss of orientation, due to wrong maneuvering. As the first aspect can be regarded as a technical restriction, the second one is grounded on individual skills and the comprehension of the connection between own movements and resulting ones of the endoscope tip.

An additional aspect of error recovery was, that especially in egocentric, most participants apparently did not even notice if they touched a wall. This assumption is based on personal communication and observations. Visible reactions and movements of participants varied strongly between the conditions. Participants generally tended to move slower in egocentric task, which was also proven by our results. Especially at the beginning when they needed to adjust to the new task features. People obviously tried to learn the association they were not seeing in front of them anymore – the relation of the movement speed of the endoscope and the distance of the tip of the endoscope and the destination (crossing a route until reaching the red feedback plate beneath it). This may not only have caused a higher workload for the new task. Concluding, this combination of technical hindrances and personal lacks in performance can maybe broadly summarized as effect as an time delay through error recovery – a latent variable negatively influencing performance progresses related to wall contacts and ToT.

**Disorientation.** As an aspect described in innate spatial abilities and navigation – this is important for endoscopy and bronchoscopy and still one of the main problems also found in literature (add ref). Similar results have been proven partly here (1) qualitatively visible through comments and general impressions from participants (2) quantitatively visible through the investigation of measured disorientation with the precision variable. Errors in precision only occurred in the egocentric task. Analyzed descriptive show, that 17 participants caused at least one time a false precision and lost orientation. Precision errors could be evaluated as problems in orienting. This may also have caused further time delay – could be related to error recovery and probably to the technical complication of bad focus – which may have also induced more errors.

**Critical Reflection of Study**

**Degrees of freedom**. One of the main differences between the EndoProto tasks and the BRONCH Mentor simulator tasks, are the degrees of freedom in hand and body movements. Around 30 degrees of freedom are reported to be used by surgeons for the coordination of their hands related to the space in which they operate (Rappel, Lahiri, & Teo, 2013). Allbows are most of the time tacked in, the controlling hand is only turning slightly and the upper torso is mostly fixed, as also approved and demonstrated in real life by the interviewd Experts (Expert 1: *only moving with the thumbs and the hand rotation*; Expert 2: The *ergonomics…to remain standing and not turn to what is visible on the screen*.). In comparison to a simulator task or a real surgery, in out EndoProto tasks, participants did not tuck the elbows. The more important difference are the finger movements, which are not related to a control system as for a real bronchoscope. Two of the most observed actions of the hands and fingers were (1) to push the endoscope (deeper) inside the EndoProto and (2) to use very small scaled finger movements to turn the wire of the endoscope near the starting point of the related tasks. Concluding, the physical parameters of the hand and body movements of participants did not transfer to a real simulator task and until now represent a strong lack in resemblance.

**Technical inaccuracy**. Qualitative results already indicated that the limited sharpness of the camera is problematic for the execution of basic scope insertion tasks. This problem is related to a very close distance of beneath 2-3 cm in front of an object. At this point the sharpness of the visualization decreases, which probably fostered a loss of orientation for participants and thus induced the attempts to pull back the endoscope to get a better view. Simultaneously this may have caused a severe time delay in trial and task completion. Our results also showed that participants took more time for the trials in the egocentric condition compared to the trials in the allocentric condition, which might not only be based on the difference in perspective, but also in the limited sharpness of the camera. Moreover, previous research findings clarified that disorientation would be one of the most occurring causes for errors and restriction for success in MIS (Taylor, Brunyé, & Taylor, 2008; Cao & Milgram, 2000). The reason for this problem is the insecurity about the exact location of the endoscope's tip, which leads to problems in navigation. Based on our own observations we can conclude, that the missing sharpness adjustment of the camera probably led to a time delay in trial completion and also to increased rates of wall contacts in the egocentric task.

**Inter-rater reliability**. Related to the performance parameters ToT and wall contacts we cannot be absolutely certain that my colleague and I measured and counted the performance parameters in exactly the same way. Finally, a human performance measurement will always stay subjective, at least to a certain extent, compared to more reliable and objective measurements as the ones executed by a machine as the BRONCH mentor simulator. Nevertheless, it shall be stated that we aimed to achieve a high degree of reliability through adjusting the following steps: (1) Personal communication before and after participant supervision, (2) the construction of a research protocol for higher similarity in procedure explanation and supervision and (3) measuring the performance parameters of the pilot test together. In cases of different scores in wall contacts, we calculated a mean.

**Skill acquisition: Qualitative insights**

Since, the current study deals with unexperienced users, rather the first two phases of skill acquisition (Mohamed, Raman, Mclaughlin, Anderson, & Coderr, 2010) were expected to take place. Personal communication and observations helped to get small insights in the phases of unconscious and conscious incompetence and supplemented findings on previous research. Mohamed et al. (2010) investigated upper and lower gastrointestinal endoscopies of novices and reported highest amounts of workload for the first 10 (upper) and the first 40 (lower) procedures of endoscopy. The authors also found a later increase in workload, representative for the learning step from unconscious to conscious incompetence. It was discussed that trainees probably became more aware of their own mistakes and got a deeper understanding of the working mechanisms of the endoscopes when transferring from the unconscious to the conscious incompetence.

**Confirming awareness**. Comparable observations and qualitative aspects of personal communication with participants revealed likewise insights. Although it has to be stated, that we did not explicitly investigate workload or conscious awareness in the current study. Nevertheless, plenty of participants uttered concerns about their high amount of wall contacts. We were asked more often, something like: *Does everybody make so many mistakes?* Such paraphrases make visible that participants were cognitively busy with their error rates. Some participants made the impression of acquiring knowledge about the mobility of the endoscope-tip in accordance to own hand movements. Some participants indicated either indirectly through their behavior or directly through utterances (example: *One has to know first, that if I turn here, the endoscope may turn in a different direction.*) that they became more aware of their failures and the consequences of their movements for completing a route successfully. One example is the awareness of the relation between shivering and increased wall contacts (for example: *I am shivering too much for this*).

**Disconfirming awareness**. On the contrary, through observations it became visible, that some participants seemed to be rather unaware of the consequences of their finger or hand movements for the resulting movements of the endoscope-tip and camera. Several cases were observed which led to repeated failures of the same kind, when the endoscope was just pushed forward, without taking into account that the endoscope-tip generally was moving slightly down with a greater distance to the starting position, because of gravity and the lack of rigidity of the endoscope-cable. While every participant got the same instruction (*Try to be as precise and fast as possible*), the observed speed of their actions varied tremendously. This insight rather supports the statistical results of missing learning effects.

**Self-reflection**. Finally, these two contrary observations demonstrate the variability of participants and their first impressions for possible unconscious or conscious insights in their own mistakes and understanding of the used tool (provisional endoscope). The interesting question now is, if these different reflections of the participants are primarily based on innate visual-spatial abilities or if they can also give information of learned skills. This question cannot be answered here. Instead, we can at least conclude that learning experience generally is triggered by critical self-reflection (Mezirow, 1990). Self-reflection is a metacognitive state (May & Etkina, 2002), which is used in learning or problem solving, where the answer is not clear. Related to the current study, participants were confronted with a new task and problem in the endoscopic condition. That at least some of them made use of verbal self-reflection demonstrated possible tendencies towards learning. Possible future investigations with the EndoProto could offer more insights in the aspect of cognition through conscious awareness of own actions and workload. For more participant impressions we refer to Appendix G – observation protocol.

**Towards validation: Task complexity and performance**

Every participant evaluated the egocentric task as more complex than the allocentric task/s. This proves the expected and requested effect of the task sequence, which was initiated to rise in complexity (comparison between allocentric and egocentric in total) in accordance to the increased resemblance of basic bronchoscopic simulator tasks. The primary interest was not to create a tool, which can replace professional simulators and related tasks. Instead, our interest was to develop a low-fidelity prototype which can supplement professional simulator training in an early stage of skill acquisition. For this purpose the difference in complexity and combined higher rate of resemblance were necessary, in order to approach the high requirements which are constantly related to surgical training and trainees (Gallagher, Leonard, & Traynor, 2009). Related to this goal, qualitative valuations of participants could indicate a tendency of task complexity and serve as one step towards validating our tasks.

Moreover, the quantitative task performance based on reported results can even give a stronger tendency towards possible directions for learning. We also proved no learning effects regarding learning curves and the correlation between ToT and wall contacts (see Appendix H – R syntax). Nevertheless we could also show group effects for performance increases in the different sub-tasks in the allocentric condition. Participants tended to decrease the amount of wall contacts and necessary ToT from task 1 until task 3, which might be an indicator for learning.

Another aspect is the proven difference in task complexity, comparing the ToT and wall contact scores for the allocentric and the egocentric task. As expected, participants needed more time for the egocentric task, but caused less wall contacts in this condition. We assume that this difference in performance (wall contacts) is less attributed to a learning effect, but more to the difference in task complexity, which can probably be ascribed to the more demanding hand-eye-coordination due to the change in perspective. Finally, this chance in condition probably influenced the higher time needed to complete the trials. Another indicator proving our previous assumptions about the task complexity is the number of repetitions. Participants had nearly the same time-frame for the allocentric and the egocentric task (25 minutes), but generally reached fewer repetitions in the last mentioned task.

In relation to separate subtasks, participants made fewest wall contacts when they conducted task 1.3 (see Figure 26). This means also, that these participants were fast enough to conduct task 1.1-1.2 in a certain time frame, to also be able to start the additional task 1.3 at all. Although the data is very uncertain, this estimation may still be an indicator for possible learning tendencies. The reason for this argument is first, the decreasing estimation for causing wall contacts from task 1.1 until task 1.3. and second, the strong difference of causing wall contacts in task 1.3 compared to 1.1 (see also Table 3). Further, the rise in the estimation for causing wall contacts relative to ToT between task 1.3 (0.51) and the following task 2 (0.74) might also contribute to the previous argumentation about skill acquisition and workload. It is not astonishing that the error rate increased in a more demanding task, because of the different amounts of cognitive load probably related to it. Finally, this increase in estimated wall contacts also contributes to the findings above.

The argumentation above about the differences in estimated wall contacts between the two task conditions gives further indications for the resemblance spectrum and real bronchoscopies. Causing less wall contacts as possible probably is one essential goal in real bronchoscopy, which even may or must be compensated with time. And, according to the interviewed experts time is rather regarded as less relevant in a real bronchoscopy (Expert 1, personal communication, 2018, June 23). Therefore, error prevention may even be regarded as more important for MIS training, than a fast task execution. Concluding, there are tendencies which seem to give indications for validating the task complexity between the allocentric and the egocentric conditions. Nevertheless, discussed degrees of freedom and technical inaccuracy still make the EndoProto less resembling a professional simulator tasks.

## Outline for Future Research

Based on statistical reasoning no learning curves could be estimated, but this does not necessarily mean that no learning took place at all. There is still a possibility for unknown and not investigated variables, which could indicate signs of learning. As the original focus of the study was hand-eye-coordination, the planned Arduino sensor glove would still offer an opportunity to find out more about hand movements. The bronchoscopy simulator was definitely a missing piece in the procedure of skill acquisition in bronchoscopy. Under the aspects of missing learning effects with the EndoProto, a deeper investigation of basic scope insertion tasks with the glove on the box trainer and the simulator tasks would be interesting. Such a procedure could help to get a deeper understanding of the possible commonalities and differences between the box trainer and the BRONCH Mentor. For possible future adjustments and validation procedures in relation to the resemblance spectrum, we tabulated the main comparisons and differentiations between the EndoProto, the VR-simulator (BRONCH Mentor) and a real bronchoscopy in Appendix J.

**Feedback.** Although feedback converting was no focus of the current study, we tried to cover simplistic aspects, as for example sounds when touching a wall. For a better adjustment to the planned BRONCH Mentor task 1 – an additional red coloring of the internal walls of the box may have given a closer impression for approaching a wall. Another interesting aspect of giving feedback for wrong actions, could be a linking with a sensor glove as the originally planned

Arduino. If such a glove could be equipped or programmed to give a non-harmful impulse as haptic feedback or induce a disturbing noise when touching a wall, this might influence a participants actions and maybe foster learning effects. Fast moving participants would maybe get slower and more careful. Otherwise, a repeated noise can also increase the workload and influence the amplitude ToT.

## Conclusion

No learning curves could be estimated, not on a population nor on a participant level. Participants visually did not learn to improve their ToT nor did they learn to decrease their rate of wall contact over the session. Although individual differences were available, practically no participant proved to be a visible learner, according to our quantitative analyses.

We were able to supplement previous findings about performance differences between MIS tasks in allocentric and egocentric perspectives. An interaction effect proved the strong influence of how routes affected wall contacts of participants conditional to tasks. We interpret this interaction effect of the difficulty parameter routes as one possible reason why no learning curves could be estimated. Moreover, exchangeable trials helped us to investigate this assumption. Another advantage of exchangeable trials was the possibility to investigate several independent questions as for example the correlation between the performance parameters ToT and wall contacts, which hint at possible speed-accuracy tradeoffs.

Finally, qualitative investigations through observations and expert interviews helped to get insights in possible qualitative aspects of learning, as for example self-reflection. On the contrary the strong movement degrees of freedom and the technical inaccuracy of the endoscope revealed current lacks in resemblance compared to professional VR-simulator tasks. Further, expert interviews helped us to evaluate and adjust the potential of the EndoProto in relation to the resemblance spectrum (Appendix J) for possible future investigations. The validation of the EndoProto still needs to be proven. We advise to repeat the current study, completed through the original focus and adjusted to the reported limitations and results, in order to further evaluate the possibilities of the EndoProto for training endoscopic and bronchoscopic skills.

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin,, 102*(1), 3-27.

Alvarado, C. J., & Reichelderfer, M. (2000). APIC guideline for infection prevention and control in flexible endoscopy. *American Journal of Infection Control, 28*(2), 138-155.

Anastakis, D. J., Hamstra, S. J., & Matsumoto, E. D. (2000). Visual-spatial abilities in surgical training. *The American Journal of Surgery, 179*(6), 469–471.

Arendt, A. (2017). Towards reliable and valid prediction of MIS-performance with basic laparoscopic tasks in the LapSim and low-fi dexterity tasks. University of Twente: Master's thesis.

Basdogan, C., Sedef, M., Harders, M., & Wesarg, S. (2007). VR-based simulators for training in minimally invasive surgery. *IEEE Computer Graphics and Applications, 27*(2), 54-66.

Buser, M., Pouw, H., & Kusters, R. (2016-2017). *Pre-oral endoscopic myotomy: Analysis of the scientific relevance and training methods of this upcoming advanced endoscopic procedure.* MDO . Enschede: University of Twente.

Colt, H. G., Crawford, S. W., & Galbraith , O. (2001). Virtual reality bronchoscopy simulation: A revolution in procedural training. *CHEST, 120*(4), 1333-1339.

Cuschieri, A., Francis, N., Crosby, J., & Hanna, G. B. (2001). What do master surgeons think of surgical competence and revalidation? *The American Journal of Surgery, 182*(2), 110-116.

Darzi, A., & Munz, Y. (2004). The impact of minimally invasive surgical techniques. *Annual Review of Medicine, 55*, 223–37.

Degani, A., Choset, H., Zubiate, B., Ota, T., & Zenati, M. (2008, Aug). A highly articulated robotic probe for minimally invasive surgery. *Conf Proc IEEE Eng Med Biol Soc, 2006*(1642343), pp. 4167–4172.

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin, 53*(2), 134-140.

Fitts, P. M., & Posner, M. I. (1967). *Human Performance.* Oxford: Brooks/Cole.

Gallagher, A. G., & Smith, D. D. (2003). Human-factors lessons learned from the minimally invasive surgery revolution. *Seminars in Laparoscopic Surgery, 10*(3), 127-139.

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*(3), 108-113.

Gmeiner, D., von Rahden, B. H., Meco, D., Hutter, J., Oberascher, G., & Stein, H. J. (2007). Flexible versus rigid endoscopy for treatment of foreign body impaction in the esophagus. *Surgical Endoscopy, 21*, 2026–2029.

Groenier, M., Schmettow, M., Huijser, S., & Gallagher, A. G. (n.d.). *Predicting laparoscopic skill acquisition on a simulator.* Unpublished draft.

Gronemeier, B. (2018, June 23). Personal Communication.

Haluck, R. S., Webster, R. W., Snyder, A. J., Melkonian , M. G., Mohler, B. J., Dise, M. L., & Lefever, A. (2001). A virtual reality surgical trainer for navigation in laparoscopic surgery. In W. J. al., *Medicine Meets Virtual Reality* (pp. 171-176). IOS Press.

Heathcote, A., Brown, S., & Mewhort, D. K. (2000). The power law repealed: The case for an exponential law of practice. . *Psychonomic bulletin & review, 7*(2), 185-207.

Jones, L. V., & Thissen, D. (2006). 1 A History and Overview of Psychometrics. *Handbook of statistics*, 1-27.

Keehner, 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: The Official Journal of the Society for Applied Research in Memory and Cognition, 20*(4), 487-503.

Kemp, S. V., El Batrawy, S. H., Harrison, R. N., Skwarski, K., Munavvar, M., Roselli, A., . . . Sha, P. L. (2010). Learning curves for endobronchial ultrasound using cusum analysis. *Thorax, 65*, 534-538.

Kennedy, C. C., Maldonado, F., & Cook, D. A. (2013). Simulation-based bronchoscopy training: systematic review and meta-analysis. Chest, 144(1), 183-192. *Chest, 144*(1), 183-192.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In Spatial cognition. In H. C. Freksa C., *Lecture Notes in Artificial intelligence: Spatial Cognition* (Vol. 1404, pp. 1-17). Berlin, Heidelberg: Springer.

Kovaleva, J., Peters, F. M., van der Mei, H. C., & Degener, J. (2013). Transmission of infection by flexible gastrointestinal endoscopy and bronchoscopy. *Clinical Microbiology Reviews, 26*(2), 231–254.

Kramp, K. H., van Det, M. J., Hoff, C., Veeger, N. G., ten Cate Hoedemaker, H. O., & Pierie, J. P. (2016). The predictive value of aptitude assessment in laparoscopic surgery: A meta-analysis. *Medical Education, 50(4), 409–427., 50*(4), 409–427.

Lechtzin, N., Rubin, H. R., White, P., Jenckes, M., & Diette, G. B. (2002). Patient satisfaction with bronchoscopy. *American journal of respiratory and critical care medicine, 166*(10), 1326-1331.

Lin, C. K., Lai, C. L., Chang , L. Y., Wen, Y. F., & Ho, C. C. (2018). Learning curve and advantages of endobronchial ultrasound - guided transbronchial needle aspiration as a first - line diagnostic and staging procedure. *Thoracic cancer, 9*(1), 75-82.

Mason, R. A. (1992). Learning fibreoptic intubation: fundamental problems. *Anaesthesia, 47*(9), 729-731.

May, D. B., & Etkina, E. (2002). College physics students' epistemological self-reflection and its relationship to conceptual learning. *American Journal of Physics, 70*(12), 1249-1258.

Medford, A. R. (2008). *Impact of the European Working Time Directive on speciality training.*

Medford, A. R. (2013). Learning curves for bronchoscopy and simulation. . *Clinical Medicine, 13*(4), 418-419.

Mezirow, J. (1990). How critical reflection triggers transformative learning. *Fostering critical reflection in adulthood, 1*, 1-20.

Mohamed, R., Raman, M., Mclaughlin, K., Anderson, J., & Coderr, S. (2010). Transition From Novice to Expert Endoscopist: Mapping the Evolution of Endoscopic Skills Using Workload Assessment. *Gastrointestinal Endoscopy, 71*(5), AB273.

Naur, T. H., Konge, L., Nayahangan, L. J., & Clementsen, P. F. (2017). Training and certification in endobronchial ultrasound-guided transbronchial needle aspiration. *Journal of thoracic disease, 9*(7), 2118-2123.

Ohuchida, K., Kenmotsu, E. H., Yamamoto , E. A., Sawada, E. K., Takasugi, E. K., Morooka, E. K., . . . Hashizume , E. M. (2009). The frontal cortex is activated during learning of endoscopic procedures. *Surgical endoscopy, 23*(10), 2296-2301.

Ost, D., De Rosiers, A., Britt, , E. J., Fein, A. M., Lesser, M. L., & Mehta, A. C. (2001). Assessment of a bronchoscopy simulator. *American journal of respiratory and critical care medicine, 164*(12), 2248-2255.

Sangani, H., Bradley, K. A., & Fidopiastis, C. (n.d.). *The Effectiveness of Training Mental Rotation and Laparoscopic Surgical Skills in a Virtual Environment.*

Schmettow, M. (in prep.). *New statistics for the design researcher.* Accessed via: https://bookdown.org/schmettow/new_statistics_for_the_design_researcher/.

Schmettow, M., & Groenier, M. (2017, 3 14). Towards simulator-based selection for minimally invasive surgery: Research framework and first results. Enschede, Netherlands.

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal,, V. K., Andersen, D. K., & Satava, R. M. (2002). Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Annals of surgery, 236*(4), 458-464.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359-1366.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. Presence: Teleoperators & Virtual Environments, 6(6), 603-616. *Teleoperators & Virtual Environments, 6*(6), 603-616.

Srour, F. J., Kiomjian, D., & Srour, I. M. (2018). Automating the Use of Learning Curve Models in Construction Task Duration Estimates. *Journal of Construction Engineering and Management, 144*(7), 04018055-1-14.

Srour, J. F., Kiomjian, D., & Srour, I. M. (2015). Learning curves in construction: A critical review and new model. Journal of Construction Engineering and Management, 142(4), 06015004. *Journal of Construction Engineering and Management, 142*(4), 06015004-1-5.

Stather, D. R., MacEachern, P., Rimmer, K., Hergott , C. A., & Tremblay, A. (2011). Validation of an endobronchial ultrasound simulator: differentiating operator skill level. Respiration. *Respiration, 81*(4), 325-332.

Sun, J., Xie, F., Zheng, X., Jiang, Y., & Zhu, L. M. (2017). Learning curve of electromagnetic navigation bronchoscopy for diagnosing peripheral pulmonary nodules in a single institution. Translational Cancer Research, 6(3), 541-551. *Translational Cancer Research, 6*(3), 541-551.

Taylor, H. A., Brunyé, T. T., & Taylor, S. T. (2008). Spatial mental representation: implications for navigation system design. *Reviews of human factors and ergonomics, 4*(1), 1-40.

Tendick, F., & Cavusoglu, M. C. (1997, Oct.). Human-machine interfaces for minimally invasive surgery. *In Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE, 6*, pp. 2771-2776.

Truong, S., Böhm, G., Klinge, U., Stumpf, M., & Schumelick, V. (2004). Results after endoscopic treatment of postoperative upper gastrointestinal fistulas and leaks using combined Vicryl plug and fibrin glue. *Surgical Endoscopy and other Interventional Techniques., 18*(7), 1105-1108.

Wanzel, K. R., Hamstra, S. J., Caminiti, M. F., Anastakis, D. J., Grober, E. D., & Reznick, R. K. (2003). Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. *Surgery, 134*(5), 750-757.

Westerhof, M. (2018). *Individual differences in MIS: Towards predicting bronchoscopic skill acquisition using basic bronchoscopic simulator tasks.* Enschede: University of Twente.

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? *Trends in cognitive sciences, 14*(3), 138-146.

## Appendix A - Supplementary method for session II: Simulator training

### Design

Based on reasons of time and availability of the simulator, the first 6 participants completed a simulator session of one hour.

### Procedure

**Second session – Simulator task.** In this session participants trained on the BRONCH Mentor (see Figure A1), a professional simulator for the domain of bronchoscopy. The simulator was made available by the ECTM of the University of Twente. It is produced and shared by 3D Systems, a global team which supports healthcare development through technology for professional training methods of medical staff. It offers a wide range of realistic training procedures for bronchoscopic skills, as well as educational information and guidelines.

For the current study, the video tutorials *posture and scope maneuvering* and *bronchoscopic navigation fundamentals* were demonstrated to every participant, before the start of the session. They explained the right posture and handling of the endoscope. Together these took around 5 minutes. Participants got another moment to test their new movements themselves. If necessary, the researcher again demonstrated and helped with the right posture and handling of the endoscope. For more details of the procedure we refer to the created guidelines for session 2 on the following pages.

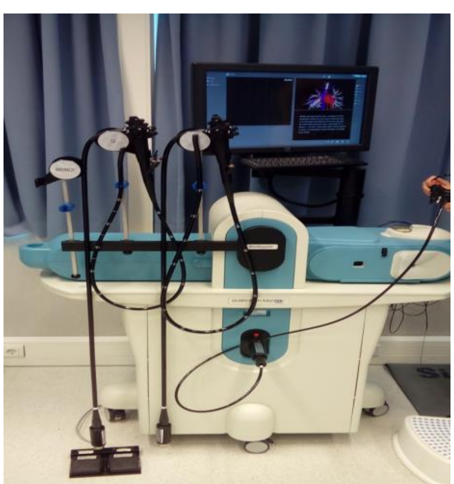


*Figure A1.* BRONCH Mentor simulator for executing bronchoscopy training. Own photo.

***Basic Scope Manipulation.*** As a first task from essential bronchoscopy, the basic scope manipulation was conducted. This task is designed for new users and people who train to acquire bronchoscopic skills. It was developed in cooperation with the American Association for Bronchoscopy and Interventional Pulmonology (AABIP), (3D Systems, 2017).

In this task participants basically trained to move the tip of the endoscope as close as possible to a blue ball moving through the center of a long tunnel, representing the human bronchia in a science fiction like setting (see Figure A2). If participants were able to keep the movement as close to the ball as possible, they prevented wall contacts and got better results. If they got too near to a wall the screen turned red and if they touched the wall additionally a loud crashing noise appeared and the screen blinked dark red for a moment. Participants were induced to follow the blue ball until the end of the route/bronchia. This task was repeated 15 times. Routes were generated randomly by the system and were 8 different ones in total.

*Figure* A2. Essential bronchoscopy task 1. - Basic scope manipulation (3D Systems, 2017).

***Guided Anatomic Navigation***. As a second task from essential bronchoscopy, the guided anatomic navigation was executed. This task also resembles a basic training method for beginners and therefore just demands for dexterity and hand-eye-coordination skills. The setting here now looks like the innards of a human lung (see Figure A3) and the participants were induced to move the endoscope through the long ways. For orientation a picture of the position of the endoscope inside the bronchia was presented next to the task screen (see Figure A4). The task resembles a search task, in which cards of light bulbs have to be found and crossed as precisely as possible. If the destination was reached sufficiently, a positive sound appeared and the light bulb blinked green. Participants were instructed to collect 10 light bulbs for every round and repeat this task 15 times. Researcher and participants counted the light bulbs together, because otherwise the task could have been executed endlessly.

In total the second session took around 60-120 minutes, depending on the speed and skills of the participant. A break of 5 minutes was granted between the first and the second task. Generally it was aimed to keep at least one day between the first (EndoProto) and second (simulator) session of the study, in order to not hinder learning effects.

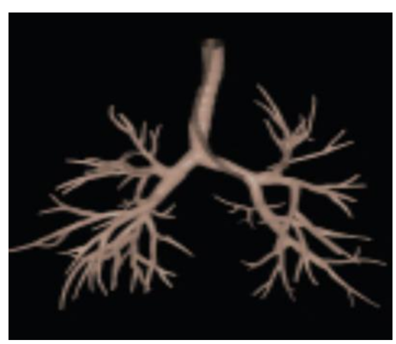


*Figure* A3. Bronchoscopic view for a participant moving with the endoscope inside the human bronchia for task 2 (left picture). Own photo.
*Figure* A4. Essential bronchoscopy task 2 - Guided anatomical guidance. Map of the human bronchia giving orientation for the position of the endoscope (right picture). (3D Systems, 2017).

**Location.** All simulator tasks took place at the University of Twente in the ECTM's (Experimental Center for Technical Medicine) simulator room 2.

## Measurements

The BRONCH Mentor provides plenty of parameters, but for the current study, the relevant variables of interest were task duration (time-on-task), number of repetitions (trials) and error rates (wall contacts). Time-on-task was measured with a stopwatch for task 2 or automatically saved via the simulator for task 1.

Collected data was written down in the following research protocol for session II. In order to get further exploratory insights in individual differences and impressions, observations and comments of participants were collected in an observation protocol (Appendix E). Data was saved in Excel tables on Google Drive.

## Participants

In order to insure no previous experience in endoscopy and to prevent interaction effects, one exclusion criteria was the participation in the study of the control group. For the simulator tasks color vision was not relevant.

**Research protocol session 2 – BRONCH Mentor**

Participant No.: ____________ Date: ___________________ Sona No.: ______________

*Task 2: Guided Anatomic Navigation*

| Light Bulb | Round 1 | Round 2 | Round 3 | Round 4 | Round 5 |
|---|---|---|---|---|---|
| Vocal Chord | | | | | |
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |

| Light Bulb | Round 6 | Round 7 | Round 8 | Round 9 | Round 10 |
|---|---|---|---|---|---|
| Vocal Chord | | | | | |
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |

**Notes and observations:**

______________________________________________________________

______________________________________________________________

______________________________________________________________

| Light Bulb | Round 11 | Round 12 | Round 13 | Round 14 | Round 15 |
|---|---|---|---|---|---|
| Vocal Chord | | | | | |
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |

Notice: Start the stopwatch when the endoscope enters the airway and stop again when it passes the vocal chords. Add a new round on the stopwatch after each repetition until the 10th light bulb was found. Repeat the task 15 times.

**Notes and observations:**

**Guidelines session 2 - BRONCH Mentor**

| | |
|---|---|
| Preparation | 1. Look up the participant number and log in, into the BRONCH Mentor system. Find the prepared schedule made up for our study. Scroll down at the tasks and courses and search for the course: "Lisa&Ace!" There you find all relevant tasks for the session. |
| Briefing (10 min) | *Welcome back! You will again train your surgery skills. Today You will use a professional simulator for bronchoscopy for this.*<br>● Give the participant the endoscope and reach him/her a small stool to stand on. Adjust the monitor to the participant's satisfaction.<br><br>*This is Your surgery tool for today. In order to know how to handle it, I will show You a short video introduction. (in total ~ 5 Minutes):*<br>● 1. Posture and scope maneuvering<br>● 2. Bronchoscopic navigation - fundamentals<br><br>Go with the participants through the following check-list/Control for:<br>● Is the tool in Your dominant hand and your thumb on the hand gear?<br>● Is the tip of the scope in the center of the airway?<br>● Please hold Your elbow in and just move the wrist, instead of the whole body.<br>○ Let participants demonstrate the movements and control if the scope is at its minimum - if the cable of the endoscope is straight. Let them try out the movable tip of the endoscope. |
| Preparation | Go in the menu to task 1 (Basic Scope Manipulation)<br>1. Go to BronchMentor, then scoll down completely to find the customized course: Lisa&Ace - Chose Task 1 Basic Scope Manipulation<br>2. Let the participant read the instructions |
| Instructions: Task I | *Okay, Are You ready for your first challenge?*<br>*In this first task You will insert the endoscope through the mouth of the patient into his/her lung and guide through it. Try to not hit the walls. The ball in the center before you helps You to reach precision. Try to be as precise and fast as possible. Do not feel discouraged from low scores. Surgeons have to train years to reach profession.*<br>● After You reached a goal, extract the endoscope completely out of simulator before starting a new trial.<br>● Please repeat the task 15 times. |

| | |
|---|---|
| | ● If you have any questions, do not hesitate to ask<br>● If you are ready, we will start now. Good luck! |
| Researcher - Attention | ● **Control if the participant repeats the task 15 times!**<br>● **Control if the participant holds the endoscope in a straight line. As necessary correct his/her posture to enable save handling of the equipment.**<br>● **Save the data after each trial!!! (close>save data)** |
| Break (5 Min) | Use the opportunity to answer open questions |
| Preparation | Start the second task (Guided Anatomic Navigation)<br>1. Go to BronchMentor, then scoll down completely to find the customized course: Lisa&Ace - Chose Task 2 Guided Anatomic Navigation<br>2. Let the participant read the instruction.<br>3. Provide additional instructions below: |
| Instructions: Task II | Your goal for this task is to find light bulbs. You can choose freely where to navigate to. If You successfully reached a light bulb a noise will inform You. In total You have to find 10. I count with you. The same rules apply as for the first task:<br>● Extract the endoscope completely after each trial.<br>● Try to be as precise and fast as possible.<br>● If you are ready, we will start now. Good luck! |
| Researcher Attention | ● **Start the stopwatch!**<br>● **Make a screenshot of the time on the screen!**<br>● **Control if the participant repeats the task 15 times. Count the light bulbs!**<br>● **Control if the participant holds the endoscope in a straight line**<br>● **Save the data after each trial!!! (close>save data). Participants have to click FINISH and SAVE!** |
| Debriefing | *Do You have any questions left? Would You like to get the results of Your surgery skills?*<br>→ Hand out contact information if necessary. Thank the person for his/her participation. |

## Appendix B - Informed consent

Participant/Protocol number: __________ Sona number: ___________________

Date: __________________ (Day/month/year) Participant name: _________________

Dear participant,

We are Ace Küpper and Lisa Mührmann. We are currently writing our master- and bachelor thesis in Human Factors and Engineering Psychology at the University of Twente. Our topic is "Learning minimally invasive surgery" and we want to test whether a specific training of dexterity tasks on a low-fi endoscopic prototype can influence the simulator task performance of bronchoscopy. Please ask us if You have any questions and we'll answer them.

Minimally invasive surgery (MIS) is one of the most used surgery methods today. In comparison to conventional open surgery, MIS offers advantages like a reduced blood loss, pain, complications, hospitalization time, and improved cosmetic results. However, these differences make performing minimally invasive surgery a great challenge for surgeons, who need a broad spectrum of cognitive and psychomotor skills. With the current research we want to investigate inter- and intrapersonal skill differences.

Your participation in this research is entirely voluntary. You may stop participating in the research at any time you choose. As You wish, Your collected data will then be deleted. In a first session You will train different dexterity tasks on an endoscopic prototype. In a second session You will train on a professional simulator for surgeons form the ECTM of the University of Twente. Every session will take approximately one hour.

We ensure that data collected for this research project will be made anonymous and kept confidential. In our final theses just anonymous data will be reported. The anonymized data will be provided to third parties, which are exclusively the University of Twente and our supervisors Dr. Martin Schmettow and Dr. Marleen Groenier. If You wish to get insight in Your data results, we gladly provide it to You via E-mail. For this purpose just contact us after Your participation.

**Declaration of agreement:** I have read and understood the preceding information. I have had the opportunity to ask questions. Every question was answered to my satisfaction. I consent voluntarily to participate in this study.

**Signature of Participant: ____________________ Date: ________________________**

I have witnessed the participant reading the consent form accurately. The participant had the opportunity to ask questions. I confirm that the individual has given consent freely. As requested a copy of the informed consent has been provided to the participant.

**Signature Researcher: _____________________ Date: _______________________**

## Appendix C – Pilot test

Participant/Protocol No.: 100 Date: 27.3.2018 Sona No.: anonymous

**Measurement table for dexterity task I – Allocentric**

Sub-Task 1.1. : Plate in Line II, colors: yellow & green (No.:1-9)

| Repetition | Route | Precision | Wall contact | Time | Skipped |
|---|---|---|---|---|---|
| 1 | 1 | True | 0 | 13 | False |
| 2 | 2 | True | 1 | 74 | False |
| 3 | 7 | True | 0 | 85 | False |
| 4 | 3 | True | 0 | 21 | False |
| 5 | 5 | True | 0 | 26 | False |
| 6 | 4 | True | 0 | 82 | False |
| 7 | 6 | True | 1 | 25 | False |
| 8 | 8 | True | 1 | 21 | False |
| 9 | 9 | True | 1 | 48 | False |
| 10 | 2 | True | 1 | 63 | False |
| Total | - | True | 5 | 403 | False |

Sub-Task 1.2. : Plate in Line III, colors: all (No.:1-20)

| Repetition | Route | Precision | Wall contact | Time | Skipped |
|---|---|---|---|---|---|
| 1 | 6 | True | 0 | 11 | False |
| 2 | 17 | True | 0 | 25 | False |
| 3 | 15 | True | 1 | 61 | False |
| 4 | 6 | True | 0 | 25 | False |
| 5 | 12 | True | 1 | 31 | False |
| 6 | 2 | True | 0 | 16 | False |
| 7 | 13 | True | 0 | 23 | False |
| 8 | 1 | True | 0 | 15 | False |
| 9 | 20 | True | 0 | 178 | False |
| 10 | 7 | True | 0 | 17 | False |
| 11 | 8 | True | 0 | 12 | False |
| 12 | 18 | True | 0 | 68 | False |
| 13 | 11 | True | 0 | 19 | False |
| 14 | 4 | True | 0 | 16 | False |
| 15 | 10 | True | 0 | 32 | False |
| 16 | 19 | True | 0 | 75 | False |

| | | | | | |
|---|---|---|---|---|---|
| 17 | 3 | True | 0 | 36 | False |
| 18 | 9 | True | 0 | 15 | False |
| 19 | 14 | True | 0 | 39 | False |
| 20 | 16 | True | 0 | 15 | False |
| Total | - | True | 2 | 744 | False |

**Notes & observations:**

- Participant tried to work very accurately and cause less wall contact. He rather moved slower in order to be precise – this behavior was also observable during the simulator task.
- He got the instruction: *Try to be as precise and fast as possible*. Nevertheless his trials took too much time, so that he could not execute task 1.3. [This also induced us to offer the skip option when a trial took more than 2 minutes.]
- He felt challenged and commented the task to be tricky and that a "calm hand" is needed for this. He also commented to be exhausted.
- Task 1.1. was more difficult for him than task 1.2., because of the restriction in movements due to a smaller distance. [This induced us to change the order of these tasks]

**Measurement table for dexterity task II – Egocentric**

Sub-Task 2.1. : Plate in Line III, colors: yellow, green & red (No.: 1-9, 16-18)

| Repetition | Route | Precision | Wall contact | Time | Skipped |
|---|---|---|---|---|---|
| 1 | 9 | True | 0 | 218 | False |
| 2 | 15 | True | 0 | 412 | True |
| 3 | 4 | True | 0 | 120 | False |
| 4 | 18 | True | 1 | 366 | False |
| 5 | 13 | True | 0 | 327 | True |
| 6 | 6 | True | 0 | 120 | False |
| 7 | 16 | True | 1 | 199 | True |
| 8 | 14 | True | 0 | 39 | False |
| Total | - | True | 2 | 1827 | 3 skipped |

Participant No.: 100 Date: 29.3.2018 Sona No.: anonymous

**Session 2 task 2: BRONCH Mentor - Guided Anatomic Navigation**

| Light Bulb | Round 1 | Round 2 | Round 3 | Round 4 |
|---|---|---|---|---|
| Vocal Chord | - | - | - | - |
| 1 | 111 | 12 | 21 | 30 |
| 2 | 157 | 23 | 5 | 2 |
| 3 | 130 | 24 | 9 | 3 |
| 4 | 140 | 12 | 35 | 4 |
| 5 | 126 | 18 | 4 | 13 |
| 6 | 245 | 8 | 10 | 15 |
| 7 | 125 | 10 | 33 | 21 |
| 8 | 123 | 11 | 13 | 20 |
| 9 | 88 | 31 | 5 | 7 |
| 10 | 87 | 23 | 23 | 27 |
| 11 | 86 | 58 | 6 | 2 |
| 12 | 72 | 33 | 22 | 75 |
| 13 | 94 | 12 | 12 | 5 |
| 14 | 131 | 26 | 11 | 17 |
| 15 | 82 | 41 | 28 | 14 |
| total | 1797 | 342 | 237 | 255 |

**Notes & observations:**

- Participant showed similar behavior as for the EndoProto training: he tried to be as precisely as possible and tried to prevent wall contact on costs of time. The aim for the simulator session was later to reach 15 repetitions in about an hour of time.
- Difficulties with the right sided-movements and mirror-inverted (to guide the endoscope to the right)

## Appendix D – Demographic survey

**Demografic Survey for BRONCH-Mentor**

Welcome!

This is the first part of the Twente Bronchoscopy Skills study. It is a short deomographic questionnaire. Data is worked with anonymously.

It You have any questions, feel free to ask the researcher.

1

**Demografic Survey for BRONCH-Mentor**

*** 1. Fill in participant number (SONA-Number)**

**This number You got from the researcher. It is 5 digits long.**

*** 2. Please fill in:**

Date / Time

MM/DD/YYYY hh mm -

*** 3. What is Your gender?**

- male
- female
- other

*** 4. Fill in Your birth date (MM-DD-YYYY).**

*** 5. Fill in Your nationality.**

*** 6. Please enter Your study.**

## Demografic Survey for BRONCH-Mentor

*** 7. Fill in Your dominant hand.**

- ◯ left-handed
- ◯ right-handed
- ◯ both

*** 8. Do You have impaired vision?**

- ◯ Yes, I wear glasses.
- ◯ Yes, I wear contact lenses.
- ◯ Yes, but I do not war glasses or contact lenses.
- ◯ Yes, I am color-blind.
- ◯ Yes, I have eye cataract (grijze staar).
- ◯ Yes, I have glaucoma (groene staar).
- ◯ No.

Other (please explain)

## Demografic Survey for BRONCH-Mentor

*** 9. Do You have dyslexia?**

**(Dyslexia is signalised by a difficulty in speech comprehension: reading and understanding written texts.)**

- ◯ Yes
- ◯ No
- ◯ I don't know

*** 10. Do You have experience in the field of endoscopy ?**

- ◯ No
- ◯ Yes

If YES,
Please specify in hours per week and extend:

*** 11. Do you have experience in video/computer gaming?**

- ◯ No
- ◯ Yes

If yes,
please specify Your gaming hours per week:

**Demografic Survey for BRONCH-Mentor**

Thank You for filling in Our survey.
We wish You joy with the upcoming surgery-challenges.

Ask the researcher for further instructions as needed.

**Appendix E – Technical drawing of the EndoProto**

**Figure E1**

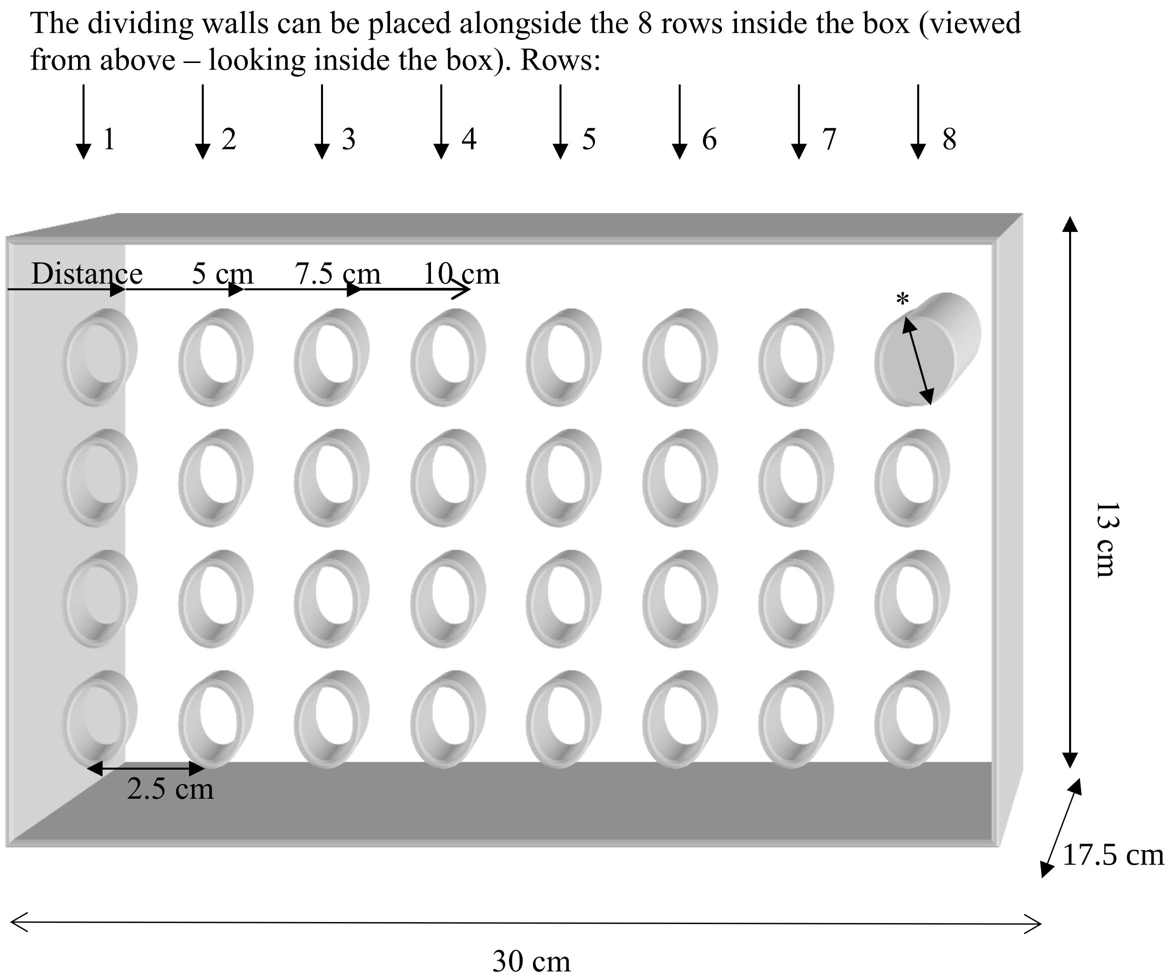



*Figure E1.* * The diameter of the holes differ between the frame of the box (6 mm) , the dividing wall for the allocentric task (6 mm) and the dividing wall for the egocentric task (11 mm). Own design.

**Appendix F – Research protocol**

**1. Test time-frame.**

| Session | Content | Duration |
|---|---|---|
| 1 | Debriefing (Explanation, Informed Consent) | ~ 5 Min |
| | Task 1.1. → aim for 50 repetitions<br>With the walls 3 different distances:<br>Line 2 - 10 trials,<br>Line 3 - 20 trials,<br>Line 4 - 20 trials or more if time permits | 20-25 Min |
| | Break | ~5 Min |
| | Task 1.2. → aim for 50 repetitions | 25-30 Min |
| | Total: | 60 Min |
| 2 | Debriefing (Explanation, Video tutorials) | 10 Min |
| | Task 2.1. 15 Repetitions | ~ 30-50 |
| | Break | 5 Min |
| | Task 2.2. Light bulbs - after 10 found pieces data saving, then starting a new trial. | 30-55 Min |
| | Total: | until 120 Min |

Participant/Protocol No.: __________ Date: __________________ Sona No.: ______________

**2. Measurement table for dexterity task I – Allocentric**

| Sub-Task 1.1. : Plate in Line III, colors: all (No.:1-20) | | | | | |
|---|---|---|---|---|---|
| Repetition | Route | Precision | Wall contact | Time | Skipped |
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |
| 11 | | | | | |
| 12 | | | | | |
| 13 | | | | | |
| 14 | | | | | |
| 15 | | | | | |
| 16 | | | | | |
| 17 | | | | | |
| 18 | | | | | |
| 19 | | | | | |
| 20 | | | | | |

| Sub-Task 1.2. : Plate in Line II, colors: yellow & green (No.:1-9) | | | | | |
|---|---|---|---|---|---|
| Repetition | Route | Precision | Wall contact | Time | Skipped |
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |

| Sub-Task 1.3. : Plate in Line IV, colors: all (No.:1-20) | | | | | |
|---|---|---|---|---|---|
| Repetition | Route | Precision | Wall contact | Time | Skipped |
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |
| 11 | | | | | |
| 12 | | | | | |
| 13 | | | | | |
| 14 | | | | | |
| 15 | | | | | |
| 16 | | | | | |
| 17 | | | | | |
| 18 | | | | | |
| 19 | | | | | |
| 20 | | | | | |

**Notes & observations:**

**3. Measurement table for dexterity task II – Egocentric**

| Sub-Task 2.1. : Plate in Line III, colors: yellow, green & red (No.: 1-9, 16-18) | | | | | |
|---|---|---|---|---|---|
| Repetition | Route | Precision | Wall contact | Time | Skipped |
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |
| 11 | | | | | |
| 12 | | | | | |
| 13 | | | | | |
| 14 | | | | | |
| 15 | | | | | |
| 16 | | | | | |
| 17 | | | | | |
| 18 | | | | | |
| 19 | | | | | |
| 20 | | | | | |
| 21 | | | | | |
| 22 | | | | | |
| 23 | | | | | |
| 24 | | | | | |
| 25 | | | | | |
| 26 | | | | | |
| 27 | | | | | |
| 28 | | | | | |
| 29 | | | | | |
| 30 | | | | | |
| 31 | | | | | |
| 32 | | | | | |
| 33 | | | | | |
| 34 | | | | | |
| 35 | | | | | |
| 36 | | | | | |
| 37 | | | | | |
| 38 | | | | | |
| 39 | | | | | |
| 40 | | | | | |

**4. Glossary**

**Attention**:

- Take care, that the participant moves the endoscope (in egocentric) after each trial back to the starting point - signalized through a visible 2 in a white square on the desktop screen, which can be found at the outer side of the EndoProto.
- Pay attention that the brightness of your desktop stays constant and that your screensaver is turned off during the experiments. Use the rubber component to locate the box on a table, so that it does not get out of place suddenly and disturbs the movements of participants.
- **Most important:**

  **Write down all measurements at once as you observe them!**

  **Do not wait for distractions as questions to let them make you forget. Always provide some additional time between two participants to note your observations.**

**Task**:

Every wall has 20 openings or routes. Depending on the task, participants get one route a time by the researcher and shall try to insert the endoscope in this hole. Crossing an opening/route is counted as one repetition. Routes are randomly generated. Based on our experience with the pilot testing (Appendix C), reasons of complexity, feasibility and efficiency and aims of training, every sub-task has a different amount of routes (task 1.1.: 20, task 1.2.: 9, task 1.3.: 20, task 2.1.: 12). A general aim is to reach as many repetitions as possible for task 1.3. and 2.1. in the given time.

**Time**:

The time is stopped for every trial/repetition. The easiest way to do this is to add new rounds on the stopwatch for detailed timestamps for every trial and the whole sub-task.

**Skip:**

If a participant takes longer than two minutes to complete a given route, an option to skip is presented. Then the participant gets a new destination. Participants can also choose to complete the present route.

**Wall contact:**

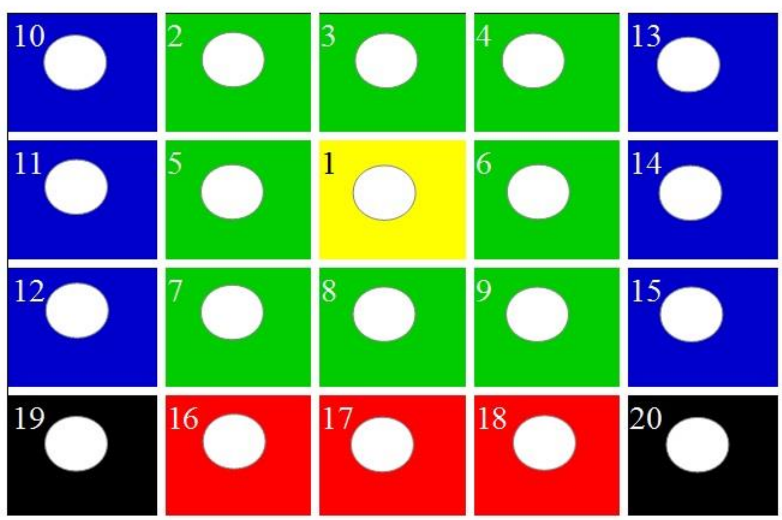
Count the number of touching points of the endoscope with the walls. These signalize wall contacts to the 'human tissue' and shall be prevented. Every contact with any of the walls inside the box is counted as one wall contact. Scratching along a wall inside the box is also counted as one point. The head of the endoscope may get stuck at one opening, but slip off again, without reaching success. In this case one wall contact point is being counted. Extracting the endoscope out of one opening may lead to unwanted touching points of the endoscope with the environment. These contacts are not counted, because they can hardly be controlled by the participant. Moreover, they happen when the endoscope is moved back again to the starting point, which signalizes that the trial is over. A trial ends, when a participant reaches the destination – crosses a route.

**Precision**:

The precision indicates if the participant was able to insert the endoscope through the right routes/holes, or if the participant went through a wrong hole or into the same one without instructions. If a wrong route was crossed, the number of this route is written down in the cell for precision. Then the trial is automatically skipped and a new destination is given to the participant.

**Coloring scheme:**

Every wall is colored and numbered as the following graphic shows. For the tasks of allocentric this wall is turned upside down and attached inside the box. Aspects of pattern recognition or working memory are not part of this study, but could be easily tested with the EndoProto.



**5. Guidelines – EndoProto session**

<table>
<tr><td>Preparation beforehand</td><td><ul>
<li>Look up the Sona-No. and prepare an informed consent and test protocol with a new anonymized participant number.</li>
<li>Generate random routes for the sub-tasks of session 1 before the session starts and write them down in the test protocols. Make use of http://www.zufallsgenerator.net/<br>or use R to program them yourself.</li>
<li>Always bring your laptop (loaded) and control if screensaver and brightness regulations are deactivated. No disturbing messages should appear (e.g. Skype messages)</li>
<li>Make sure that your mobile phone (stopwatch) is loaded.</li>
<li>Bring enough copies of test protocols and informed consents.</li>
<li>Make sure that you have all necessary test equipment: Laptop, stopwatch, rubber component, protocols, informed consents, pens, EndoProto, plates, booked room for session 1 or simulator room (Simroom 2, at the ECTM)</li>
<li>If you just have participants for session 1 book a different test room, that ensures more silence and fewer distractions!</li>
</ul></td></tr>
<tr><td>Briefing</td><td>Welcome!<br>Today You will perform 2 different dexterity tasks on a MIS-prototype. The duration is 1 hour. You will get 1 Sona credit for Your participation. If You agree, you can take part in another session, which takes around 1 ½ until 2 hours. You can earn 2 additional Sona credits for Your participation. The second session will take place at the earliest 1 day after the first session.<br>☐ Would You like to take part in the study?<br><br>➔ Control if the participant has completed the demographic questionnaire (survey)!<br>◆ If not, let him/her complete it now (duration 4 Minutes) or ask him/her to do it later.<br>➔ Let the participant read and sign the informed consent!</td></tr>
<tr><td>Preparation of task</td><td>Prepare the stopwatch. Let the participant sit on the table and show him/her the materials (cable and box) and explain the following:</td></tr>
<tr><td>Task instructions:</td><td>You may use one or even both hands for the task, but your hands have to stay outside the box all the time. The box may not</td></tr>
</table>

| | |
|---|---|
| Dexterity task I | *be moved. You have to sit on a chair while performing this task and begin at the starting point '1'. After each trial you have to go back to the starting point for the next trial. If the endoscope gets stuck inside an opening, I can help You to extract it.*<br>● *Do you have any questions so far?*<br>*If you have any urgent questions, please ask first, after you have reached a destination and turned back to the starting point. If the question is not that urgent, please wait until the end of the first task when you get a short break.*<br><br>***Try to be as precise and fast as possible.***<br>***Have fun.*** |
| Break - 5 Minutes | Use this time to answer open questions and to prepare the equipment for the second task. |
| Task preparation | Connect the endoscope to the PC and start the ViewPlayCap program. Make sure that it works correctly. Administer the brightness of the endoscope to the highest stage. Place the screen in a pleasant position for the participant. Prepare the stopwatch. |
| Task instructions: Dexterity task II | *Please stand up now and grasp the cable. Look at the square with the '2'. This is Your new starting point for this task. Please remember as for the first task: Always go back to the starting point before starting a new trial. You will now repeat the same tasks as in the first half of our session, but now you keep standing and your view has to be fixed on the screen, as You maneuver the endoscope inside the box.*<br>● *Do You have a good view on the screen?*<br>*If the endoscope gets stuck inside an opening, I will help You to extract it. If you completely lose your orientation, You may always go back to the starting point to re-orientate and start again. I let You know, when You have mastered a route.*<br><br>***Again, try to be as precise and fast as possible.***<br>***Good luck!*** |
| Announcement for researcher: | The current study did not focus on working memory abilities. Therefore, the destination (number of a route) was re-mentioned, if participants forget and/or asked for it. |

| | |
|---|---|
| Debriefing | Thank the participant and ask if he/she would like to take part in the second session of the study. If yes, request them to subscribe for session 2 or assign them manually to ensure their participation. If you do the latter, always make sure the participant agrees to the next appointment. |
| After session 1 | For researchers:<br>Complete the protocol as necessary. Add relevant observations or thoughts on the protocol. Make sure the information you have written down is also clear to you in case you have to look it up later again! Try to fill in the collected data as fast as possible in the prepared Excel documents.<br>Organizational :<br>- Grant Sona credits for completed sessions.<br>- Control your agenda - coordinate sessions and participants. |

**Appendix G – Observation protocol**

| No. | Participant | Observation |
|---|---|---|
| General impressions | | • Most people seemed to take more time for the trials in the task of egocentric. They also seemed to make less wall contacts in this condition.<br>• Precision was rather a problem in dexterity task II, because of a loss in orientation and the difficulty to estimate the distance between the tip of the endoscope and the target hole – although exactly the same distance was trained before (task 1.1.).<br>• All participants evaluated the egocentric task as being more demanding than the task of allocentric. Own design. |
| Unspecified comments | | • *Does everybody make so many mistakes?*<br>• *I will not become a surgeon like this.*<br>• |
| 0 | 100 (Pilot test) | • Participant tried to work very accurately and cause less wall contact; rather moved slowly in order to be precise.<br>• We instructed: *Try to be as precise and fast as possible*. Nevertheless the repetitions took too much time, so that task 1.3 could not be executed [This induced us to offer .the skip option when a trial took more than 2 minutes.]<br>• Participant felt challenged, exhausted and commented the task to be tricky and that a "calm hand" is needed for this.<br>• Task 1.1 was more difficult than task 1.2, because of the restriction in movements due to a smaller distance. [This induced us to change the order of these tasks]<br>• <u>Simulator</u>: similar behavior as for the EndoProto training: tried to be as precisely as possible and tried to prevent wall |

| | | |
|---|---|---|
| | | contact on costs of time. <br>• Difficulties with the right sided-movements and mirror-inverted (to guide the endoscope to the right) |
| 1 | 103 | • Movements were rather fast, which led to a high amount of wall contacts and lacks in precision <br>• Simulator: Fast. |
| 2 | 109 | • Fast movements of the endoscope, which led to a high amount of wall contacts. |
| 3 | 110 | • Rather reluctant to take the option of skipping a route, which caused high time intervals and wall contacts. <br>• Simulator: Difficulties in guiding the endoscope to the left side – rather moved it involuntarily to the right side. |
| 4 | 112 | • Fast movements and less wall contacts, but poor precision. |
| 5 | 113 | • High amount of wall contacts; lacks in precision and rather slow movements. Impression: uncoordinated. <br>• Simulator: Difficulties in guiding the endoscope. Movements were fast. Hands were shaking. Impression: restless and impatient. |
| 6 | 117 | • Good performance parameters: fast movements, less time, less wall contacts, high precision rates; impression: Concentrated and calm <br>• Simulator: Same concentrated and calm impression as for the box training. Fast progress and movement speed. |
| 7 | 101 | • Slow movements, scared to make errors, but had high rates of wall contacts. <br>• Difficulties in guiding the endoscope, needed long thinking breaks before starting to maneuver the endoscope, after assurance to maneuver in the right direction the speed increased, but precision was lacking |
| 8 | 104 | • Very accurate and precise, |

| | | |
|---|---|---|
| 9 | 105 | • Slow movements, high wall contacts, bad precision,<br>• Often changed positioning, agitated |
| 10 | 106 | • In box training in allocentric worse than in egocentric! |
| 11 | 107 | • Participant often makes wall contacts directly near to the destination and then crosses route; rather takes time to continue instead of skipping a route,<br>• Describes routes 16 and 7 as very complicated in egocentric task; uses small finger movements and fingertips to maneuver most of the time in task I and II,<br>• Sometimes the camera is being bent outside of the box in such a way that does not affect the movements inside (displacement activity?)<br>• Participant comments:<br>- *It is complicated to find the orientation for the direction of the camera. It is like in gaming, when flying a helicopter.*<br>- *One has to know first, that if I turn here, the endoscope may turn in a different direction.*<br>- *This view-task is definitely more complicated than the first one.* |
| 12 | 108 | • Dexterity task I: very fast, but causes many wall contacts<br>• Dexterity task II: Difficulties in orientation and moving the endoscope, made use of small and hard movements |
| 13 | 111 | • Comment of participant: *Bad view from the camera*.<br>• Needs many attempts until finally crossing a route; loses vision because of the physiology and view of the camera. |
| 14 | 114 | • Fast movements, high amount of wall contacts; Impression: impatient; wanted to skip a route very quickly |
| 15 | 115 | • Fast movements, but bad precision |

| | | |
|---|---|---|
| 16 | 116 | • Low wall contact, calm hands and precise movements<br>• Left side appeared easier to the participant |
| 17 | 118 | • Very impatient and hectic; asked plenty of times for help and about the needed time; caused many wall contacts |
| 18 | 119 | • Fast movements but also much wall contact |
| 19 | 120 | • Lacking in precision; movement of the endoscope was quick and hard in allocentric, but rather slow in the egocentric<br>• Problems to orientate in egocentric |
| 20 | 121 | • Fast, precise, calm and focused |
| 21 | 122 | • Tasks of allocentric: fast and good orientated, but for egocentric: loss of orientation, bad precision |
| 22 | 123 | • High wall contact rate; dexterity task II: precision worsened |
| 23 | 124 | • Left side routes appeared easier; very slow movements<br>• Shaking hands, caused much wall contact<br>• Person first tried to manage the direction of the endoscope and went back to the starting point, as it changed suddenly – which induced low ToT |
| 24 | 125 | • Good precision; right side routes appeared more difficult<br>• Described dexterity task II as much more difficult than the first task; concentrated, focused, calm hands |

*Note*: Observations were made per participant during and after a session took place. Categorization was made afterwards, in order to classify possible types of participants. Observations were subjective and made by the researcher who was responsible for the related participant at the session. Due to the fact, that sessions may have been supervised in different languages, all citations shall be regarded rather as paraphrases.

# Appendix H – R syntax

## Data analysis

Martin Schmettow

09 June, 2018

```
knitr::opts_knit$set(warning = F, message = F)
```

```
purp.data = F
purp.mcmc = T
```

```
library(tidyverse)
```

```
## -- Attaching packages ------------------------------------------------------
--------------------- tidyverse 1.2.1 --
```

```
## v ggplot2 2.2.1     v purrr   0.2.4
## v tibble  1.4.2     v dplyr   0.7.4
## v tidyr   0.8.0     v stringr 1.3.1
## v readr   1.1.1     v forcats 0.3.0
```

```
## -- Conflicts ---------------------------------------------------------------
--------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(readxl)
```

```
library(brms)
```

```
## Loading required package: Rcpp
```

```
## Loading 'brms' package (version 2.3.0). Useful instructions
## can be found by typing help('brms'). A more detailed introduction
## to the package is available through vignette('brms_overview').
## Run theme_set(theme_default()) to use the default bayesplot theme.
```

```
library(rstanarm)
```

```
## rstanarm (Version 2.17.4, packaged: 2018-04-13 01:51:52 UTC)
```

```
## - Do not expect the default priors to remain the same in future rstanarm ve
rsions.
```

```
## Thus, R scripts should specify priors explicitly, even if they are just the
 defaults.
```

```
## - For execution on a local, multicore CPU with excess RAM we recommend call
ing
```

```
## options(mc.cores = parallel::detectCores())

## - Plotting theme set to bayesplot::theme_default().

##
## Attaching package: 'rstanarm'

## The following objects are masked from 'package:brms':
##
##     exponential, kfold, lasso, ngrps
```

```
options(mc.cores = 6)
library(mascutils)
library(asymptote)
```

```
##
## Attaching package: 'asymptote'

## The following objects are masked from 'package:mascutils':
##
##     inv_logit, logit
```

```
library(bayr)
```

```
##
## Attaching package: 'bayr'

## The following objects are masked from 'package:rstanarm':
##
##     fixef, ranef

## The following objects are masked from 'package:brms':
##
##     fixef, ranef

## The following object is masked from 'package:stats':
##
##     predict
```

```
load("AK18.Rda")
```

### Data preparation - Box study

```
box_files <-
  dir(path = "raw_data/AK/",
      pattern = "^Participant\\d{3}_Box\\.xls",
      full.names = T,
      recursive = T)
```

```
check_box <- function(x){
  colnames(read_xls(x))
}
```

```
read_box <- function(x){
  print(x)
  read_xls(x) %>%
    select(-Date) %>%
    #select(Participant, Task, Repetition, Route, Precision, Wrong_Route, Skip
ped, Wall_contact, Time, TimeOnTask) %>%
    mutate(Route = as.character(Route),
           #Precision = as.character(Damage),
           Wall_contact = as.character(Wall_contact),
           Time = as.numeric(Time),
           TimeOnTask = as.numeric(TimeOnTask))
}


AK18 <-
  set_names(box_files) %>%
  map_df(read_box) %>%
  select(-TimeOnTask) %>%
  filter(!is.na(Time)) %>%
  tidyr::separate(Task,
                  into = c("View", "Distance", "Marvin"),
                  sep = "\\.") %>%
  mutate(Part = str_extract(Participant, "\\d+"),
         trial = as.integer(Repetition),
         Route = as.character(Route),
         Wrong_Route = as.numeric(Wrong_Route),
         Setup = "Box",
         Wall_contact = as.integer(Wall_contact),
         View = if_else(View == "1", "allo", "ego"),
         Task = str_c(View, Distance, sep = "_"),
         ToT = Time/60,
         Precision = if_else(Precision == "TRUE", T, F)) %>%
  filter(!is.na(Repetition), is.na(Wrong_Route)) %>%
  select(Setup, Part, View, Distance, Task,
         Route, trial, ToT, Wall_contact, Precision) %>%
  print()

save(AK18, file = "AK18.Rda")
```

### Data exploration

```
load("AK18.Rda")
```

#### Descriptives

Number of observations

```
AK18 %>%
  group_by(Setup, Part, Task) %>%
```

```
  summarize(N_trials = n()) %>%
  ungroup() %>%
  group_by(Setup, Task) %>%
  summarize(N_Part = n(),
            min(N_trials), median(N_trials), max(N_trials), sd(N_trials)) %>%
  knitr::kable()
```

| Setup | Task | N_Part | min(N_trials) | median(N_trials) | max(N_trials) | sd(N_trials) |
|---|---|---|---|---|---|---|
| Box | allo_1 | 24 | 20 | 20 | 20 | 0.000000 |
| Box | allo_2 | 24 | 5 | 10 | 10 | 1.020621 |
| Box | allo_3 | 13 | 14 | 20 | 20 | 1.664101 |
| Box | ego_1 | 24 | 14 | 25 | 39 | 6.192416 |

```
AK18 %>%
  group_by(Route, View, Distance) %>%
  summarize(n_obs = n()) %>%
  spread(View, n_obs)
```

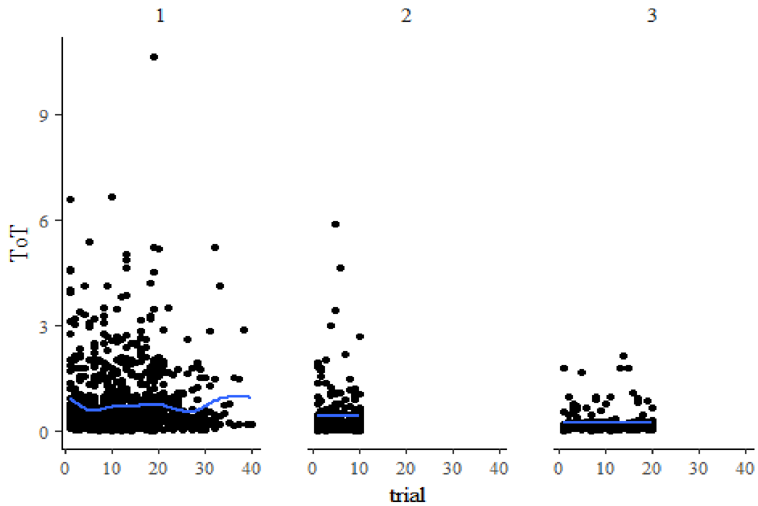
```
## # A tibble: 50 x 4
## # Groups:   Route [20]
##    Route Distance  allo   ego
##    <chr> <chr>    <int> <int>
##  1 1     1           24    53
##  2 1     2           26    NA
##  3 1     3           11    NA
##  4 10    1           25     2
##  5 10    2            2    NA
##  6 10    3           10    NA
##  7 11    1           27    NA
##  8 11    3           12    NA
##  9 12    1           22     3
## 10 12    3           11    NA
## # ... with 40 more rows
```

## Explorative analaysis

### Raw learning curves

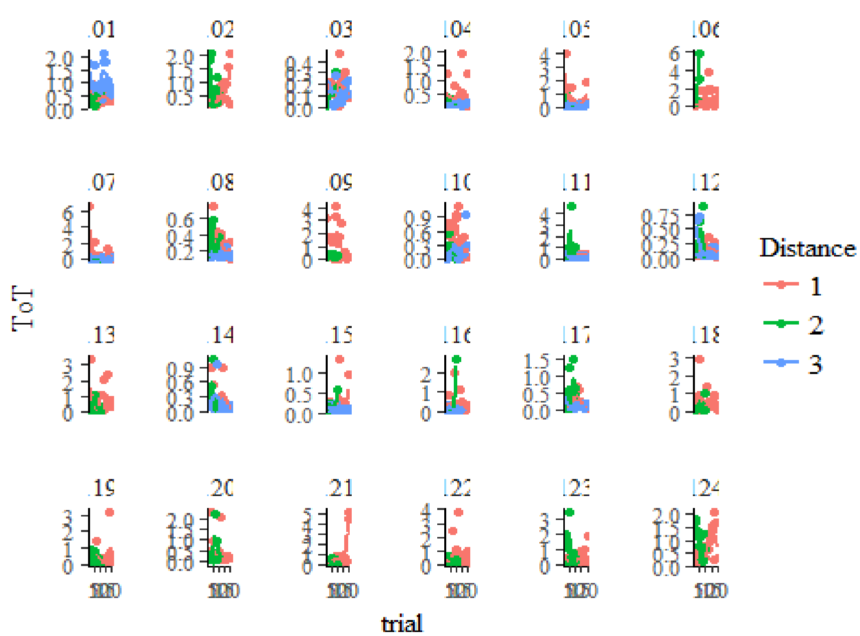
```
AK18 %>%
  ggplot(aes(x = trial, y = ToT)) +
  facet_grid(~Distance, scale = "free_y") +
  geom_point() +
  geom_smooth(se = F)
```

```
## `geom_smooth()` using method = 'gam'
```




```
AK18 %>%
  filter(View == "allo") %>%
  ggplot(aes(x = trial, color = Distance, y = ToT)) +
  facet_wrap(~Part, ncol = 6, scale = "free_y") +
  geom_point() +
  geom_smooth(se = F)
```

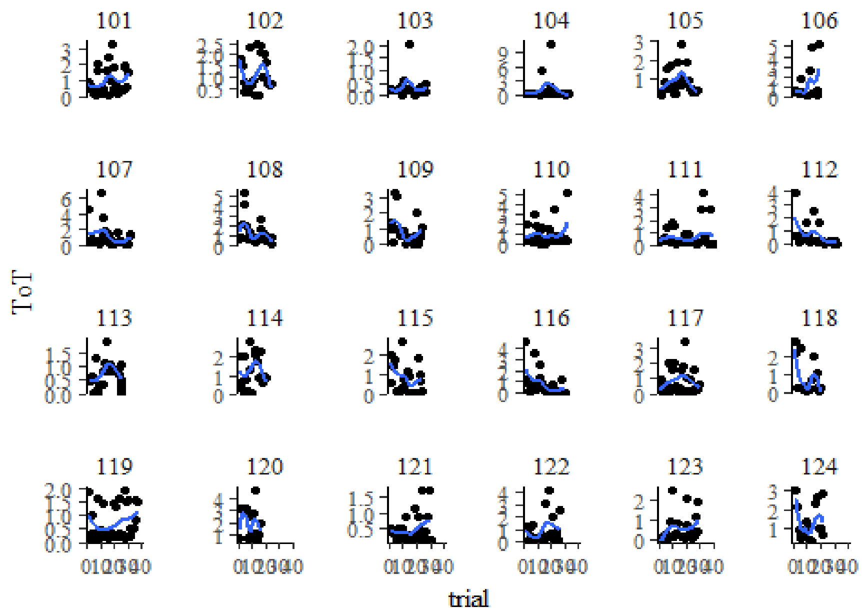
```
## `geom_smooth()` using method = 'loess'
```




```
AK18 %>%
  filter(View == "ego") %>%
  ggplot(aes(x = trial, y = ToT)) +
  facet_wrap(~Part, ncol = 6, scale = "free_y") +
```

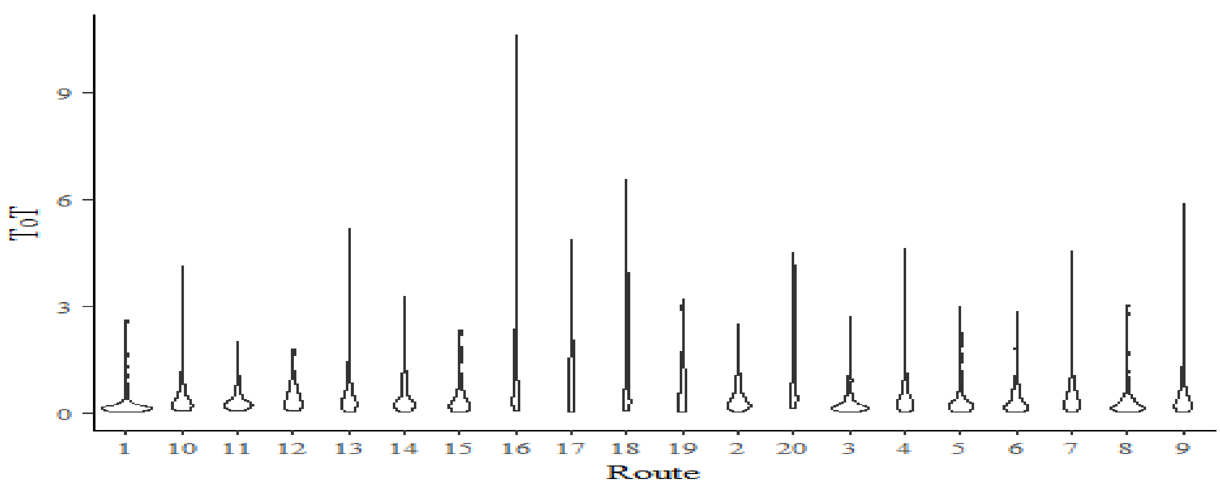
```
  geom_point() +
  geom_smooth(se = F)
```

```
## `geom_smooth()` using method = 'loess'
```




**Examining Routes**

```
AK18 %>%
  ggplot(aes(x = Route, y = ToT)) +
  geom_violin()
```



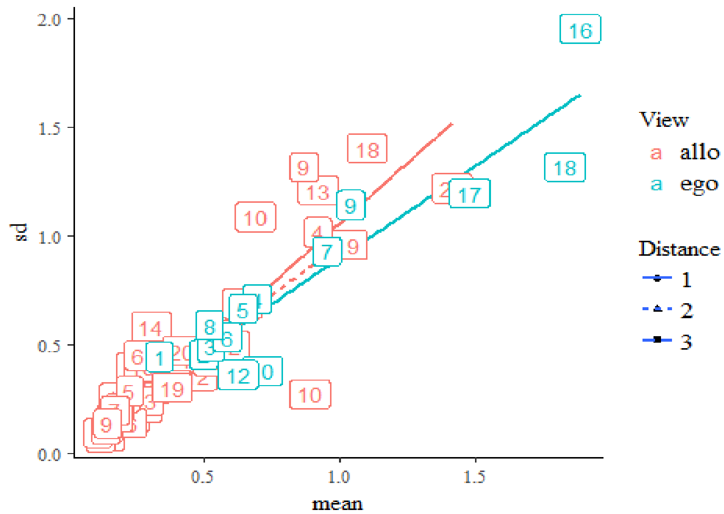

```
AK18 %>%
  group_by(View, Distance, Route) %>%
  summarize(mean = mean(ToT),
            sd = sd(ToT)) %>%
  ggplot(aes(x = mean, y = sd, color = View, linetype = Distance, shape = Dist
```

```
ance, label = Route)) +
  geom_point() +
  geom_smooth(se = F, method = "lm") +
  geom_label()
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing missing values (geom_label).
```



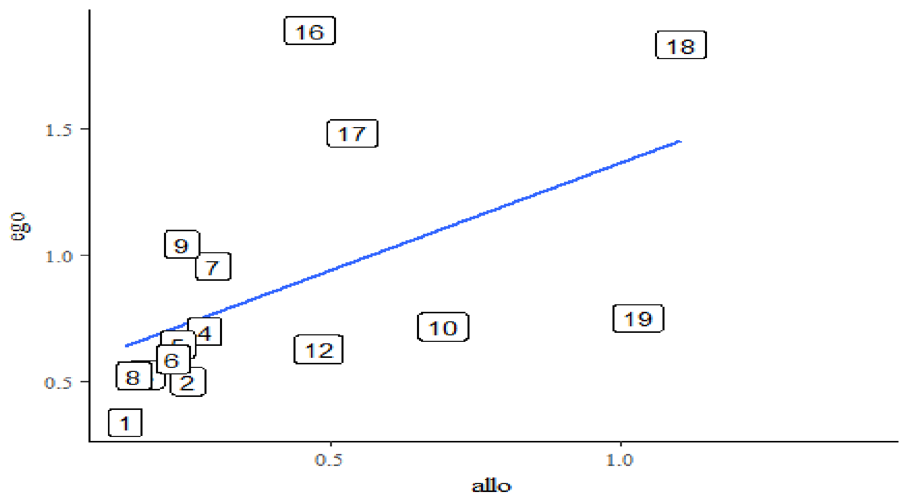

- linear relationship between mean and sd. That should be covered by the model.

```
AK18 %>%
  filter(Distance == 1) %>%
  group_by(View, Route) %>%
  summarize(mean = mean(ToT)) %>%
  spread(View, mean) %>%
  ggplot(aes(x = allo, y = ego, label = Route)) +
  geom_point() +
  geom_smooth(se = F, method = "lm") +
  geom_label()
```

```
## Warning: Removed 5 rows containing non-finite values (stat_smooth).

## Warning: Removed 5 rows containing missing values (geom_point).

## Warning: Removed 5 rows containing missing values (geom_label).
```




- potentially poor correlation between difficulty of routes across View. We may need an interaction effect *or* dismiss egocentric perspective data *or* remove some routes (see below)
- Routes 16, 17, 18 seem to be extremely difficult under ego view *and* produce extreme variance

```
AK18_allo <-
  AK18 %>%
  filter(View == "allo")
```

### Learning curves

Any atte,pt to estimate proper learning curves failed. See below for an analysis. In brief: not a single participant showed signs of learning with the box.

Setting up the LARY model:

#### Estimated curves

#### Effects

Individual differences as standard deviations by task and parameters:

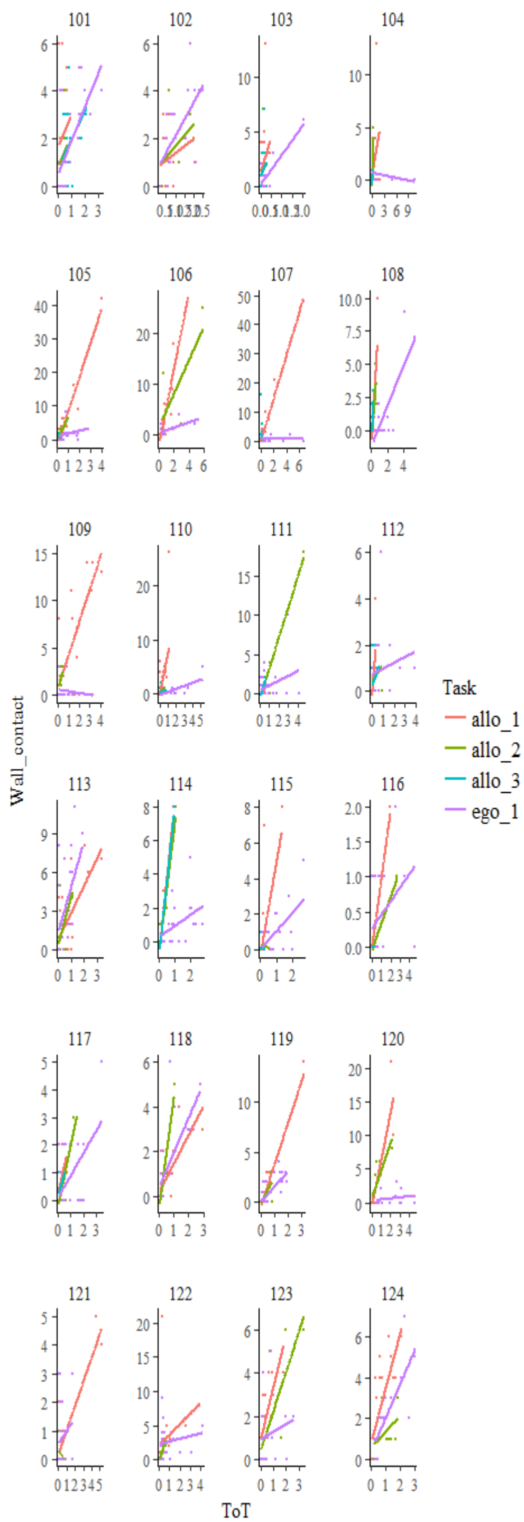
### Association between performance measures

#### Exploratory analysis

Is there any sign of learning when avareging over all participants?

```
AK18 %>%
  ggplot(aes(x = ToT, y = Wall_contact, color = Task)) +
  facet_wrap(~Part, scale = "free", ncol = 4) +
```

```
  geom_point(size = .4) +
  geom_smooth(se = F, method = "lm")
```



## Regression

### *M_6: Checking for learning effects*

The following regression model estimates the *individual* associations between number of wall contacts by ToT. We control for trial, Distance and View.

This analysis will, among others, produce intercept effects (participant-level and population-level). To make this parameter more meaningful, we center `ToT` at the population average.

Similarly, the variable `trial` is shifted by one to make the intercept represent the first trial.

```
tot_pop_avg <- mean(AK18$ToT)
```

```
AK18 <-
  AK18 %>%
  mutate(cToT = ToT - tot_pop_avg,
         trial = trial - 1) %>%
  as_tbl_obs()
```

First, we examine if there is any learning effect at all, meaning on the population-level and the participant-level.

```
M_6 <- stan_glmer(Wall_contact ~ 1 + trial +
                  (1 + trial|Part) +
                  (1|Route) +
                  (1|Obs),
             family = poisson,
             data = AK18,
             init = "0")
```

```
P_6 <- posterior(M_6)
```

```
## Warning in sqrt(value): NaNs produced
```

```
save(M_6, P_6, file = "M_AK_6.Rda")
```

```
load("M_AK_6.Rda")
```

```
fixef(P_6, mean.func = exp)
```

```
##
##
## Table: Estimates with 95% credibility limits
##
## fixef             center       lower      upper
## ----------  ----------  ----------  ---------
## Intercept    0.8288882   0.5342585   1.268346
## trial        0.9919659   0.9760908   1.008340
```

The intercept effect is the average number of wall contacts, at the first trial.
By exponentiation, that trial estimate is multiplicative. The effect of trial is practically 1; in multiplication that means: no change. Also given the tiny CIs We can be rather certain that *on average*, people do not learn to have fewer wall contacts.

Is that true for all individuals? We take a look at the participant level. This is more easily done on the *linear predictor scale*, which is the logarithm of the multiplicative scale. Here, things are additive, where 0 means no change. We compare the pop-level effects to the standard deviation of the individual deviations.

```
left_union(fixef(P_6) %>% discard_redundant(),
           grpef(P_6) %>%
             filter(re_factor == "Part") %>%
             rename_if(is.numeric, funs(str_c(., "_sd"))) %>%
             discard_redundant()
)
```

```
## # A tibble: 2 x 7
##   fixef        center   lower   upper center_sd lower_sd upper_sd
##   <fct>         <dbl>   <dbl>   <dbl>     <dbl>    <dbl>    <dbl>
## 1 Intercept -0.188    -0.627  0.238      0.719   0.495     1.04
## 2 trial     -0.00807 -0.0242 0.00831    0.0277  0.00972   0.0475
```

Compared to the very small effect of trial there is substantial variation. However, that by no means makes anyone in the sample a visible learner. We can exclude learning (and exhaustion effects) on average and for practically all individuals. Note the pronounced variation of performance at trial 1 (Intercept).

### *M_5: ToT and wall contact*

In any case, we can exclude the predictor trial from teh following model, which will answer our main question: how are the two performance indicators, wall contact and ToT related. Very importantly, we have to control for the different routes, as they strongly effect both: ToT and wall contacts.

```
M_5 <-
  stan_glmer(Wall_contact ~ 1 + cToT + Task +
                (1 + cToT|Part) +
                (1|Route) +
                (1|Obs),
             family = poisson,
             data = AK18)
```

```
P_5 <- posterior(M_5)
```

```
## Warning in sqrt(value): NaNs produced
```

```
save(M_5, P_5, file = "M_AK_5.Rda")
```

```
load("M_AK_5.Rda")
```

The effect is dramatic: On average, with every minute longer, the number of wall contacts multiplies by 1.79, or increases by almost 80%, with good certainty.

How is the variation between individuals. The following effects are on a logarithmic scale,

making things additive, with 0 being no effect.

```
left_union(fixef(P_5) %>%
             filter(fixef  %in% c("Intercept", "cToT")) %>%
             discard_redundant(),
           grpef(P_5) %>%
             filter(re_factor == "Part") %>%
             rename_if(is.numeric, funs(str_c(., "_sd"))) %>%
             discard_redundant()
)
```

```
## # A tibble: 2 x 7
##   fixef     center  lower upper center_sd lower_sd upper_sd
##   <fct>      <dbl>  <dbl> <dbl>     <dbl>    <dbl>    <dbl>
## 1 Intercept -0.122 -0.495 0.228     0.525    0.377    0.734
## 2 cToT       0.586  0.443 0.734     0.278    0.166    0.435
```

Compared to the average effects size (0.58), the variation is existent, but not overwhelming. By far most participants have a positive relation between ToT and wall contact.

***M_7: ToT and wall contact, interaction effect Route by Task***

The following model etnends the previous by adding the possibility that the route effects differ by task. In addition, the model accounts for overdispersion by using the negative binomial distribution instead of Poisson.

```
M_7 <-
  brm(Wall_contact ~ 1 + cToT + Task +
                (1 + cToT|Part) +
                (1 + Task|Route),
             family = negbinomial(link = log),
             iter = 4000,
             warmup = 3000,
             data = AK18)
```

```
## Compiling the C++ model
```

```
## Start sampling
```

```r
P_7 <- posterior(M_7)
```

```r
save(M_7, P_7, file = "M_AK_7.Rda")
```

```r
load("M_AK_7.Rda")
```

```r
fixef(P_7, mean.func = exp)
```

```
##
##
## Table: Estimates with 95% credibility limits
##
## fixef              center       lower       upper
## -----------  ----------  ----------  ----------
## Intercept     1.2474457   0.8273745   1.8173290
## cToT          1.7726131   1.5533064   2.0606988
## Taskallo_2    1.2207438   0.6539765   2.1196508
## Taskallo_3    0.5067341   0.3619497   0.6950428
## Taskego_1     0.7398900   0.5305070   1.0052777
```

```r
grpef(P_7)
```

```
##
##
## Table: Estimates with 95% credibility limits
##
## fixef         re_factor       center       lower       upper
## -----------  ----------  ----------  ----------  ----------
## Intercept    Part         0.5525471   0.3941567   0.7952082
## cToT         Part         0.2411365   0.1045807   0.4163128
## Intercept    Route        0.6506511   0.4464556   0.9400092
## Taskallo_2   Route        0.5235004   0.0874271   1.2007623
## Taskallo_3   Route        0.4120467   0.0470928   0.8848689
## Taskego_1    Route        0.4610276   0.1991460   0.7875025
```

While there is substantial variation in how routes effect wall contacts conditional on tasks, the average association with ToT practically remains the same and the same holds for teh individual variation.

```r
left_union(fixef(P_7) %>%
             filter(fixef  %in% c("Intercept", "cToT")) %>%
             discard_redundant(),
           grpef(P_7) %>%
             filter(re_factor == "Part") %>%
             rename_if(is.numeric, funs(str_c(., "_sd"))) %>%
             discard_redundant()
)
```

```
## # A tibble: 2 x 7
##   fixef     center  lower upper center_sd lower_sd upper_sd
```

```
##   <fct>       <dbl>  <dbl> <dbl>     <dbl>    <dbl>    <dbl>
## 1 Intercept  0.221 -0.189 0.597     0.553    0.394    0.795
## 2 cToT       0.572  0.440 0.723     0.241    0.105    0.416
```

***Model selection***

Does M_7 fit the data better than M_5?

```
waic(M_5)
```

```
## Warning: 300 (19.0%) p_waic estimates greater than 0.4. We recommend trying
## loo instead.
```

```
##
## Computed from 4000 by 1582 log-likelihood matrix
##
##           Estimate   SE
## elpd_waic  -1999.4 38.1
## p_waic       416.4 11.2
## waic        3998.7 76.2
```

```
## Warning: 300 (19.0%) p_waic estimates greater than 0.4. We recommend trying
## loo instead.
```

```
waic(M_7)
```

```
##
## Computed from 4000 by 1582 log-likelihood matrix
##
##           Estimate   SE
## elpd_waic  -2180.5 47.4
## p_waic        76.8  6.7
## waic        4361.1 94.8
```

```
## Warning: 32 (2.0%) p_waic estimates greater than 0.4. We recommend trying
## loo instead.
```

It does.

## Visualization of results

```
load("M_AK_7.Rda")
```

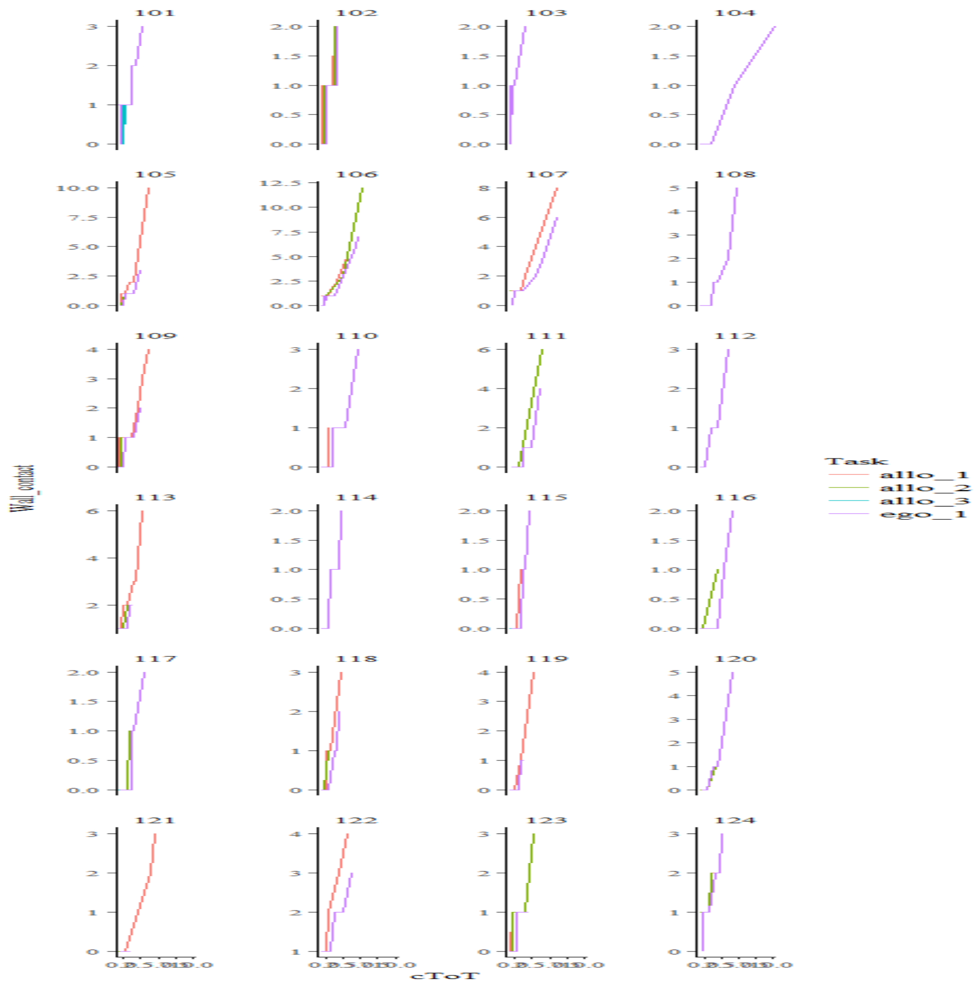
```
AK18_pred <-
  AK18 %>%
  mutate(Route = 1)
```

```
T_predict_Route1 <-
  predict(M_7, newdata = AK18_pred) %>%
  mutate(pred = "Rt_1")
```

```
T_predict_Route1 %>%
  left_join(AK18) %>%
  ggplot(aes(x = cToT, y = Wall_contact, col = Task, col = pred)) +
  facet_wrap(~Part, ncol = 4, scale = "free_y") +
  geom_line(aes(y = center))
```

```
## Joining, by = "Obs"
```

```
## Warning: The plyr::rename operation has created duplicates for the
## following name(s): (`colour`)
```




This confirms that the number of wall contact is positively associated with ToT for all and every participant. However, one issue remains to discuss: could the relationship purely be caused by time for error recovery. That would be the case, if a wall contact causes a severe delay in task completion.

## Additional Plotting

### Individual differences

````
```{r}
AK18 %>%
  ggplot(aes(x = Part, y = ToT)) +
  geom_boxplot()
```
````

### Differences between conditions – population level

````
```{r}
AK18 %>%
  ggplot(aes(x = Task, y = ToT)) +
  geom_col()
```

```{r}
AK18 %>%
  ggplot(aes(x = View, y = ToT)) +
  geom_col()
```

```{r}
AK18 %>%
  ggplot(aes(x = Task, y = Wall_contact)) +
  geom_col()
```

```{r}
AK18 %>%
  ggplot(aes(x = View, y = Wall_contact)) +
  geom_col()
```
````

### Differences between conditions – individual levels

````
```{r}
  AK18 %>%
  ggplot(aes(x = View,col=View, y = ToT)) +
  geom_col()+
  facet_wrap(~Part, nrow=4)
```

```{r}
  AK18 %>%
  ggplot(aes(x = View,col=View, y = Wall_contact)) +
````
````

```
  geom_col()+
  facet_wrap(~Part, nrow=4)
```

```{r}
  AK18 %>%
  ggplot(aes(x = Task,col=View, y = Wall_contact)) +
  geom_col()+
  facet_wrap(~Part, nrow=4)
```

**Interaction effect reoute by task**

```{r}

AK18 %>%

  ggplot() +

  aes(x = Route, y = Wall_contact, color = View) +

  geom_line(aes(group = View)) +

  geom_point()

```

## Appendix I - Expert Talks

### Expert Meeting I

Table I1

*Expert I*

| Citation Original | Category of Content |
|---|---|
| Bronchoskopie war einfach…damit ist man in die Endoskopie gegangen, da hat man das Ding in die Hand gedrückt gekriegt, in der Lungenfachabteilung. Du hast ja immer jemanden neben dir gehabt und dann bist du direkt da angefangen. | Training |
| Bei uns ist noch keinem übel geworden. | Sickness, Comparison to real setting |
| Das entspricht nicht der Realität, das ist schon mal klar. (Bezug auf die Android Cam) | Comparison to real bronchoscopy |
| Das wird ja unscharf. Das ist ja bei nem normalen Endoskop nicht. Dieser Unscharfe-effekt muss natürlich weg, sonst kann man damit nicht arbeiten | EndoProto - endoscope |
| Dieses Teil ist ja vorne dran, aber wir spielen ja hinten damit, mit den Reglern und dann muss man in den Oberlappen rein und so und dafür braucht man dieses Instrumentarium, diese Regelelemente. Man kann es ja nach oben und unten und den Rest macht man ja durch Drehungen mit der Hand. | Endoscope, handling |
| Ich drehe mit der Hand halt und wenn es ganz schwierig ist, kommt auch eine gewisse Drehung vom Oberkörper mit rein, weil wenn du in den Oberlappen rein willst, musst du manchmal extreme abwinkeln. Sonst siehst du ja nicht alle Ostien. Ich bronchoskopier ja nicht von vorne, sonst von hinten, vom Kopf her, vom Kopfende. Wenn vom Körper bronchoskopiert wird, ist das Bronchialsystem ja umgekehrt. Dann ist rechts links und links rechts. | Elbow and body posture (Questioning comparing simulator) |
| Kann gut sein, dass der Oberkörper mit kommt, habe ich noch nicht drauf geachtet. Weiß ich nicht. | Body twisting |
| Hier sieht man auch dass er das nur mit dem Daumen bewegt und mit der Handrotation. | Dr shows me a video of a bronchoscopy |
| Vorstufe des Simulators….aber das Problem ist ja: es wird unscharf. Es müsste zumindest ne Kamera sein, die in der Nähe nicht unscharf wird. | EndoProto |
| Ja, das könnte vielleicht sein. Wenn man sag, die müssen halt durch die Dings da durch gehen und die müssen versuchen das zu drehen. Aber richtig realistisch ist das hier so nicht. | Fingertips, maneuvering with the fingers, training with the EndoProto |
| (als Krankenschwester gearbeitet und vielfach bei Bronchoskopien zugesehen): Wenn Sie jetzt vom Handling ausgehen, etwas dicker. (das Kabel) | Comment of wife |
| Das ist hier zu klein. | Diameter of openings |

| | |
|---|---|
| Das geht irgendwann nicht mehr weiter, weil das Ding zu groß für die kleinen Bronchien ist. Die verzweigen sich ja immer mehr, die werden immer kleiner. | in the EndoProto for allocentric – Bronchia |
| Im Grunde sollte man hier ne Grundfertigkeit üben, bevor man dann an den Simulator mit Instrumentarium geht. Ist richtig? | |
| Das Problem ist immer diese Unschärfe. Das verwirrt. | [showing wall and tasks for egocentric] diameter for holes |
| Das ist ja gar nicht so verkehrt, von Anfang an auf den Bildschirm zu gucken, weil das ist ja das, was ihr letzendlich machen must. [Expert 1 bestätigt: Ja klar. ]<br>Ehefrau: Und das Handling dann. …<br>N dünnes Bronchoskop ist von der dicke auch nicht wahnsinnig mehr, vielleicht ne hälfte aber das ist ja nicht das entscheidende. Es ist halt nur von der Unschärfe nicht zu gebrauchen. | Comment of wife Thickness of the Endoscope-cable |
| Das ist realistischer.<br>Das andere war ziemlich ziemlich klein. | Diameter of holes egocentric |
| Aber trotzdem, es ist fast unmöglich darein zu kommen, in die Löcher. Irgendwann hat mans geschafft, aber es ist nicht so das mans bewusst geschafft hat. Man stochert rum und irgendwann trifft man mal. Das entspricht nicht der Realität. Es ist nicht richtig anzusteuern. | Critique on EndoProto |
| Du kannst nach oben, unten, kannst drehen…es ist schwierig | Dr.  tests EndoProto |
| Als Einstieg klar, aber mit ner Kamera die nicht unscharf wird, in der Nähe. | EndoProto |
| Der Druck entsteht schon, wenn man hier über die Zunge geht, über den Nasen-Rachen-Raum. Da hat man schon Widerstand. … Bis es dann letztendlich nicht mehr weiter geht. Natürlich spüre ich was, aber das Bronchoskop liegt ja frei in dem Rohr in den Ostien. Der Widerstand entsteht hauptsächlich oben. Manchmal nimmt man auch ein bisschen Gel, damit es besser flutscht über die Zunge, dann gehts ja in die Stimmritze, über die Luftröhre, die ist ja so breit, da hast du eigentlich Überhaupt kein Widerstand…dann gehst du ins rechte Bronchialsystem rein, winkelst ab, gehst weiter nach unten, Unterlappen, guckst dir die ganzen Oszien an…aber das ist kein richtiger Widerstand, weil das bronchoskop liegt ja in der Regel frei und dann hörts halt irgendwann auf. | Feedback of wire camera compared to real setting |
| Das ist ja hier die einfachste vorstufe einer Endoskopie. Wenn Sie das so machen wollen, müssen das zumindest so hinkriege, das die Kamera scharf ist und sie das anjustieren können… Der Rest ist ja: Das Bronchoskop geht nach oben und nach unten un der Rest muss durch Drehung erzielt werden, geht ja nicht anders. | Professional Bronchoscopy |
| Firma die Endoskope herstellt: Olympus führend. | Professional Endoscopes |
| wife.: Was ist wenn sie noch eine zweite Box bauen, und das man erst nur nach rechts und nach links gehen muss, vom Handling her!? | EndoProto |
| Als ersten Einstieg um erst mal so bildlich irgendwo etwas | EndoProto |

| | |
|---|---|
| ruchzuschieben könnte ich mir das vorstellen, aber es muss scharf sein. Und son basic Instrumentarium wäre ja nicht schlecht | |
| Hier sieht man den Oberlappen (video demo) und manchmal kommt man da besser rein, wenn man von vorne reingeht. | Professional bronchoscopy |
| In der Regel ist es so das man da ruhig bei steht, da macht man keine Verrenkung. Das macht man alles so mit der Hand. Und hier (Box) ist das nicht realistisch, wenn man das nur mit der Hand so dreht. | Posture and movement of upper body |
| Wie Videoendoskopie funktionieren könnte, als ersten Einstieg, als kleine Fingerübung. So wäre das gedacht, oder? Aber dann muss zumindest die Kamera scharf sein. | EndoProto |
| Je nachdem wie stark empfindlich die Bronchialschleimhäute sind. Wenn man da dranstößt, kann man Blutungen auslösen, aber die sind nicht dramatisch. Nicht zum Beispiel wie beim Durchstoßen in der Koloskopie. Durchstoßen kann man nicht.<br>Blutungen: Das trübt halt nachher die Sicht und man muss dann mit Kochsalz spülen, aber richtig verletzen so nicht.<br><br>Es ist ein relativ stabiles System. Mit nem starren Bronchoskop klar.<br>Die verbluten ja nicht, die ersticken die Leute. Und dafür ist das starre Bronchoskop da.<br><br>Und deshalb ist es auch logisch das in den großen Lungenkliniken von oben bronchoskopieren. | Damage and risk |
| Für das erste hands-on, wäre son Ding (Box) gar nicht schlecht, wenn es besser bedienbar wäre. Wenn es zumindest annähernd realistische Bedingungen bietet, vom Handling her und natürlich von der Schärfe des Bildes. Das müsste man so fordern. | Evaluation of the EndoProto |
| Frage nach timing Beschränkung:<br>Ne, nur weil der nächste Patient kommt.<br>Sonst, gibt es keinen Zeitdruck. Man muss nur schnell werden, wenn e seine massive Blutung gibt, aber das will man hier nicht erleben. … und was sagen die Bronchoskopie Schwestern dazu.<br>Aber vom Ablauf her: Ne. Man sediert die Leute so lange es nötig ist und dann werden sie irgendwann wach. | Time restriction |
| Das Bronchialsystem ist fest, das ist starr, da kann man sich drin bewegen. Da ist schon besser, son starres System wie hier. Da kann man sich drin bewegen. | Material rigidity |
| Wie gesagt, Simulatoren, das ist für mich ein Fremdwort.Wie alt bin ich.. | Simulators |
| Ich bin mit der Bronchoskopie erst in Kontakt gekommen als ich nach der Internisten-Ausbildung ins Teilgebiet gegangen bin, in ne Lungenheilkunde. Da hat man gastroskopiert, koloskopiert…Ich weiß, da waren Übertragungen aus der Endoskopie, die uns gezeigt wurden. Und ich meine da war auch irgendwie ne Puppe. | Training and education |

*Note*: Expert I is doctor for internal medicine, pneumology, allergology, somnology and environmental medicine. Personal communication took place at the 23th of June 2018.

**Expert Meeting II**

Table I2

*Expert II*

| Paraphrasing* | Category of Content |
|---|---|
| The real size in a bronchoscopy lies between the two of the walls. Or using first the broad ones and then the thin ones. | Size of openings in walls |
| The size of the camera wire is more like one endoscope for children. | Camera |
| It is a nice instrument, for training hand-eye-coordination, but the focus has to be nearer.<br>The overview is missing, which leads to disorientation. | Camera function and distance between openings! |
| It is rather nearly not possible to harm a patient in a bronchoscopy, because you do not pierce through the human tissue as for example in a coloscopy. | Variables: Wall contacts |
| It is possible to let them train and investigate if they get faster. | Variable: time |
| Het is wel goed een faste box. De bronchia zijn ook fast, niet zo als een colon. Er zit beweging in. | Material of EndoProto |
| Citation Original | |
| Er zijn een paar dingen die je met de bronchosckopie moet denken. Je kunt hem op het zitten doen. Je kunt hem op liggen doen. Op dat moment verandert ontmiddelijk het anatomie, so als je naar de patient kijkt. (laat het zien met echte instrumentarium uit de simroom). En ik moet daar uitkomen, in de long en dat is wel belangrijk. | Real bronchoscopy and visualization (shows me the procedure!) |
| Ik kann met de scope alle kanten op. Wat zij (de studenten) doen is, Ze gaan zo.<br>Je moet alleen in de introductie direct kijken, maar naar dit moment moet je niet meer erna kijken, maar on de scherm.<br><br>En dann moet ik deze doen – hand-oog coordinatie, maar die doe ik niet meer op de patient maar on de scherm. Dat is belangrijk. Maar wat ze fout doen….Ze letten er niet meer goed op hoe de draaimechanisme is. En wat zit ik nu…. Dus, de introductie moet goet zijn | Real bronchoscopy & training:<br><br>Introduction, hand-eye-coordination<br><br>Summary:<br>Students have problems to shift their vision to the screen and maneuver the scope during viewing the screen |
| Dus je introductie moet goed zijn. Als je binnekomt moet je meteen een goed overzicht hebben, waar ik heen moet. | Comparison EndoProto |
| [demonstratie op de menselijke pop met endoscope] En daan | Anatomy |

| | |
|---|---|
| moet<br>Ik er kennis van de anatomie hebbben.<br>En dann is er daar een takje en daar en takje, en dat is maar zoo een klein gebied war ik op zie. | |
| Die is dicker dan jou scope hier. Dat is een normale scope zo als wij die bij volwassene mensen gebruiken. Bij kinderen is die nog wat dünner. Dann kommt die ongeveer met het niveau wat jej hebt. | Scope |
| Dat betekend de hele anatomie draait om 180°. En nu ga ik anders in. | Anatomy and position of bronchoscopy: from head or body |
| Ons doel van dit endovak is nu, dat zee en beetje leeren van de basis. Wa is nu een bronchgoscopy, wat is nu en gastroscopy, wat is een coloscopy. Ik maak von hun geen scopisten. Ik gaa niet zeggen: Je hebt het examen gehaalt. Je kann nu alle bronchoscopien doen.<br>Maar dat ze, als ze in de ziekenguis gaan aan de patient kunnen uitleggen hoe en wat. Wat is een bronchoscopy…Dat je begrijpen wat de longartz aan het doen is en wat het betekend voor de patient. | Educational purpose |
| (demonstratie echte bronchoscopy) Als ze geen overzicht hebben, dann zijn eindlos lang bezig om er heen te komen. | Comparison EndoProto and real bronchoscopy |
| Maar als je nu weet hoe je ding draait. En dan zie je hier nu op het realistisch model hoe dicht alles bij elkaar zit, waar die takjes zitten.<br><br>De anatomie moeten ze kennen en ze moeter er in kunnen. Ze moeten die dakken in kunnen. | Overview, knowledge, skills, training |
| EN wat je kunt oefenen, … je kunt in jouw box niet alleen de hand-oog coordinatie oefenen, maar da kunt je ook met de hapjes doen. | EndoProto |
| (laat een biopteur zien)<br>Je hebt geen echte scope.<br>Maar dat kann je ook laten oefenen. | training |
| Dat is niet naar de patient kijken, maar permanent naar het beeld kijken en permanent weten wat je aan het doen bent. | Bronchoscopic procedure |
| Ja, je moet hier blijven zitten en niet darnatoe draaien wat je ziet. Dus, dat leer ik ze ook, da ergonomie. Ze moeten dus gewoon als ze bezig zijn…Je bijvt rustig staan.<br>Anders is het een economische fout. Je kreegt rugklachten en alles van.<br>Ik zie hier ook sommige, die gaan zo mee. Nee. Blijv rechtop staan. | Body posture - no twist |
| Ze moeten zich bewust gemaakt worden. Blijv rechtop staan. Denk aan je ergonomie. | Training and students |

| | |
|---|---|
| Donker is die ook, maar als je dichterbij kommt is die goed. | Camera of the scope |
| Ik denk maar, dat je bij die kartjes, alle openingen kleiner, dichter bij elkaar moet zetten en dann kann je weer merdere kanten, wat breederen en wat smallere openingen. Dat is prima. Maar, van broen, naar blauw naar geel. En ze moeten zo maneuvrieren op grond van het beeld wat ze zien op de comuter. Prima. Alleen dat maneuvrieren is wat lastiger bij jouw dit niet hebt, die stuuring system. Die is een hele stuk, ik kann er alle kanten op (laat zien echte brochoscop). Maar die van jou is een hele dünne draat en die gaat die doen (nachgeben). Die is een probleem. Of je moet er een wat stijvere mantel omheen doen. Zeg maar, een weer achter het system. Als je hem stuurt, dat hij echt die kant op gaat. | EndoProto |
| Begin daar eerst mee, dan maak je jezelfs makkelijk. (De introductie van bronchoscopie). Dan kom je hier, bij alle die oszie, en daar kann je alle kanten op. Maar bij je heb je geen overzicht. | Bronchoscopic procedure |
| Die benaderen van belangrijk deel dat system uit het ziekenhuis. | Simulators |
| Voor het oefenen is dat prima. Maar met jouw camera, met dit slangetje, dit ist e supple. Je kunt niet sturen. In kan alleen in. | Box - camera |
| Ja, daar kan je hem heel goed gebruiken, voor het insturen, maar alleen voor dit stukje. Mar darnaa is het afstand van je camera te groot, dan kann je hem niet meer sturen. | Box-camera, insertion |
| Het is heel leuk en het is goed bedacht, maar je loopt tegen de praktijk aan. Tegen hoe werkt zon ding. Voor het begin is het prima. Maar als ik verder ga met jou camera, kan het niet meer, dann moet je het stugger, stijver maken. | Box evaluation, camera |

*Note*: The personal communication with Expert II took place at the 2nd of July 2018 in the University of Twente in Enschede. Expert II is specialized in technical medicine. Some of his working fields are life support, endoscopic skills, injections, punctures and catheterize, cardiorespiratory system and technology.
*Because of technical problems, the recording of the first part of the communication broke down. Therefore, the paraphrasing is based on memory retrieval of the communication and does not represent perfect citations.

**Appendix J – Comparison between EndoProto - BRONCH Mentor – Reality**

| Category | EndoProto (Prototype) | BRONCH Mentor (Simulator) | Bronchoscopy (Reality) |
|---|---|---|---|
| Set-up for Participants: | | | |
| - Posture | Sitting or standing (adjustable) | Standing | Comparable to simulator - depending |
| - Hand-elbow | Personal and depending on endoscope – less realistic | Realistic: elbow near the ribs, arm tacked in, just hand twists, | Comparable to simulator – but more flexible adjustments necessary |
| - Fingers | Different handling based on missing control system (endoscope) | comparable: Sim-reality | comparable: Sim-reality |
| Economics | | | |
| - Budget | Low | High | Highest |
| - mobility | Manually movable | Locally fixed (Express available) | Patient-depending |
| Risk | None! | None! | Health-lifedepending! |
| Tasks | Task 1 – Allocentric (non-bronchoscopic dexterity task)<br><br>Task 2 – Egocentric (endoscopic view task) | 1. Basic Scope Manipulation (BSM)<br><br>2. Anatomical Guidance (AG) | |
| Routes: | | | |
| - Task 1 | Randomly assigned | Randomly assigned | Reality setting - never 100% controllable |
| - Task 2 | Randomly assigned | Guided – self directed | |
| Variables | | | |
| - ToT | As performance parameter, Detailed and comparable measurements from the insertion of the endoscope until crossing a route | As performance parameter, Detailed measurements from insertion of the endoscope into the mouth until reaching the vocal chord and until reaching destinations | Time is less relevant for bronchoscopic procedure per se, it is more relevant for organizational and economic factors |
| - Wall_contacts | Appear frequently, because of small scaling. Slightly acoustic and haptic feedback when hitting a wall. | Task 1 (BSM): pre-cuing (red alert) when the scope moves near the walls. Acoustic error noise and visual blinking | Less relevant in realistic setting, because the risk to harm a patient is rather low. The physician |

| | | | |
|---|---|---|---|
| | Task 2: broader scaling leads to lower rates. Lower acoustic and visual feedback, because of a loss of camera visibility | screen when it hits a wall. | gets feedback comparable to a simulator task when the bronchia ends. |
| Feedback | Task 1: visible, hearable<br><br>Task 2: less visible!, slightly hearable | | |
| Scaling and composition | Scaling for task 1 is not comparable to a real setting<br><br>Scaling of task 2 is more realistic! But for an even more realistic setting, it should become smaller in the depth.<br><br>BroPro as a stir system box seems adequate for training | No expert opinion available!<br>But, compared to the endoscopic BroPro task, the simulator tasks are of much bigger scale – probably simplified for training purposes | |
| Endoscopic camera | | | |
| - resolution | Comparable | Comparable and partly even better! | Comparable |
| - focus | Lack of quality! | Comparable | Comparable |