MASTER THESIS

TOWARDS PREDICTING BRONCHOSCOPIC SKILL ACQUISITION USING BASIC BRONCHOSCOPIC SIMULATOR TASKS

Marlise W. Westerhof

FACULTY OF BEHAVIORAL, MANAGAMENT AND SOCIAL SCIENCES (BMS)

DEPARTMENT OF COGNITIVE PSYCHOLOGY AND ERGONOMICS (CPE)

30/08/2018

Supervisors
dr. Martin Schmettow (CPE)
dr. Marleen Groenier (TM)

UNIVERSITY OF TWENTE.

Abstract

Introduction: Bronchoscopy is a very complex skill, and surgeons seem to differ at learning this skill. In order to minimize patient risk, it is crucial to only enable proficient surgeons to perform bronchoscopies. However, adequate selection and assessment methods do not exist at present. A promising solution would be to include simulator-base prediction of surgical performance. The current study aims at exploring whether bronchoscopy simulator performance can be used to individually assess prospective trainees based on the maximum performance they are expected to reach after training, based on learning curves. By means of this assessment method, it is explored which individuals are more suitable for becoming a bronchoscopist and which individuals will experience more difficulties. First, it was analysed whether learning curves exist in the performance of basic bronchoscopy simulator tasks. Moreover, it was explored whether individual differences exist in acquiring bronchoscopic skills to confirm the need for individualized training. Finally, it was tested whether the simulator tasks together form a reliable test to select prospective trainees based on their performance.

Method: Nineteen students took part in a repeated-measure study. Participants had to perform two or three basic bronchoscopy simulator tasks fifteen times each on the GI-BRONCH Mentor ™. Exponential learning curves have been estimated per task and participant. Maximum performance, based on time-on-task, was the primary parameter of interest. Correlations between the asymptote for each task on population-level were calculated.

Results: All participants improved performance over time for all three tasks, as the time to complete each trial decreased. Moreover, individual differences in acquiring bronchoscopic skills have been found. The internal consistency between the three simulator tasks was low and highly uncertain.

Discussion: The internal consistency of the three simulator tasks was low. However, all findings were highly uncertain due to a small sample size. A possible limitation of the current study is the lack of an external criterion to validate simulator performance. Moreover, using time-on-task as a performance measure might not be the most suitable performance variable for bronchoscopic skills. Finally, the Basic Scope Manipulation and Step-by-Step Diagnostic Manoeuvres tasks might be too easy to able to discriminate between the levels of psychomotor ability. We encourage medical education institutes to implement and further develop performance-based selection. A useful extension to the current study would be to include expert performance as cut-off scores. The most promising task as predictor of future surgical performance is the Guided Anatomical Navigation, however results were highly uncertain. On the long run, it is recommended to explore the relationship between basic simulator tasks and the whole bronchoscopic procedures.

Table of Content

The number of performed minimally invasive surgeries (MIS) has rapidly increased over the last years (Tsui, Klein, & Garabrant, 2013). A frequently performed MIS procedure is bronchoscopy. Bronchoscopy enables visual inspection and tissue sampling from the airways by inserting a bronchoscope through the nose or mouth (Evison & Munavvar, 2016; Shah, 2003). MIS offers several short-term and long-term advantages over traditional open surgery, such as less blood loss, reduced pain and shorter length of hospital stay (Pache, Hubner, Jurt, Demartines, & Grass, 2017). However, in contrast to traditional surgery, surgeons need to master different complex skills while performing MIS procedures (Gallagher, Leonard, & Traynor, 2009). Challenges in learning MIS include interpreting 3D information about the human body from 2D images presented on a screen, a lack of haptic feedback, visual-spatial ability and psychomotor skills (Gallagher et al., 2009). As a consequence, surgeons showed prolonged learning curves, in contrast to conventional open surgery (Fuchs, 2002). Thus, patient safety during MIS might be endangered since surgeons may not have achieved competency (yet).

It has been found that surgical residents differ in the amount of performed procedures necessary to acquire proficiency in MIS skills, with estimations that some individuals will not even reach proficiency at all despite practice (Grantcharov & Funch-Jensen, 2009; Louridas et al., 2017). Identifying the expected performance level and skill acquisition rate of residents is crucial to ensure effective individualized training. Residents who show proficiency could be allowed for taking the next step in surgery training, while those who experience difficulties may benefit from either providing intensified training or the recommendation to pursue a different career (Hofstad et al., 2013). To ensure learning efficiency, it is crucial to predict applicant's future surgical performance at an early stage in surgical training.

However, international accepted benchmarks on how to select, train and assess (future) pulmonologists do not exist at present (Ernst & Herth, 2013). Current selection of

prospective trainees still relies on cognitive selection tools, such as pre-university performance and aptitude tests, and non-cognitive selection tools including unstructured interviews, reference letters, personal references and psychometric questionnaires (Epstein & Hundert, 2002). However, as current selection tools do not align with skills necessary to perform bronchoscopy, the validity of these objective selection methods is widely criticized (Fielding, Maldonado, & Murgu, 2014).

Moreover, surgical trainees are still being inconsistently assessed based on the number of performed MIS procedures and supervisors' subjective evaluations (Fielding, Maldonado, & Murgu, 2014). For example, according to the American College of Chest Physicians, pulmonary trainees should perform at least 100 supervised flexible bronchoscopies to achieve basic competency (Ernst, Silvestri, & Johnstone, 2003). However, according to Konge et al. (2011), such threshold numbers are neither scientifically proven nor universally accepted by bronchoscopists. While one trainee might need only 25 simulation procedures to reach a pre-set performance level, another one still struggles with acquiring technical bronchoscopy skills after the 50th procedure (Wahidi et al., 2010). By developing a method for predicting skill acquisition, it may also become possible to predict the amount of procedures an individual needs before mastering technical bronchoscopic skills in the future.

An alternative approach to the current selection and assessment in bronchoscopy education would be to include simulator-based prediction of surgical performance and skill acquisition. Besides providing bronchoscopy training in a zero-risk environment, bronchoscopy simulators enable new means to tailor surgery training to the needs of an individual. Performance metrics provided by simulators, such as procedural time and wall contact, could provide objective and structured assessments about a trainee's proficiency level, track the training progress over time and provide continuous feedback and motivation for learning and evaluating the training curricula (Epstein & Hundert, 2002; Grantcharov,

2008). In addition, feedback based on performance metrics enhances MIS learning (Samia, Khan, Lawrence, & Delaney, 2013). However, little research has been conducted about using simulators as assessment and selection instruments within the field of bronchoscopy (Ernst & Herth, 2013).

The present study explores individual differences in learning bronchoscopy while aiming for a valid and reliable tool to predict future bronchoscopic performance among novices. We aim at exploring whether bronchoscopy simulator performance can be used to individually assess prospective trainees based on the maximum performance they are expected to reach after training, based on learning curves. By means of this assessment method, we want to identify which individuals are more suitable for becoming a bronchoscopist and which individuals will experience more difficulties. First, we are interested in whether learning curves exist in the performance of basic bronchoscopy simulator tasks. Moreover, we will explore whether individual differences exist in acquiring bronchoscopic skills to confirm the need for individualized training. Finally, we will explore whether the simulator tasks together form a reliable test to select prospective trainees based on their performance.

1. Background

1.1 Risks and Challenges of Bronchoscopy

Flexible bronchoscopy is a frequently performed MIS procedure in respiratory medicine, which enables visual inspection and tissue sampling from the respiratory tract by inserting a bronchoscope through the nose or mouth (Evison & Munavvar, 2016; Shah, 2003). A flexible bronchoscope contains a fibreoptic system which transmits live images captured by the tip of the instrument to a video screen. The bronchoscope can be controlled by moving the lever to flex or extend the bronchoscope's tip and by gently turning the wrist for clockwise and anti-clockwise rotation of the scope. If an abnormality is detected, biopsies can be taken by

inserting sampling tools, such as brushes, forceps and needles, into the working channel (Leiten, Martinse, Bakke, Eagan & Gronseth, 2016) (see Figure 1).

Figure 1. Visualization of a flexible bronchoscopic procedure. Retrieved from http://iacharitygolf.com/project/electromagnetic-bronchoscopy/ on 16-08-2018

1.1.1 Complications. Flexible bronchoscopy is a generally save procedure, but rare complications often have life threatening consequences. The most common complication during flexible bronchoscopy is hypoxemia, an abnormal low level of oxygen in the blood (Schramm et al., 2017; Shah, 2003). Other bronchoscopy-induced injuries include pneumothorax, pneumonia and intrabronchial bleedings, although these major complications occur in less than 1 % of the cases (Ernst, Silvestri, & Johnstone, 2003). Because lungs are one of the vital organs of the human body, exchanging gas between the lungs and the blood, it is crucial to avoid complications during and after bronchoscopy. However, as MIS in general is a very hard-to-learn skill, bronchoscopists with less experience showed higher complication rates, higher procedures time and higher amounts of sedation required than bronchoscopists with more experience (Stather, MacEachern, Chee, Dumoulin, & Tremblay, 2013). More

specific, novice bronchoscopists show increased complication rates when performing bronchoscopy during the acquisition phase, the first trimester, of bronchoscopist training (Ouellette, 2006). Notably, even after several years of MIS training complications persist among bronchoscopic surgeons (Wahidi et al., 2010).

1.1.2 Psychomotor ability. In contrast to traditional open surgery minimally invasive surgeries, such as bronchoscopy, require a completely different set of complex skills. During bronchoscopy, the surgeon needs to perform several complex actions simultaneously: withdrawing or advancing the shaft, flexing or extending the scope's tip, and rotating the scope's shaft and tip. Simultaneously, the surgeon has to observe a screen for visual feedback (Colley & Freund, 1997). During open surgery, the surgeon has direct visual and manual access to the operation side, whereas MIS-procedures, such as bronchoscopy, involve remote manipulation of instruments together with an indirect view of the surgical field provided by a camera (Fuchs, 2002). Interpreting 3D information about human anatomical structures from 2D images presented on a screen results in a lack of depth perception. Moreover, as the surgeon no longer manipulates the instrument directly, haptic feedback is reduced. Since bronchoscopy requires the integration of, among others, spatial perception, dexterity and muscle function, it is a very complex technical skill (Silvennoinen, Mecklin, Saariluoma, & Antikainen, 2009).

The success of a surgery depends to a large extent upon the use of adequate psychomotor skills by the surgeon. Psychomotor ability is defined as "the relatively innate potential to acquire psychomotor skill after practice" (Kaufman, Wiegand, & Tunick, 1987, p. 1). Psychomotor skills required to keep the bronchoscope in the centre of the airway to avoid scope-induced discomfort and injury include fine and gross motor functioning, hand-eye coordination, bimanual dexterity and spatial perception (Herth, Shah, & Gompelmann, 2017; Kaufman et al., 1987). Due to indirect vision and manipulation during bronchoscopic surgery,

depth perception and eye-hand coordination are two of the main challenges for surgeons (Breedveld, Stassen, Meijer, & Stassen, 1999). For this reason, these two components of psychomotor ability will be explained in more depth.

Compensation for insufficient depth cues. Due to indirect sight of the operative field, the surgeon's depth perception is reduced during bronchoscopy. Spatial perception is the ability to "judge the relations of objects in space, to judge their shapes and sizes, to manipulate them mentally, and to visualize the effects of putting them together of turning them around." (Breedveld, Stassen, Meijer, & Stassen, 1999, p. 3). Spatial perception is an important factor in psychomotor functioning in bronchoscopic skills, on the contrary, it is crucial during bronchoscopy to deal with a loss in depth perception (Fielding et al., 2014; Madan et al., 2005). Three-dimensional anatomical structures of the airways are presented on a two-dimensional video screen which results in a lack of depth indicators, such as pictorial cues, visuomotor cues and the comparison of pictures seen by the right and left eye. Additionally, bronchoscopists cannot use shadows as depth cues, since the tip of the bronchoscope includes a light source to increase brightness of the picture. The lack of depth-perception cues interfere with accurately determining spatial relations of the scope with respect to anatomical structures and performing. As a result, surgical residents require time-consuming and intensive training curricula to learn to compensate for the absence of depth cues (Breedveld et al., 1999).

Disturbed hand-eye coordination. As a consequence of reduced depth perception hand-eye coordination is impaired as well. Hand-eye coordination refers to "the ability to produce goal-oriented hand- and arm movements that are guided by visual information from the eyes" (Lee, Junghans, Ryan, Khuu, & Suttle, 2014, p. 51). Effective and efficient hand movements depend on accurate coordination between visual-, cognitive- and motor- systems. To plan, control and adapt hand- and arm movements sensory input of the environment is

required, which includes e.g. the shape, motion, orientation and size of objects. The visual cortex sends visual-spatial information to the motor cortex, which elicit muscle controlling signals to coordinate contractions hand- and arm muscles. Information about perceived positions of the hand and the target are used to send motor signals to the muscles (Breedveld et al., 1999; Lee et al., 2014). In contrast to direct observation of the instruments and operative field by looking down during open surgery, a surgeon indirectly observes the operative field via a camera view displayed on a screen during bronchoscopy. The distance between the operative field, displayed on a screen, and the surgeon's hands, makes it impossible to observe both the operative field and the hands simultaneously (Breedveld et al., 1999; Wentink, Breedveld, Meijer, & Stassen, 2000). In addition, the use of a long scope results in a lack of direct haptic feedback, the bronchoscope's camera has a different point of view than when the surgeon had direct visual access to the surgical field, and the scope as well as the screen magnify the operative field (Lee et al., 2014). Consequently, similar to compensate for insufficient depth cues, surgeons need intensive training to adapt to impaired eye-hand coordination during bronchoscopy (Konge et al., 2011). Thus, both impaired depth perception and eye-hand coordination suggest that bronchoscopy is a procedure of higher cognitive demand in contrast to traditional open surgery, which indicate the need of adequate long-term training.

1.2 Psychomotor Skill Acquisition in MIS

According to the theory of Fitts and Posner (1967), three stages are completed during psychomotor learning: a cognitive, an associative and an autonomous stage. During the cognitive stage, the surgeon relies on past experience, reasoning and instructions to cognitively understand the requirements of performing the task. However, findings of studies of students at the University of Twente gave no indication of a cognitive stage (Huijser, Schmettow, & Groenier, 2015; Kaschub, Schmettow, & Groenier, 2016; Arendt, Schmettow,

& Groenier, 2017). This might be due to the fact that the used simulator tasks were rather simple to understand, which could indicate that the cognitive stage is only applicable to complex procedures.

After gaining experience with the task, cognitive strategies will be improved during the associative stage. The surgeon works on reducing errors and making skill performance more efficient by optimizing and integrating strategies. According to Fitts and Posner, the ultimate goal of the learner is being able to perform the task at the autonomous stage: achieving advanced levels of performance by requiring little or no attention with few or no errors. Although bronchoscopists aim for as few errors as possible, it is crucial to keep optimizing strategies which means that total automatization of bronchoscopic procedures is not desirable. As a bronchoscopist gains more experience with bronchoscopy, less attention is required to perform these skills. Due to limited working memory's capacity, the automatization of motor skills frees up cognitive resources for other processes, such as communication with nurses or diagnosis of anatomical abnormalities (Gallagher et al., 2009).

Learning new skills should not only aim at achieving successful task performance, but an individual should also be able to apply these skills to a wide range of contexts. Successful learning of a new skill which enables both successful task performance and transferability to other situations, is called holistic skill acquisition. For instance, skills that a surgeon acquired by simulation-based training should be transferable to real-life surgery in the operating room. A study on motor sequence learning in acquiring complex motor skills by a student of the University of Twente emphasized the importance of variability in simulator tasks to improve the transferability of simulator-learned laparoscopic skills to real surgery (David, Schmettow, & Groenier, 2018). It was assumed that memorization of a specific motor sequence has a small role in the improvement of performing complex motor procedures, such as we have in bronchoscopy.

1.3 Predicting MIS-performance Using Learning Curves

The current study investigates how individuals learn complex bronchoscopic skills by exploring individual learning curves for different simulator tasks. According to Wanzel, Ward and Reznick (2002) learning curves are suitable as prediction and assessment tools for surgical performance. Learning curves enable quantitative exploration of an individual's learning process of skills over time, including previous experience, improvement and maximum performance. Rather than one-time performance measurements, learning curves focus on repeated measurements of performance over time. Previous research on learning curves in minimally invasive surgery mainly focused on visualizing (averaged) data using learning curves. In contrast, to our knowledge, statistical estimation of learning curves is actually rather new. An exponential learning curve model based on the law of practice was already applied to explore learning of laparoscopy in previous theses of students of the University of Twente (Huijser, Schmettow, & Groenier, 2015; Kaschub, Schmettow, & Groenier, 2016; Arendt, Schmettow, & Groenier, 2017)

Improvements in performance due to practice and training proceed in a non-linear way. During early trials, improvement rates in performance outcomes, such as reduced error and increased speed, will be high but slows down with increased practice, which is known as a learning curve (Heathcote, Brown, & Mewhort, 2000). Once an individual reached its maximum performance, additional training won't lead to performance increasement due to boundaries in performance caused by physical constraints, including neural integration time and motor response time. This is called a saturation effect: the more training is performed, the closer the level of performance gets to the individual performance limits and the less additional training adds (Schmettow, 2018b). These boundaries of performance measures result in non-linear relationships between predictors and outcome.

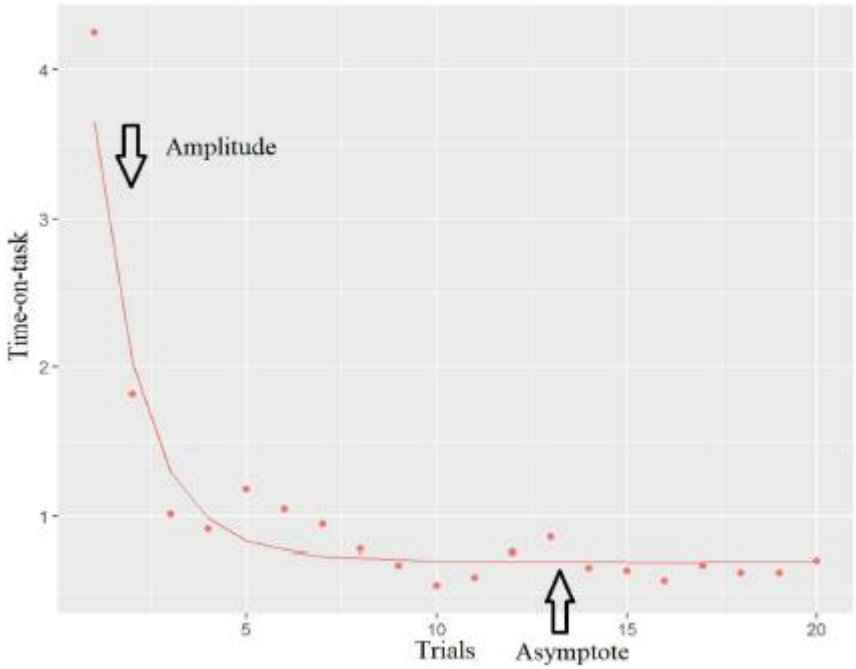


Figure 2. Exponential learning curve. Trials is depicted on the x-axis, whereas the y-axis shows time-on-task (Arendt et al., 2017).

Factors influencing the shape of a surgeon's learning curve include the nature of the task, experience, manual dexterity and anatomical knowledge (White, Rodger, & Tang, 2016). For instance, it has been found that minimally invasive procedures show longer learning curves than open surgical procedures (White et al., 2016). According to Heathcote, Brown and Mewhort (2000) the exponential law of practice learning curve's structure consists of tree parameters: amplitude (δ, amount of learning), rate (ρ, speed of learning) and asymptote (ω, maximum performance) (See Figure 2). Each of these parameters represents a person-specific learning parameter. Amplitude represents the individual's amount of improvement, so how much someone learns (initial performance minus asymptote). The rate represents the speed of learning; the faster an individual reaches maximum performance, the higher his or her rate. Finally, the asymptote represents the participant's expected maximum performance which will be reached after ongoing practice. Since we are interested in predicting whether someone will reach proficiency, the asymptote, or maximum performance, is our main parameter of interest.

Learning curves based on performance data could provide valuable information about surgical performance during simulator training. Simulator-based prediction of surgical performance and skill acquisition is a more structured and performance-based alternative to the current unstructured selection and assessment in bronchoscopy education. Besides providing training in a zero-risk environment, simulators enable new means to tailor surgery training to the needs of an individual. Individual learning curves based on performance metrics provided by simulators, such as procedural time and wall contact, could provide objective and structured assessments about a trainee's proficiency level, track the training progress over time and provide continuous feedback and motivation for learning and evaluating the training curricula (Epstein & Hundert, 2002; Grantcharov, 2008). Unfortunately, simulator training is highly expensive which impede the global implementation of simulator-based training and assessment. In addition, since most evidence on minimally invasive surgery is derived from laparoscopic surgery, less is known about using simulators as assessment and selection instruments within the field of bronchoscopy (Ernst & Herth, 2013). This indicates the need for research on using simulators as selection and assessment methods.

1.4 Individual Differences in Learning and Performing Bronchoscopy

Surgical trainees differ in the learning rate of acquiring technical skills, while some students are unable to achieve competence despite training. Recent studies were able to discriminate subsets of students who differed in learning curves patterns for laparoscopic skills training (Alvand, Auplish, Khan, Gill, & Rees, 2011; Grantcharov & Funch-Jensen, 2009; Louridas et al., 2017; Schijven & Jakimowicz, 2004). Studies differed in among others in tasks and performance parameters. High performers (6-22%), students who were hypothesized to have an innate psychomotor ability, acquired skills and reached proficiency after just a few trials. However, the majority of trainees (37-70%) were moderate performers who improved their performance by practice and eventually will reach proficiency. A striking

finding is that 8-24% of the individuals, low performers, experienced difficulties to learn technical skills and were unable to achieve an acceptable performance level despite training. More precisely, Grantcharov and Fuch-Jensen (2009) identified percentages of high, moderate and low performers of 5.4, 70.3 and 24.3, respectively, whereas Louridas et al. (2017) identified percentages of 21, 71 and 8, respectively.

At present, evidence for individual differences in surgical skills, as well as for the cost-effectiveness, improved patient safety and efficacy of simulator-based training within other medical fields exceeds the amount of studies covering simulator-based training in bronchoscopy (Pastis et al., 2014). In general, most evidence on minimally invasive surgery skill acquisition is derived from laparoscopic surgery, whereas bronchoscopic skill acquisition is far less covered. Therefore, careful translations from laparoscopic research to bronchoscopy are necessary. Nevertheless, individual differences within technical bronchoscopic skills have been found for learning bronchoscopic skills as well. A study of Wahidi et al. (2010) on competency metrics in learning bronchoscopy among novices bronchoscopy fellows found that participants still showed significant differences in performance at the 50th bronchoscopy. In addition, Dalal and colleagues (2011) demonstrated that a wide variation exists in learning curves among novice resident performing fiberoptic upper airway endoscopy. The number of attempts necessary to achieve proficiency varied between 27 and 58. Moreover, corresponding to studies within the field of laparoscopy, Dalal and colleagues (2010) were able to discriminate between three learning curve patterns: Participants who reached proficiency after minimal amount of training, those who did not achieve proficiency but showed improvement after intensified training and those who neither improved performance nor reached proficiency despite training. Since a large variation exists in the learning curves of bronchoscopy residents, bronchoscopy education should aim at competence-based learning, which enables education tailored to individual needs.

Individual differences in MIS-performance could be due to differences in innate abilities. It has been found that cognitive abilities, such as perceptual ability and visuospatial ability, are related to the learning curve of surgeons for minimally invasive surgery performance (Groenier, Schraagen, Miedema, & Broeders, 2014; Luursema, Buzink, Verwey, & Jakimowicz, 2010; Maan, Maan, Darzi, & Aggarwal, 2012). However, individuals show differences in the visual-spatial ability they possess (Madan et al., 2005). In addition, Gallagher et al (2003) were able to identify different levels of laparoscopic basic psychomotor skills among experienced surgeons by means of basic laparoscopic tasks. Notably, among experienced surgeons (> 50 laparoscopic procedures) between 2 and 12% of surgeons showed deficits in basic psychomotor skills. Psychomotor ability affects the rate at which actions automatize: it has been found that when task complexity increases, surgical trainees with lower psychomotor ability experience more difficulties in automating their performance (Jabbour, Reihsen, Sweet, & Sidman, 2011). To compensate for lesser psychomotor ability, an individual needs more practice in order to reach necessary levels of competence in psychomotor skills. But the acquisition of psychomotor skills is unattainable if an individual does not possess psychomotor ability. According to Jabbour et al (2011), prospective bronchoscopic surgical trainees with low levels of psychomotor ability, who will need high amounts of training which extend regular training curricula, should be advised to pursue different careers during screening of prospective surgeons.

1.5 History of Bronchoscopy Education

During the last years, bronchoscopy education underwent major developments to ensure patient safety and education efficiency (Fielding et al., 2014). Formal training for surgeons was achieved through observing procedures before acquiring skills and competence by training on real patients, following the apprenticeship model. However, because of patient safety concerns training on patients is now undesirable (Fielding et al., 2014). Moreover,

since surgery training curricula strive for more efficient education, there is a rising interest in the prediction of surgical skill acquisition and performance. On the contrary, as previously discussed, surgeons differ in the amount of innate ability they possess and the acquisition of technical surgical skills. Early prediction of surgical skill performance and acquisition enables both training tailored to individual skill acquisition rates and baseline performance, as well as recommendations to pursue different career paths for those who will not meet the requirements despite practice (Ernst et al., 2015). However, as described earlier, international accepted and scientific proven benchmarks on how to select future pulmonologists do not exist yet (Ernst & Herth, 2013). Fortunately, during the last years, a growing interest exists towards factors predicting bronchoscopy performance and international objective benchmarks at which to aim (Konge et al., 2012; Wahidi et al., 2010).

The virtual reality simulator is a ground-breaking invention within the field of medical education. As a response to the need for training and assessment within the medical field, VR-training simulators provide a complete different approach by enabling bronchoscopic skills to be trained and assessed in a zero-risk and efficient environment (Herth, Shah, & Gompelmann, 2017). Bronchoscopy simulators provide a more patient centred and individual tailored alternative to the conventional training- and assessment methods due to several reasons. First off all, simulation-based training of technical bronchoscopy skills improves surgical skill proficiency (Ost et al., 2001). Secondly, simulators provide objective performance metrics, such as procedural time and scope-wall contact, which enables structured and equal performance level evaluation for every trainee. Moreover, performance measures also track the training progress of an individual over time and provide feedback and motivation for learning and evaluating the training programs (Epstein & Hundert, 2002). In contrast to e.g. cognitive aptitude tests, bronchoscopy simulators assess the integration of complex cognitive and psychomotor skills into a holistic view by incorporating a realistic

environment. Last but not least: simulator-based prediction of skill performance increases patient safety (Konge et al., 2012). Instead of procedural training and assessment on alive patients, residents can train in a zero-risk environment. Moreover, by identifying which individuals will become proficient surgeons and who will not, and by assessing residents based on proficiency, residents will be allowed to operate on patients only after possessing the necessary skills. Thus, bronchoscopy simulation education ensures patient safety, and learning effectiveness and efficiency. However, due to the scarcity of these highly expensive simulators, working hours restrictions and variable training durations due to individual differences simulator-based training and assessment is not widely implemented in the field of medical education yet (Sadideen, Hamaoui, Saadeddin, & Kneebone, 2012).

1.5.1 Current state of research on the prediction of MIS-performance.

The misconception of cognitive aptitude testing. Recently, there has been growing interest in the predictive value of innate aptitude for minimally invasive surgery performance, to enable controlled and systematic trainee selection. Innate abilities, including visual-spatial and perceptual abilities, of surgeons are related to the duration of learning to acquire competency in minimally invasive surgery (Gallagher, Cowie, Crothers, Jordan-Black, & Satava, 2003; Luursema et al., 2010; Schlickum, Hedman et al. 2011; Groenier, Schraagen et al. 2014). Based on the hypothesis that aptitude predicts surgical performance, cognitive aptitude tests are widely applied selection methods within the medical field. Since most studies covering the prediction of MIS performance focus on laparoscopy, we will discuss laparoscopic skill prediction first before we continue to bronchoscopy. A systematic review of studies exploring surgical performance predictors by Maan, Maan, Darzi and Aggarwal (2012) examined the predictive value of attributes which have impact on laparoscopic performance. According to this systematic review, psychomotor aptitude, as well as intermediate- and high-level visual-spatial perception are suitable criteria for assessing

prospective laparoscopy trainees for surgical training. More specifically, Stefanidis (2006) stated that psychomotor testing by means of innate ability tests has limited value in prediction of baseline laparoscopic performance, but these innate tests would be more suitable for predicting the skill acquisition speed.

However, research of Groenier, Schraagen, Miedema and Broeders (2014) concluded that cognitive aptitude tests did not predict laparoscopic performance as the relationship between MIS performance and cognitive aptitude is very complex. Groenier et al (2014) showed that independently examined individual cognitive abilities only mediated parts of MIS performance. Therefore, they suggested the importance of more general reasoning and cognitive abilities. Given the mixed results in the literature for cognitive aptitude, a study of Huijser, Schmettow and Groenier (2016) aimed at providing more insight in the relationship between cognitive aptitude and laparoscopic simulator task performance. However, no predictive relationship was found for visual spatial ability and maximum performance, neither for spatial memory and learning speed. In 2004, Veenman, Wilhelm & Beishuizen already stated that independently assessing single cognitive abilities does not predict the amount of training needed to reach proficiency or the learning rate on laparoscopic tasks. Thus, it can be concluded that the predictive value of cognitive aptitude tests is insufficient to allow for laparoscopy applicant solution.

As far as we know, no previous research has investigated whether single cognitive tests predict bronchoscopy performance. In contrast, most studies covering this topic concern laparoscopic procedures. We concluded that splitting up cognitive aptitude into separate abilities by means of cognitive aptitude tests may not resemble the complexity of abilities when integrated in a task like a bronchoscopic procedure. Since bronchoscopy and laparoscopy are both complex, cognitive demanding procedures, which require psychomotor skills, such as eye-hand coordination for scope manipulation and visual-spatial ability to

compensate for a lack of depth perception (Hofstad et al., 2013; Simoff, 2018), we tend to expect that cognitive aptitude tests may not resemble the complexity of bronchoscopy. To perform bronchoscopic procedures successfully, it requires the coordination of several skills, so it is impossible to attribute bronchoscopic performance to one single skill. Therefore, our surmise is that cognitive aptitude tests may not be valid prediction tools for future bronchoscopy performance either, whereas taking a more holistic approach could be more adequate. To explore a holistic approach to resident selection, we will apply the resemblance spectrum.

The resemblance spectrum. The resemblance spectrum provides a framework for exploring the resemblance of test suites to reliably predict future surgery performance. Previous studies used it to categorize psychometric tests in terms of resemblance to real laparoscopic procedures (Arendt et a., 2017; David et al., 2018). While innate ability tests resemble real-life laparoscopy surgery the least and real-life laparoscopic tasks the most, basic simulator tasks are in between both (see Figure 3). Thus, the spectrum varies from low-fidelity to high-fidelity techniques. It is expected that these different techniques rely on different combinations of cognitive skills and different cognitive demand levels, instead of relying on a single construct. Although existing research on the resemblance spectrum focused on laparoscopy, we expect the resemblance spectrum framework also to be applicable to bronchoscopy, since both procedures are highly complex and rely on different cognitive skills.

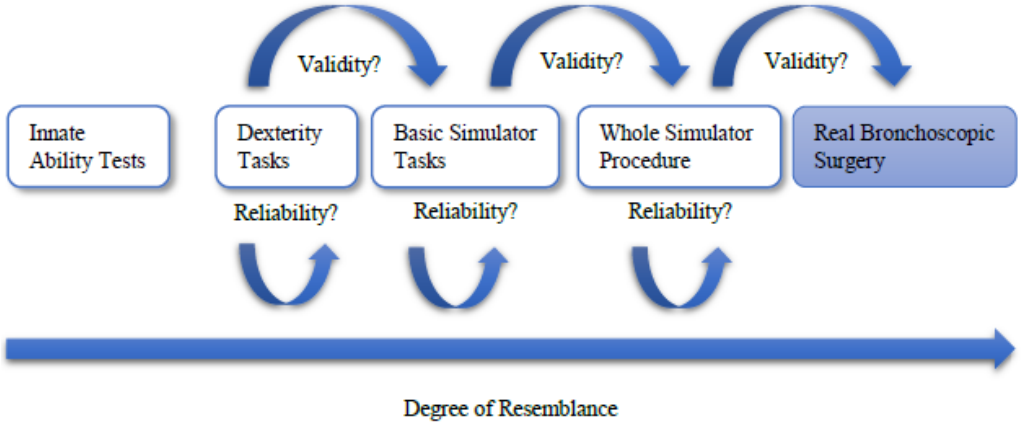


Figure 3. The resemblance spectrum.

The resemblance spectrum could be an explanation for the finding of Groenier et al. (2014) that cognitive aptitude tests do not predict laparoscopic performance, since cognitive aptitude tests are not as complex as real laparoscopic tasks. According to Arendt et al. (2017), it might be that cognitive aptitude tests do not cover the combination of manual dexterity and cognitive demand which is necessary for performing laparoscopic skills. However, it remains unclear how far psychometric tests should be on the right part of the spectrum to reliably predict laparoscopic performance. In order to evaluate the applicability of the resemblance spectrum, the concepts validity and internal consistency will be applied.

Validity refers to whether multiple tests measure the same underlying construct. Suites that are placed next to each other on the continuum have a more valid link to each other than those which are placed further apart. High correlations between tests indicate that tests can replace each other. Thus, the current study aims at exploring which suites on the left part of the spectrum show high correlation with real procedures. In addition, internal consistency relies on the combination of multiple items to obtain a reliable prediction of performance (Groenier, Schmettow & Huijser, 2017). A high correlating set of items results in low measurement error, which is the underlying construct of psychometric tests as intelligence tests. Only high-correlating items, in this case simulator tasks, will be added to the test suite. The current study aims at identifying which tasks contribute to a valid and reliable assessment

suite. The concepts validity and internal consistency are necessary to evaluate the resemblance of a test suite.

1.6 Low-fidelity Assessment

Several studies by students of the University of Twente focused towards (subparts of) the prediction of minimally invasive surgery skill acquisition (Arendt, Schmettow, & Groenier, 2017; Mührmann, Schmettow, & Groenier, 2018; Warnke, Schmettow, & Groenier, 2018; Küpper, Schmettow, & Groenier, 2018). A recent study of Arendt et al. (2017) focused on exploring what it takes to adequately predict laparoscopic performance. They estimated exponential learning curves on time-on-task to explore whether basic laparoscopic tasks and dexterity tasks enable valid and reliable prediction of MIS-performance to allow for systematic selection of prospective surgeons. In accordance with previously described studies, Arendt et al. (2017) found individual differences in maximum performance. However, they concluded that low-fidelity assessment, such as dexterity tasks, might be less feasible, whereas laparoscopy simulator tasks are more promising for surgeon selection due to higher resemblance to the real surgical tasks. Thus, in terms of the resemblance spectrum, dexterity tasks are not resembling enough for applicant selection. Moreover, studies of Mührmann et al. (2018) and Küpper et al. (2018) originally planned to test whether a low-fidelity boxtrainer can improve bronchoscopic simulator task performance. However, due to unforeseen circumstances Mührmann et al. (2018) shifted their focus on exploring the association between time on task and wall contact. They found that as participants made more mistakes, they needed more time for completing a simulator task. Moreover, the low-fi boxtrainer is not suitable as a substitute of a high-fidelity VR-simulator. Thus, as well as cognitive aptitude tests, low-fidelity tasks such as boxtrainers and dexterity tasks may not be adequate prediction tools for MIS-performance either. This means, in terms of the resemblance spectrum, both innate ability tests and dexterity tasks have proven to be not resembling enough to real

laparoscopic surgery to predict laparoscopic performance. On the contrary, Kramp (2016) suggested the implementation of simulator-based assessment. While innate ability tests only predict part of individual differences in MIS performance, simulator-based assessment provides a more holistic approach by measuring several innate abilities in a more realistic environment at once. Consequently, the aim of the current study is to take one step right on the resemblance spectrum and to explore the test suite of basic simulator tasks.

1.7 High-fidelity Simulator Assessment

Although research has illuminated the usefulness of simulator-based training in the medical field, only few studies have examined the use of simulators as prediction and assessment tools of bronchoscopic skill acquisition. The studies of Ost (2001) and Pastis (2014) demonstrated construct validity of two different bronchoscopy simulators in differentiating between levels of experience. Ost (2001) was able to distinguish between novices (0 bronchoscopies, n = 11), intermediates (25 to 500 bronchoscopies, n = 8) and expert bronchoscopists (> 500 bronchoscopies, n = 9) in terms of wall collisions, percentage of segments visualized and procedure time measured by a bronchoscopy simulator. In addition, Pastis et al. (2014) demonstrated construct validity of the Simbionix GI-BRONCH Mentor™ in discriminating skill level in scope manipulation and airway anatomy among novice (<10 bronchoscopies, n = 7), experienced (200 to 1000 bronchoscopies, n = 6) and expert pulmonologists (> 1000 bronchoscopies, n = 7).

A closer look to the literature on simulators in bronchoscopy education, however, reveals a number of gaps and shortcomings. In the study of Pastis et al. (2014) it remains unknown how individuals learn bronchoscopic skills on a simulator over time, as participants only performed each task once. To estimate an individual's performance with high certainty, repeated performance measures are necessary (Schmettow, 2018b). In contrast, the study of Ost (2001) did include twenty repetitions of each task, but learning curves were estimated

over the population mean. On the other hand, studies which do report a wide variation in bronchoscopic skill acquisition used either manual performance assessment tools (Wahidi et al., 2010) or an unknown performance algorithm including time and number of collisions (Dalal et al., 2011), instead of raw simulator metrics such as time or wall-contact. Since predicting bronchoscopic performance, or minimally invasive surgery performance in general, by means of individual simulator performance is still in its infancy, the current study will cover this approach from the ground up.

1.8 Research Questions

The aim of our work is to explore whether basic bronchoscopy simulator tasks can be used to allow for a valid prediction of bronchoscopic skills. Bronchoscopic skills are very complex and hard to learn: bronchoscopists show increased complication rates when performing bronchoscopy during the first trimester of training (Quellette, 2006). Moreover, individuals differ at learning bronchoscopic skills which might be due to differences in innate psychomotor ability (Dalal et al., 2011; Ernst et al., 2015). In order to ensure patient safety, adequate selection, training and assessment are crucial. To explore individual differences in bronchoscopic skill acquisition, the current study uses three basic simulator tasks which are expected to rely on psychomotor ability.

As current research in simulator-based learning of bronchoscopic skills remains limited due to methodological shortcomings, this study addresses learning bronchoscopic skills on a simulator by starting from scratch. We will estimate individual exponential learning curves based on time-on-task, our main parameter of interest is maximum performance (asymptote). Firstly, we will check whether individuals show learning in all three simulator tasks by decreased time-on-task using estimated learning curves. Next, we will investigate whether individuals differ in acquiring basic bronchoscopic tasks by exploring individual differences in the asymptote parameter. Finally, in order to answer the question whether the

basic simulator tasks are predictors of individual performance of technical bronchoscopic skills, we will test whether the asymptote of the three tasks correlate with each other in the performance parameter time-on-task. The current study is exploratory in nature, no hypotheses were formulated.

2. Method

2.1 Participants

The current study consists of a convenience sample of nineteen students of the University of Twente (68.4% female, M_{age} = 20.2 years, age range: 19-24 years) who participated voluntarily in turn for three course credits. Five participants were Dutch, thirteen were German and one was Bulgarian. Two participants reported to be left-handed. One participant wore eyeglasses. All participants were novices without prior endoscopy experience. Fifteen participants were recruited by use of the test subject pool SONA of the faculty of Behavioural, Management and Social sciences of the University of Twente. Four participants were recruited from the direct environment of the researchers. Participants were included if they had the minimum age of 18 years, speak and understand English and have normal or corrected to normal vision. Prior to participation, all participants gave written permission by signing an informed consent form (See Appendix C). Ethical approval was obtained from the Ethics Committee at the Faculty of Behavioural Sciences of the University of Twente, The Netherlands (request number: BCE18198).

2.2 Design

The current study has a repeated-measures within-subjects design. All participants carried out at least two tasks of a training curriculum on the virtual-reality bronchoscopy simulator GI-BRONCH Mentor™. Due to time constraints and technical problems with the simulator, only five participants completed all three tasks. All tasks were repeated fifteen times. First, participants completed the Basic Scope Manipulation task, after which they

completed the Guided Anatomical Navigation task. The routes in the Basic Scope Manipulation task were randomly assigned. Participants had to identify ten lightbulbs for the Guided Anatomical Navigation task. During an optional second session, participants performed the Step-by-Step Diagnostic Manoeuvres task, first taking a forceps sample and then taking a brush sample, but only after having completed the first and second task. For all three tasks, a workload questionnaire was conducted after each uneven trial, starting with the first. [1]

2.3 Materials

2.3.1 GI-BRONCH Mentor™ . The three simulator tasks were administered with the GI-BRONCH Mentor™ virtual reality simulator. The GI-BRONCH Mentor™, developed by 3D Systems, Cleveland, OH, USA (formerly Simbionix), is a flexible bronchoscopy training simulator used for the simulation of a wide range of basic bronchoscopy tasks and clinical bronchoscopy procedures(Simbionix, 2018b). The GI-BRONCH Mentor™ consists of a plastic mannequin with a mouth and nose entrance for bronchoscopy or upper endoscopy (see Figure 4). The simulator provides a realistic training environment including an authentic bronchoscope with tactile feedback (Pentax ECS-3804F) and bronchoscopy tools, such as a biopsy forceps, cytology brush and aspirating needle (See Figure 12 and 13 in Appendix A). The computer generates a dynamic endoscopic view, based on the bronchoscope's movement captured by a sensor on the tip of the scope. The endoscopic view is displayed on a 24-inch LCD touch screen. The software running on the simulator is MentorLearn LMS (version 1.4.0.68).

[1] The workload questionnaire is beyond the scope of the current study. The current study is part of a broader study about individual differences in learning bronchoscopic skills among novices, measured by a virtual bronchoscopy simulator and self-reported workload. Three other students from the University of Twente worked on the study with each different research questions. With respect to the completeness of the method section, parts which are irrelevant for the current study are mentioned as well.

Figure 4. BronchMentor™ (Simbionix, 2018a)

The GI-BRONCH Mentor™ incorporates a bronchoscopy training curriculum with scenarios with varying degrees of complexity of tasks and difficulty of anatomy, divided over four modules: Essential Bronchoscopy, Emergency Bronchoscopy, Essential EBUS and CHEST Standardized Curriculum. The current study only uses three basic skill tasks, Basic Scope Manipulation (Cyberscopy) and Guided Anatomic Navigation, from the Essential Bronchoscopy module. This module provides tasks to acquire and integrate bronchoscopic capabilities to accelerate the trainee's learning curve. Objectives of the module are for instance to acquire basic bronchoscope manoeuvring capabilities, to improve hand-eye coordination, to enhance 3D anatomical perception and to objectively assess bronchoscopic skills level.

The current study only focused on the Basic Scope Manipulation, Guided Anatomical Navigation and Step-By-Step Diagnostic Manoeuvres, because we expect these three tasks to be related to innate psychomotor ability. Since the current study focuses on the psychomotor ability component of bronchoscopic skills, it is not necessary to include participants with e.g. anatomical knowledge in the target group. According to Jabbour et al. (2011) complex

bronchoscopy motor skills should be trained by breaking down one task into component parts, both for training and assessment. Task deconstruction into easier subtasks improves learning because it decreases frustrations and mental demands, and provides a feeling of success when mastering easier to acquire subtasks (Gallagher, Smith, et al., 2003; Jabbour et al., 2011). In a previous study, the Basic Scope Manipulation task was able to discriminate between various degrees of skills among novice, experienced and expert bronchoscopists based on final score, total time, % at mid-lumen and total wall hits (Pastis et al., 2014).

Basic Scope Manipulation (Cyberscopy). The first skill task was Basic Scope Manipulation. This task aimed at acquiring basic scope control skills such as developing hand-eye coordination. The user should follow a blue ball in a cyber environment while trying to stay in the centre of the lumen and avoiding wall contact (see Figure 5). The computer randomly selected a path. To account for differences in difficulty of the randomly chosen paths, paths were taken into account at the data analysis.

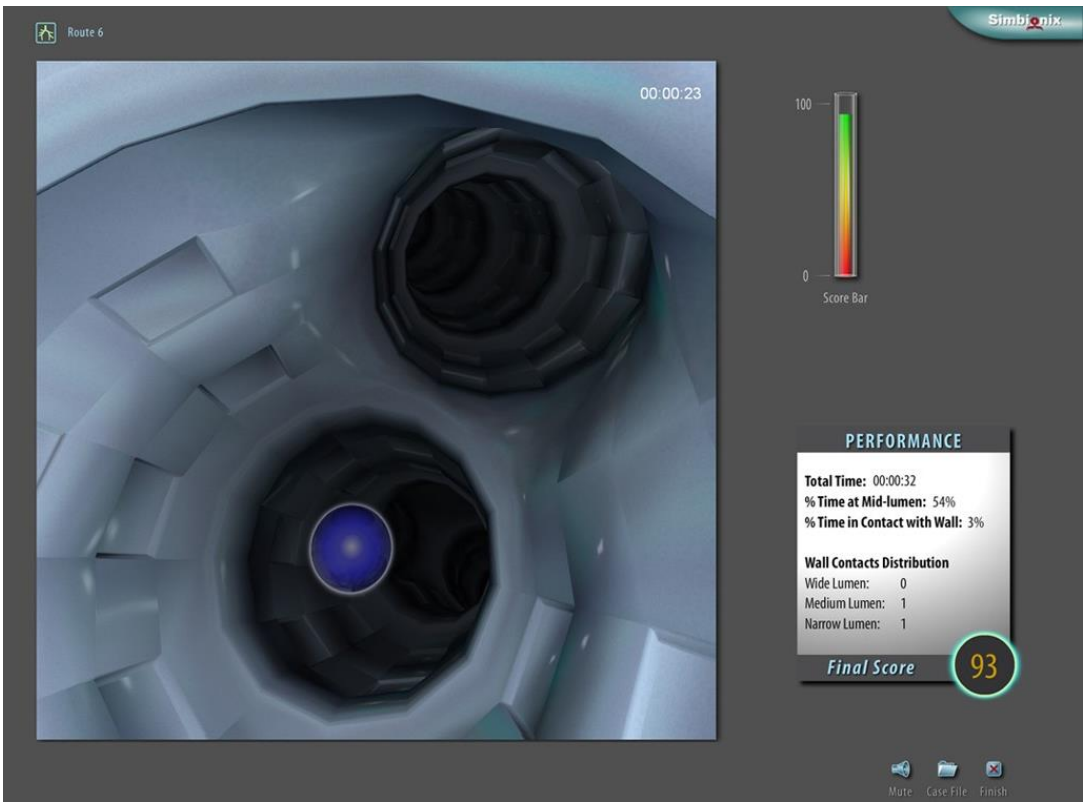


Figure 5. Simulator task 1: Basic Scope Manipulation (Simbionix, 2018a).

Guided Anatomic Navigation. The second skill task was Guided Anatomic Navigation. The goal of this task was to learn and practice correct scope navigation within the anatomical environment in order to perform a complete airway inspection by passing all bifurcations of the trachea. By matching the bronchoscope's view finder to a presented light bulb, directional guidance was provided to help the user find the right scope roll and tip flexion or extension to pass the carina (see Figure 6). Correct alignment of the scope's view finder to the presented light bulb resulted in a change in colour and auditory feedback. Wall contact was indicated by visual and auditory feedback as well.

It was expected that novice participants without prior anatomical knowledge could experience difficulties with reaching hard to reach lumen, such as right and left upper lobe, because they will overlook these carinas. Furthermore, during the task no indication is given about the amount of to-be-completed bifurcations; it is expected that novices do not know how many they have to complete. Thus, we decided that participants had to identify ten of twenty-eight available light bulbs per trial.

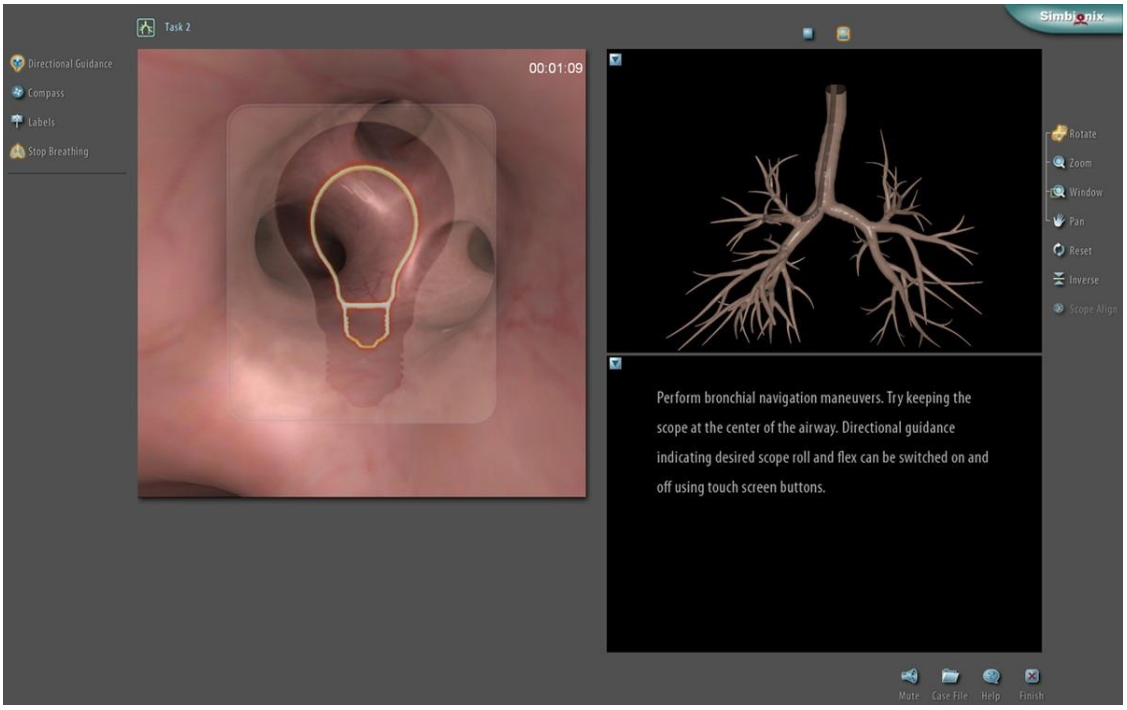


Figure 6. Simulator task 2. Guided Anatomic Navigation (Simbionix, 2018a).

Step-by-Step Diagnostic Manoeuvres. The third skill task was Step-by-Step Diagnostic Manoeuvres. The goal of this task was to acquire skills needed to perform forceps, brush and transbronchial needle biopsy from a purple-marked area of interest. However, the current study only used the tools for brush and forceps biopsy, due to the complexity of the transbronchial needle biopsy. Step-by-step instructions were provided on the screen for each diagnostic manoeuvre (See Figure 7). After completion of each diagnostic manoeuvre, feedback was provided on the screen about whether tissue was obtained successfully and whether or not the tool was used correctly during sampling. Participants took a forceps biopsy and then a brush biopsy.

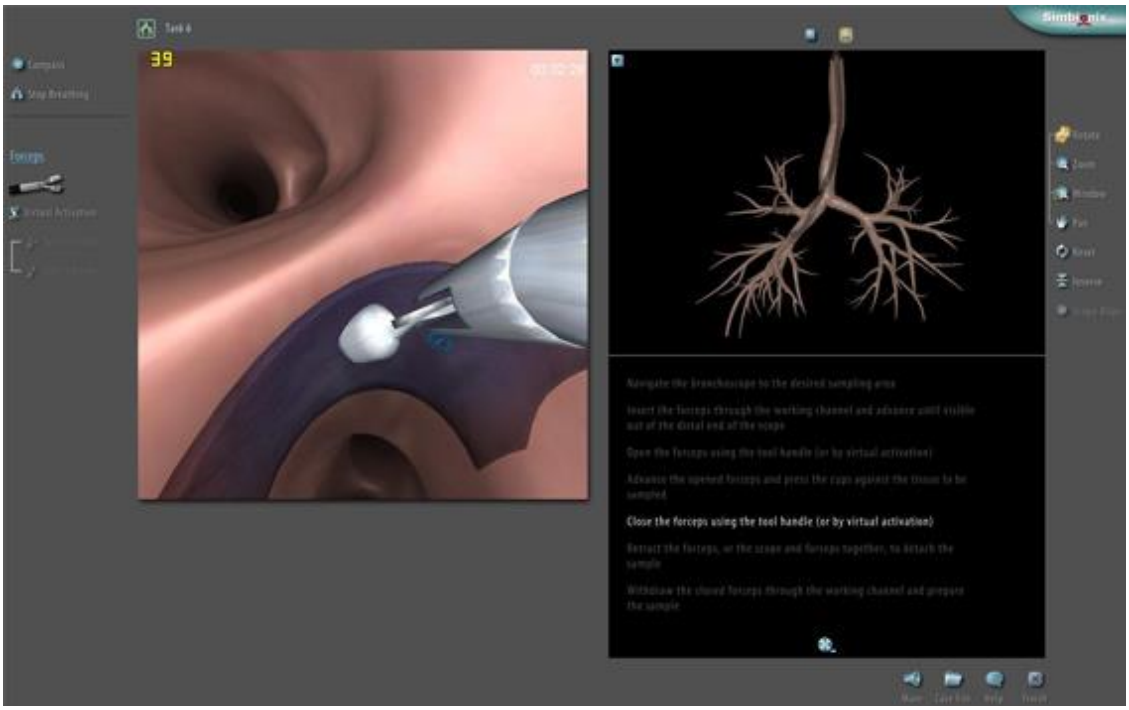


Figure 7. Simulator task 3. Step-by-Step Diagnostic Manoeuvres (Simbionix, 2018a)

2.4 Procedure

The study took place in the advanced Simulation Room 2 at the Experimental Centre for Technical Medicine (ECTM) at the University of Twente. Participants were tested individually.

2.4.1 Briefing. The experimenter instructed participants about the procedure of the research, after which informed consent was obtained (See Appendix C for the informed

consent form). Next, a demographic questionnaire was conducted with questions concerning gender, age, nationality, vision, gaming experience and bronchoscopic experience (See Appendix for the demographic questionnaire). In order to prevent fatigue, the experiment was divided over two sessions of 120 and 60 minutes, respectively. The first and second task were performed during the first session, whereas the third task was performed during the second session which was at least 24 hours after the first session, with a maximum of three weeks. Participants repeated all three tasks fifteen times. The simulator digitally stored the data in the MentorLearn cloud. After each uneven trial number of each task, starting with the first trial, participants filled out the six subscales of the NASA-TLX questionnaire on a laptop. After the fifteenth trial, participants ranked the perceived impact of these six subscales from high to low as part two of the NASA-TLX scale. All instructions which were given to the participants can be found in Appendix D.

Simulator task 1: Basic Scope Manipulation. In the first part of the first session, participants performed the Basic Scope Manipulation task. First, participants were instructed about the simulator and scope handling by means of a short video instruction, after which they started the first task on the simulator. Instruction about the basic scope manipulation task was presented on the simulator's screen. During this task, routes differed in difficulty.

Simulator task 2: Guided Anatomic Navigation. After a five-minute break, participants started with the second part of the first session by performing the Guided Anatomic Navigation task. Like the first simulator task, instruction about the guided anatomic navigation was presented on the simulator's screen. Participants had to identify ten of twenty-eight available light bulbs.

Simulator task 3: Step-by-Step Diagnostic Manoeuvres. During an optional second session, participants performed the third task. First, participants were instructed about the functioning and use of the diagnostic tools by means of a short video instruction, after which

they started the first task on the simulator. Participants were instructed to always take the left carina when arriving at a bifurcation. A step-by-step instruction about the procedure was presented on the simulator's screen. Participants had to take a forceps tissue sample first, after which they performed a brush tissue sampling. During both procedures, tissue samples should be taken from a purple-marked area of interest in the simulated lung. First, the researcher handed the sample tool over to the participant when arriving at the location of interest. After inserting the tool into the working channel, the participant had to choose the correct sampling tool on the screen. Then, the researcher handed the tool to the non-dominant hand. The tissue sample was obtained by holding the scope at the sampling area and either pushing the lever up or down. When the feedback on the screen showed a successful tissue sample, the tool needed to be rejected until the tool-selection screen was shown again. Next, the sampling brush had to be chosen. A brush sample was obtained by holding the scope at the sampling area and moving the tool back and forth in the working channel. After completing the second tissue sampling correctly, the trial was finished.

2.4.2 Debriefing. After completion of the second or third task, depending on whether the participant would participate in the second session or not, the participants were debriefed. A short explanation of the study and its aim was given, after which the participants had the opportunity to ask questions and finally were thanked for participation.

2.5 Measures

2.5.1 BronchMentor™. The BronchMentor™ provides a wide range of performance parameters. Performance parameters computed by the simulator include, but are not limited to, total performance time, final score, % time at mid-lumen, % time in contact with the wall, % time with clear visibility, list of bifurcations where manoeuvre was performed and a lobe or segment was entered satisfactorily (first, second or third and up attempt) and list of

skipped bifurcations. However, as the simulator is designed for educational rather than for research purposes, not all data were useful for our study.

The main parameter of the current study was time on task. Van Dongen et al. 2007 stated that the amount of time needed to complete a task is a suitable indicator of the participant's skill. However, using the duration to complete a task as a performance measure has both benefits and drawbacks. On one hand, time as an performance parameter does not indicate whether a procedure was performed successfully and safe. However, as time is an important patient safety factor due to possible complications and the performance under moderate sedation (Jabbour et al., 2011), time-on-task (in minutes) was used as the primary performance parameter. According to a study of Pastis and colleagues (2014) on the construct validity of the Simbionix BronchMentor™ the parameters final score, total time, % at mid-lumen and total wall hits of task 1 were able to discriminate correctly between experts, experienced and novice bronchoscopists. In order to include accuracy of performance, we also included all hits of the Basic Scope Manipulation task as a parameter. Unfortunately, the parameter Wall Hits was not available for the tasks Guided Anatomical Navigation and Step-by-Step Diagnostic Manoeuvres. Due to the unknown formula of the final score, this parameter was not used in current study. For each task and trial, time-on-task was measured in seconds. This performance parameter was used to establish individual learning curves for all tasks and participants.

For the first simulator task, the total procedural time provided by the simulator was used as time-on-task parameter. However, for the second and third task, times were administered by means of a stopwatch due to a brief delay in the time provided by the simulator in contrast to the real time-on-task. Besides the total time-on-task, for the second task the time until passing the vocal cords and for every identified lightbulb the duration in seconds was measured as well. Moreover, for the third task, we measured the duration in

seconds until passing the vocal cords, handling the instrument to the participant, completing the forceps biopsy and completing the brush biopsy.

2.5.2 Workload Questionnaire. For assessing self-reported workload of the three tasks the multi-dimensional NASA Task Load Index (NASA-TLX) scale was used. As described earlier, the workload questionnaire is described for the sake of completion, as it was not covered by the current research questions. The NASA-TLX provides an overall workload score based on a weighted average of ratings on six subscales: mental demand (How mentally demanding was the task?), physical demand (How physically demanding was the task?), temporal demand (How stressed were you by the speed of the task?), own effort (How hard did you have to work to achieve your result?) and frustration (How frustrated were you about the task or parts of it?) (National Aeronautics and Space Administration (NASA), n.d.). After completing fifteen trials, the participant had to rank the weight of the factors, beginning with the factors with the highest impact. The overall workload score per trial per task for each participant was computed by multiplying each rating by the weight given to that factor by the participant (NASA, n.d.). Then, the sum of the weighted ratings for each task was divided by fifteen, which is the sum of the weights.

2.6 Statistical Analysis

A non-linear mixed-effects model with an exponential learning curve as a likelihood function was performed to create regression models for estimating learning curves for each participant (Heathcote et al., 2000). No major data cleanings were performed.

2.6.1 Learning curve model. The analysis aims at estimating learning curves to explore skill acquisition among individuals. According to Heathcote, Brown and Mewhort (2000), the exponential learning curve model is an appropriate method to explore individual learning progress. This model uses the exponential law of practice to compute expected performance after practice trials. The function described by Heatcote et al. is (2000):

$$Y_{ptN} = Asym_{pt} + Ampl_{pt}exp(-Rate_{pt}N)$$

The exponential learning curve model includes the parameters maximum performance (asymptote), amplitude and rate, as described earlier in the introduction. A learning curve has been estimated for each participant and task based on time-on-task, including the parameters summarizing previous trials or the initial performance, the learning rate and the estimated maximum performance. The nonlinear multilevel mixed effect model was performed using the regression package 'brms' 2.3.0 (Bürkner, 2018) in the statistical computing environment R version 3.5.0 (R Core Team, 2018). The multilevel model was used to estimate individual learning curves per task and combine these separate learning curves, which allowed for an variation analysis on population-level. Non-linear functions were built using the package Asymptote (Schmettow, 2018a).

2.6.2 Non-linear regression. We set up the LARY (Linked Amplitude Rate Asymptote) model which was used to compute the expected performance (ToT) after practice trials based on the exponential law of practice (Schmettow, 2018a). Random effects indicate the variance in a population which is caused by individual performance differences. In order to explore random effects, the parametrizations amplitude and asymptote were linearized through link functions on a log-scale and logit-scale, all with a range from $-\infty$ to $+\infty$. Generalized linear mixed-effects models were estimated using the Bayr package (version 0.8.8). The Generalized Linear Models (GLM) framework consists of three elements: A random probability distribution, a linear predictor and a link function. Random distributions take into account the relationship of variance and mean, and the expected pattern of randomness, while linear predictors contain information about the model's independent variables (Schmettow, 2018b). The link function establishes the relationship between the mean of the distribution function, the expected value μ, and the linear predictor θ. As a consequence, the range of the non-linearly transformed mean ranges from $-\infty$ to ∞ and thus

linearity is established between μ and θ. In doing so, link functions meet two essential requirements: Firstly, the order in magnitude in link functions must be maintained, a so-called monotonically increasing function. Moreover, to ensure linearity, link functions are required to map variables to the range [-∞; ∞].

The LARY model estimated random effects for the amplitude, rate and asymptote parametrizations for each task, based on the parameter time-on-task with credibility intervals of 95%. Weakly informative priors were used for this data model. To allow for the variance in the difficulty of randomly chosen routes during task one, the model included route as a variable for computing the parameters asymptote and amplitude.

We used the Gamma distribution of random probability for modelling the error component within the LARY model instead of the Gaussian distribution, because time-on-task measures typically show left-skewed random patterns and by approaching maximal performance, variance of residuals decreases. Errors in time-on-task are typically negatively-skewed distributed, which means that the mean is to left of the peak. Moreover, time-on-task measures are continuous and positive, but lower boundaries are rather at the lowest human possible time to complete a task than at zero. Additionally, exponential distribution rises only if all events have the same probability to occur. However, this will never happen in behavioural research. As a solution, we use the gamma distribution since, in contrast to exponential distributions, the peak of the gamma distribution can move along the x-axis rather than fixed at zero.

Then, the parameters amplitude, rate and asymptote were computed for each task by means of the posterior probability distribution of the LARY model, both on individual- and population-level. Correlations between the asymptote parameter for each task on population-level were calculated, again with credibility intervals of 95%. Finally, learning curves on expected maximum performance (ToT) were estimated for each participant per task.

3. Results

First, exploratory plots of raw time-on-task data will be given to provide an overview of the collected data. Next, we established the LARY-model to estimate the performance parameters on population-level, followed by an analysis on participant-level. Then, we calculated correlations between the parameters of the three task. Finally, we criticised our used model to determine the robustness of the performed analyses.

3.1 Compliance and Overview of Collected Data

Nineteen participants performed task 1 and 2, whereas only five participants completed the third task due to technical difficulties with the simulator and time constraints. Moreover, since six participants took part in the study of Küpper, Schmettow and Groenier (2018) as well, the number of trials performed by these participants was constrained by the experiment's duration. The average amount of performed trials per task are 13.6 for Basic Scope Manipulation (*SD* = 2.63, range = 7-16), 12.2 for the Guided Anatomical Navigation (*SD* = 4.7, range = 4-15) and 14.0 for Step-By-Step Diagnostic Manoeuvres (*SD* = 1, range = 5-15). For a full syntax and an overview of the dataset see Appendix E.

3.1.1 Simulator tasks. Next, it is explored whether a non-linear model fitted to the time-on-task performance parameter of all three basic bronchoscopy simulator tasks. Prior to model fitting, we introduce some exploratory plots on raw time-on-task data. Figure 8 shows curves smoothed using locally weighted scatterplot smoothing (LOESS) fitted to raw time-on-task of all participants for the three simulator tasks (See Figure 15 in Appendix B for raw time-on-task per participant). The x-axis shows the number of trials. On the y-axis, the time-on-task is displayed in seconds. Colours indicate each of the three simulator tasks.

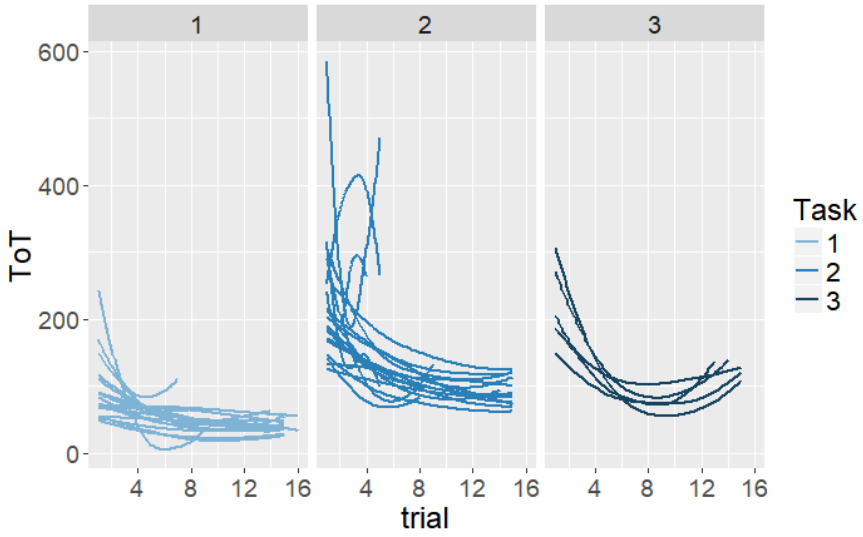


Figure 8. Raw time-on-task of participants for the three simulator tasks. Time-on-task (in seconds) on the y-axis, amount of trials on the x-axis.

The performance on the first task, Basic Scope Manipulation, looks more homogenous than on the second and third task, Guided Anatomical Navigation and Step-By-Step Diagnostic Manoeuvres, respectively. Since time-on-task is a reaction-time performance parameter, a lower score on the y-axis indicates a better performance. As all participants show decreasing time-on-task values over time for all three tasks, this indicates that performance increases and thus participants become better at performing the three basic tasks. However, the decline in ToT differs in magnitude among both participants and tasks. During the first task, only one participant shows a steep initial decrease in ToT ending in approaching the asymptote, whereas all other participants show curves approaching the asymptote initially. For the second task, we see an initially steep fall, after which curves become less steep. Moreover, almost all participants show a general parallel decline of time-on-task on the second task.

Although task 3 was only performed by five participants, results are relatively consistent. Curves fall steep during the first trials and increase again between the eight and tenth trial. Curves did not completely stabilize, so the asymptote was not reached during the third task. As all three tasks show curves which fall (steep) initially followed by approaching

the asymptote or increasing again, we can conclude that the ToT performance depicts non-linear curves.

Notable is the finding that of randomly chosen routes of task 1, route 1 shows the highest spread in ToT. Figure 9 shows the ToT distribution per route of task 1. The x-axis depicts the route, the y-axis the time-on-task in seconds. The median weights of most routes are estimated to be centred between 40 and 50 seconds, whereas routes 1 and 5 show median weights estimated around 55 and 65 seconds respectively. However, time-on-task-scores are highly variable for route 1 of the task Basic Scope Manipulation, since values lie between 20 and 205 seconds. Moreover, it is estimated that 25% of the measurements of route showed a task completion time between 55 and 120 seconds. Thus, we conclude that it is recommended to add route as a control variable to our model in order to account for variation in task duration.

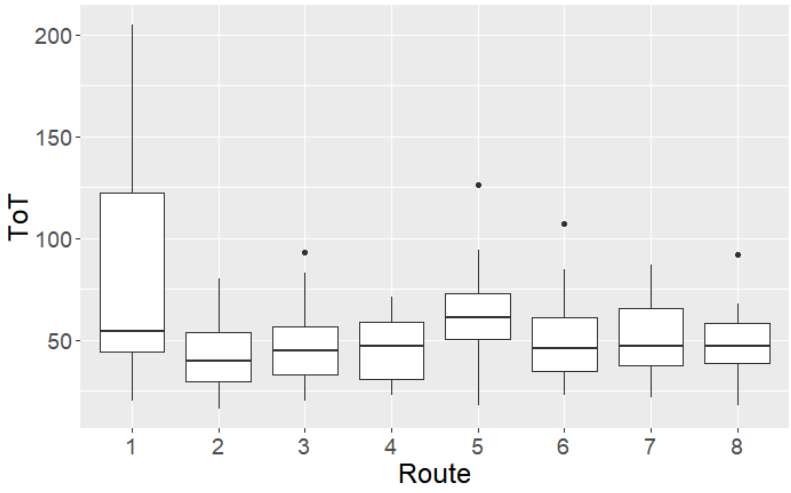


Figure 9. Time-on-Task distribution (in seconds) per route of task 1.

3.2 Bronchoscopic Skills Acquisition

To explore whether individuals show learning in all three simulator tasks by decreased time-on-task, and whether it is possible to estimate individual learning curves, we estimated non-linear multilevel mixed-effects models for the parameters time-on-task and wall contact. The LARY model estimated random effects for amplitude, rate and asymptote

parametrizations for each task, including the Gamma distribution for modelling the error component. Since we found that routes of the Basic Scope Navigation task showed variation in time-on-task, we included route as a random level effect variable for computing the parameters asymptote and amplitude. First, a visual exploration from individual learning curves was performed, after which a population-level analyse was conducted to explore the learning of individuals in all three simulator tasks.

3.2.1 Time-on-task.

In order to explore whether individuals show learning in all three simulator tasks by decreased time-on-task, learning curves were estimated for each participant. Figure 10 shows the individual learning curves of all participants for all three tasks, whereas Figure 16 in Appendix D shows learning curves per individual.

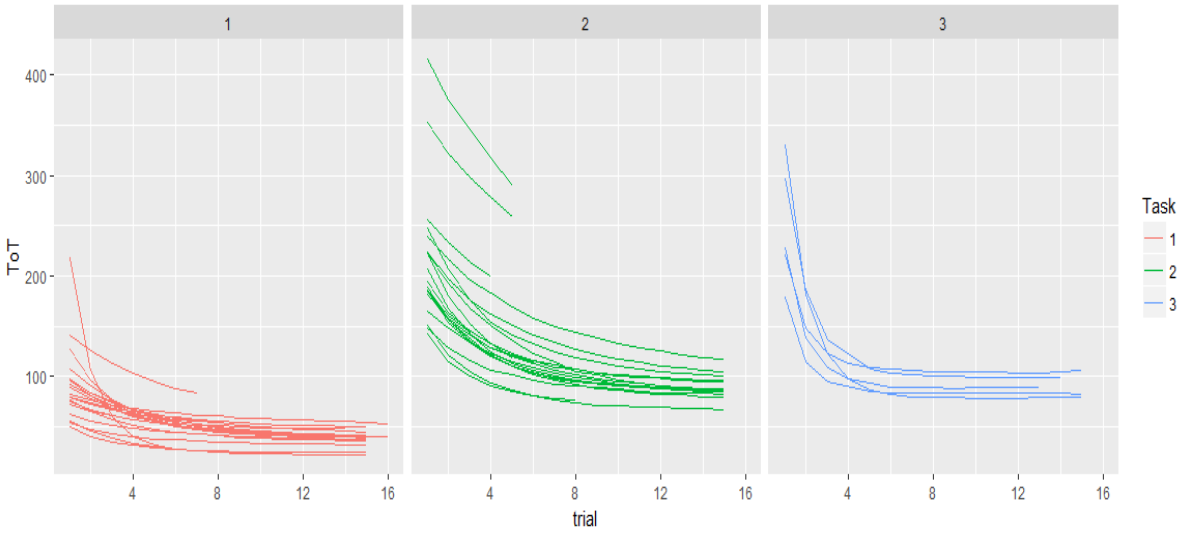


Figure 10. Estimated time-on-task learning curves of all participants for the three simulator tasks. Low values on the y-axis indicate superior performance.

As can be seen in Figure 10, learning curves were estimated for all three tasks. However, differences are visible between the three tasks' learning curves. For the first task, several participants show low improvement in performance and reach the relatively low asymptote between the eight and tenth trial during the first task. This means that some individuals are able to learn this task quite fast, which could indicate that the first task is

relatively easy to complete. In contrast to the first task, learning curves on the second task show lower initial performance, followed by a more gradual decrease resulting in reaching the asymptote between the tenth and twelfth trial. Notable is the coherence between asymptote and amplitude as curves run in parallel: low asymptote is associated with low amplitude and vice versa. Whereas only five participants performed Step-by-Step Diagnostic Manoeuvres, relatively consistent individual learning curves are visible. Notably is the low initial performance, followed by a sudden steep fall which results in reaching the asymptote between the sixth and eight trial.

Next, to establish population-level effects for the three learning parameters over three tasks, the posterior distribution of our LARY-model was used. The population's average, indicated by fixed effects, per task and parameter can be found in Table 1 on the next page. The population-level fixed effect of maximum performance for the task Basic Scope Manipulation (3.60, *95% CI* [3.34; 3.82]) was lower than for maximum performance of Guided Anatomical Navigation and Step-by-Step Diagnostic Manoeuvres (4.47, *95% CI* [4.24; 4.68] and 4.53, *95% CI* [4.26; 4.79], respectively). All three simulator tasks show low levels of certainty, as credibility limits deviate more than 0.2. In addition, the amplitude for task 1 (4.19; *95% CI* [3.69; 4.61]) is again lower than for task 2 (5.11; *95% CI* [4.75; 5.45]) and 3 (6.06; *95% CI* [4.86; 7.11]). Notably, the predicted values of amplitude and asymptote run more or less in parallel. Thus, based on the population's average, initial performance might provide some information about the expected maximum performance.

Table 1

Fixed effects and corresponding random factor variation of the three parameters per task

Parameter	Task	Fixed Effects			Random Factor Variation		
		Center	Lower CI95	Upper CI95	Center	Lower CI95	Upper CI95
Amplitude	Basic Scope Manipulation	4.19	3.69	4.61	0.77	0.44	1.33
Amplitude	Guided Anatomical Navigation	5.11	4.75	5.45	0.44	0.14	0.78
Amplitude	Step-by-Step Diagnostic Manoeuvres	6.06	4.86	7.11	0.60	0.04	2.91
Rate	Basic Scope Manipulation	-1.24	-1.96	-0.68	0.63	0.27	1.25
Rate	Guided Anatomical Navigation	-1.34	-1.98	-0.78	0.62	0.12	1.23
Rate	Step-by-Step Diagnostic Manoeuvres	-0.04	-0.72	0.66	0.33	0.01	1.78
Asymptote	Basic Scope Manipulation	3.60	3.34	3.82	0.31	0.19	0.51
Asymptote	Guided Anatomical Navigation	4.47	4.24	4.68	0.19	0.04	0.51
Asymptote	Step-by-Step Diagnostic Manoeuvres	4.53	4.26	4.79	0.21	0.04	0.81

Note. Estimates with 95% credibility

3.2.2 Wall contact. Next to estimating population-level effects for time-on-task, we also estimated fixed effects for the amount of wall contacts during task Basic Scope Manipulation. Population-level effects for amplitude (-0.18, 95% [-10.24; 7.28]), rate (0.80, 95% [-8.38; 10.83]) and maximum performance (0.89, 95% [-9.37; 1.39]) are highly uncertain, since credibility limits deviate more than 0.2.

3.3 Individual Differences in Bronchoscopic Skill Acquisition

In order to explore whether individuals show differences in learning bronchoscopic skills, we conducted the following two analysis. Since performance prediction of trainees is only required if participants differ at performance, this question will be answered next. We estimated the maximum performance parameters on participant-level using the posterior distribution of the LARY-model. As described in the data analysis section, maximum performance runs on the log scale. First, a visual exploration from individual learning curves was performed, after which a group-level analyse was conducted to explore the diversity of performance in the sample.

3.3.1 Individual learning curves. To explore whether individuals differ in acquiring basic bronchoscopic tasks, learning curves were estimated for each participant. Figure 10 on page 41 shows the individual learning curves of all participants for all three tasks, whereas figure 16 in Appendix D shows learning curves per individual. For all three tasks, participants show variation in performance. The first task depicts the most homogenous curves, which indicates a lower inter-individual variability compared to the other two tasks. In contrast, the greatest individual differences are visible within the second task, Guided Anatomical Navigation. We see that most learning curves run in parallel but differences exist in initial- and maximum performance. Since learning curves for the second task run in parallel, including the first and last trials, it may be possible that initial performance could be a predictor of maximum performance.

3.3.2 Random factor variation.

Time-on-task. Next to visually exploring individual learning curves, we explored whether individual show differences in learning bronchoscopic skills by means of standard deviations from the population-level average. Table 1 shows the population average, represented by the fixed effects, per task and parameter on log-scale. The random factor

variation, which indicates the standard deviations of the population average, can be extracted from the posterior of the mixed-effects model. Thus, the we'll explore the overall variation of performance in the population by analysing the random factor variation. The random factor variation of maximum performance was highest for the task Basic Scope Manipulation (0.31, *95% CI* [0.19; 0.51]), followed by Step-by-Step Diagnostic Manoeuvres (0.21, *95% CI* [0.04; 0.51]) and Guided Anatomical Navigation (0.19, *95% CI* [0.04; 0.81]). This means that the greatest variation from the population-average of maximum performance is within the Basic Scope Manipulation task. The first task shows a moderate level of certainty, whereas uncertainty is high for the second and third task. Fixed effects random effects variation for the three parameters are presented in Table 1.

3.4.2 Wall contacts. Examination of random effects of wall contacts for the first task suggested that variation in maximum performance exists across individuals, which indicates that the population mean of maximum performance of wall contacts is not applicable to everyone. Figure 11 displays the random effect size per parameter for each participant of the first task. Variation is highest for maximum performance (asymptote) as the deviations range from -0.5 to 0.75. Amplitude and rate are less spread since values are centred around zero. However, all values are highly uncertain, as can be seen in the appendix which includes a list of participant-level parameters for all parameters.

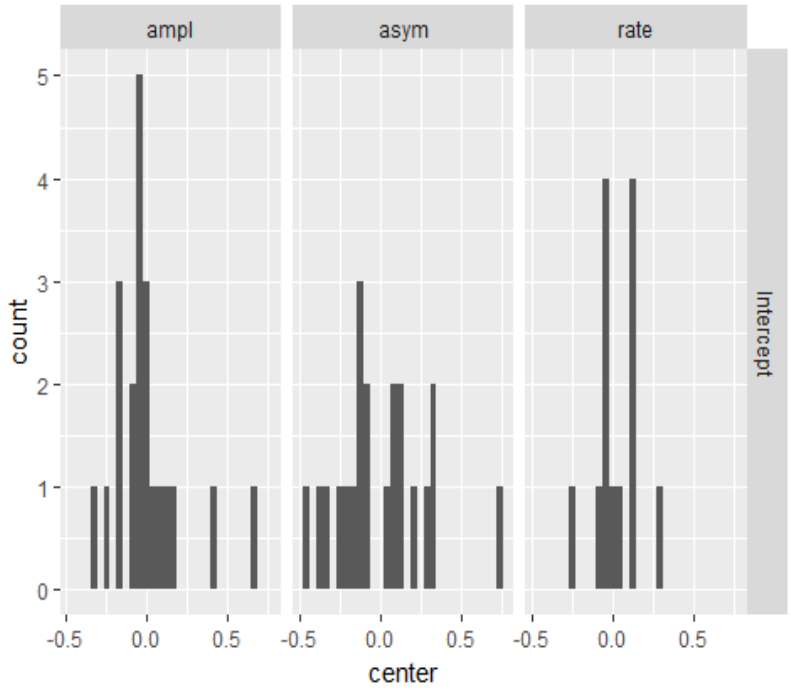


Figure 11. Participant-level random effects per parameter for wall-contact of task 2.

3.4 Internal Consistency Reliability of Simulator Tasks

In order to answer the question whether the basic simulator tasks are predictors of individual performance of technical bronchoscopic skills, we tested whether the three tasks correlate with each other. Firstly, correlations were estimated between the three tasks on population-level amplitude, rate and maximum performance for time-on-task. Table 2 presents the correlation between the three simulator tasks for the amplitude, rate and asymptote parameters. As maximum performance is our parameter of interest, we will mainly focus on this parameter. Next, correlations were estimated between population-level amplitude, rate and maximum performance for wall contacts during the Basic Scope Manipulation task.

To assess the certainty of the estimated correlations, we calculated 95% credibility limits. So, there is a 95% probability that the true value of the population parameter lies within the calculated range. Credibility limits running from the centre to higher values are most valuable, because they state that the true value could be higher than the observed value (Arendt et al., 2017). High certainty is indicated by credibility limits deviating only by a

maximum of 0.2, whereas credibility limits which include values between -0.1 and 0 indicate high uncertainty. In the current study, positive correlations were expected between the performances on the simulator tasks.

3.4.1 Time-on-task. A weak positive correlation was found between population-level maximum performance of Basic Scope Manipulation and Guided Anatomical Navigation (r = 0.32). Moreover, no correlation was found between population-level maximum performance of Basic Scope Manipulation and Step-By-Step Diagnostic Manoeuvres (r = -0.04). Finally, a weak positive correlation was found between population-level maximum performance of Guided Anatomical Navigation and Step-By-Step Diagnostic Manoeuvres (r = 0.31). All correlations of population-level ARY-parameters per task are presented in Table 2 on the next page. All correlations between the simulator tasks are highly uncertain (0.32, *95% CI* [-0.48; 0.90], -0.04, 95% *CI* [-0.90; 0.85] and 0.31, 95% *CI* [-0.88; 0.90], respectively). Although amplitude was not our main parameter of interest, we report the correlation for amplitude between Basic Scope Manipulation and Guided Anatomical Navigation since it was the highest correlation we found. A moderate positive correlation on the population-level amplitude parameter between the tasks Basic Scope Manipulation and Guided Anatomical Navigation (0.58, *95% CI* [-0.13; 0.94]). However, this outcome is highly uncertain too. Since the credibility limits for all four reported correlations include zero, it cannot be concluded with certainty if the simulator tasks are correlated at all. Although the reported correlations might suggest the possibility that simulator tasks are suitable as performance assessment tools, no further conclusions can be drawn.

Table 2

Correlation between simulator tasks for the three parameters

Parameter	Correlation1	Correlation2	Estimate	Lower CI95	Upper CI95
Amplitude	Basic Scope Manipulation	Guided Anatomical Navigation	0.58	-0.13	0.94
Amplitude	Basic Scope Manipulation	Step-by-Step Diagnostic Manoeuvres	0.25	-0.82	0.94
Amplitude	Guided Anatomical Navigation	Step-by-Step Diagnostic Manoeuvres	0.24	-0.81	0.92
Rate	Basic Scope Manipulation	Guided Anatomical Navigation	0.32	-0.48	0.90
Rate	Basic Scope Manipulation	Step-by-Step Diagnostic Manoeuvres	-0.04	-0.90	0.85
Rate	Guided Anatomical Navigation	Step-by-Step Diagnostic Manoeuvres	0.31	-0.79	0.95
Asymptote	Basic Scope Manipulation	Guided Anatomical Navigation	-0.41	-0.92	0.47
Asymptote	Basic Scope Manipulation	Step-by-Step Diagnostic Manoeuvres	-0.11	-0.88	0.91
Asymptote	Guided Anatomical Navigation	Step-by-Step Diagnostic Manoeuvres	-0.02	0.88	0.90

Note. 95% estimated credibility limits in square brackets.

4. Discussion

The aim of the current study was to explore whether basic bronchoscopic simulator tasks can be used to allow for a valid prediction of bronchoscopic simulator performance. Firstly, we were interested whether individuals show learning in all three simulator tasks by decreased time-on-task. Secondly, we explored whether individuals differ in acquiring basic bronchoscopic skills by exploring individual learning curves. Finally, it was analysed whether

the three simulator tasks correlate with each other. Individual learning curves were estimated to answer these research questions.

4.1 Findings

Bronchoscopy is a highly complex procedure which requires the integration of several skills to accurately control the bronchoscope without creating discomfort and trauma. To compensate for reduced depth perception and disturbed hand-eye coordination, surgeons are required to possess psychomotor ability. In the current study, the tasks Basic Scope Manipulation, Guided Anatomical Navigation and Step-by-Step Diagnostic Manoeuvres of the Simbionix GI-BRONCH Mentor™ were selected to explore how individuals learn complex psychomotor skills, such as bronchoscopic procedures.

4.1.1 Bronchoscopic skill acquisition. Firstly, we checked whether individuals show learning in all three simulator tasks by decreased time-on-task. We found that all participants improved their performance over time for all three tasks, as the time to complete each trial decreased. In line with previous studies, improvement rates slowed down after a few trials. Thus, we were able to estimate individual learning curves based on time-on-task as our main performance parameter.

4.1.2 Individual differences in bronchoscopic skill acquisition. Since candidate selection makes only sense if individuals differ at learning bronchoscopic skills, we explored whether individuals differ in acquiring basic bronchoscopic tasks. Based on our analysis and visually exploring individual learning curves, we were able to identify differences in performance between participants. Exploring variation of random effects revealed that maximum performance on tasks was scattered across individuals. This means that the population mean of performance is not representative for individuals, since many individuals deviate from the mean. We agree with previous studies reporting individual differences in

MIS-skill acquisition (Grantcharov & Funch-Jensen, 2009; Louridas et al., 2017). Thus, our findings emphasize the need for effective individualized training and assessment.

Based on our analysis, the largest individual differences were found in the Basic Scope Manipulation task. This finding is in line with Pastis (2014), who stated that the time-parameter for Basic Scope Manipulation differentiates between performance levels. Moreover, participants did not only differ at the maximum performance on all tasks, large individual differences were also found for amplitude and rate. These findings were highly uncertain. However, effect sizes showed that the first task was completed the fastest which may indicate that the Basic Scope Manipulation task is the easiest to complete.

For the second task, Guided Anatomical Navigation, individual differences were clearly visual in learning curves between initial performance and maximum performance. As most curves were parallel, it might indicate that initial performance predicts maximum performance. However, due to high uncertainty of our results more research is required.

4.1.3 Internal consistency of simulator tasks. To explore which of the three tasks contribute to a reliable and valid psychometric test, we tested whether the three tasks correlate with each other in the performance parameter time-on-task. Based on our analysis, we cannot state with certainty whether or not the used basic bronchoscopy simulator tasks are suitable predictors of basic bronchoscopic technical skill acquisition. Since internal consistency between the three basic simulator tasks was low and highly uncertain, it remains unclear if there is internal consistency at all between Basic Scope Manipulation, Guided Anatomical Navigation and Step-by-Step Diagnostic Manoeuvres. Notably, we found the highest internal consistency between Basic Scope Manipulation and Guided Anatomical Navigation for amplitude. Hence, certainty was low. Overall, our findings implicate weak reliability between three basic bronchoscopy simulator tasks, which questions the potential for using the simulator tasks as psychometric testing. We concluded that, at this moment, the three tests

together are not a valid objective assessment for selection of prospective surgeons. However, it is important to note that our study faces some limitations, which will be discussed later.

A possible explanation of these findings could be that the three tasks require different skills. The highest internal consistency were found between Basic Scope Manipulation and Guided Anatomical Navigation for all three learning parameters. These two tasks show the highest similarity between tasks, since both depend mainly on being able to navigate the scope. However, the second task may be more cognitive complex, since participants keep searching for unidentified lightbulbs in an environment which might feel like a labyrinth due to a lack of anatomical knowledge. In contrast, the first and third task do not show any internal consistency at all. A possible explanation is that the Step-by-Step Diagnostic Manoeuvres task deviates the most from the first task, since it requires a relatively small part of repeatedly navigating to the same tissue whereas keeping the scope still and controlling tools represent the main part of the task. The increased cognitive complexity might be also an explanation for the weak correlations between the second and third task. However, it is important to note the high uncertainty of all correlations.

We expected the third task to be the most advanced since participants had to integrate what they learned so far in terms of manoeuvring the scope while acquiring additional fine motor skills to use the diagnostic tool. However, participants reached maximum performance at a similar pace to the first task, which is against our expectations. Notably was the finding that for the Step-by-Step Diagnostic Manoeuvres, participants showed low initial performance, followed by a high increase in performance and resulting in constant performance quite early. Results for the third task are highly uncertain in particular due to an extreme small sample size, which again emphasizes the need to interpret current findings with great care. Nevertheless, a possible explanation for the fast increase in performance could be that the Step-by-Step Diagnostic Manoeuvres task placed higher demands on cognitive

resources rather than on psychomotor ability. It may be possible that in terms of psychomotor challenge the Step-by-Step diagnostic Manoeuvres was similar or even lower to prior tasks, while it was more difficult to understand the aim of the task. This task differed from the first and second task as participants were handed an additional tool to operate the forceps and brush. It was observed that participants asked questions more frequently about the third task compared to the first and second task, which may indicate that it was unclear to participants what was expected from them during the Step-by-Step Diagnostic Manoeuvres. What also may have contributed to this is that we developed an instructional video ourselves, while for the first and second task we used instruction videos developed by experts of CHEST. Thus, the low initial performance could be explained by unclarity how to perform the task. Since manoeuvring the scope was similar to prior tasks, once the aim and how-to of the task became clear, participants were able to apply frequently trained skills and complete the task faster.

4.2 Limitations of the Current Study

After having explored the results of the current study, we will address shortcomings within our study. Firstly, our concerns towards the lack of an external criterion to validate our findings will be discussed. Then, we will discuss our concerns towards using time-on-task as our main performance measure. After that, we will address limitations concerning the ability of the used tasks to discriminate between performance levels.

4.2.1 Lack of external validation. The main limitation of the current study is the lack of an external criterion to validate. Ultimately, the goal of bronchoscopic assessment is to use simulator performance as a predictor of real bronchoscopic performance. However, although the current study explored the internal consistency of the three simulator tasks, the external validity of simulator performance for real bronchoscopic performance was not covered. If a surgical test task highly correlates with valid external criteria, a surgical test can be considered valid. Both measures should reflect the same underlying construct (Stommel &

Wills, 2004). However, in the current study, we do not know whether simulator performance predicts real bronchoscopic performance, since this was not measured.

4.2.2 Time-on-task. A possible drawback of the current study is the use of time-on-task as a primary performance parameter. Although wall contact was used as a parameter for the first task, this measurement was not included for the second and third task. Procedure duration is an important patient safety factor due to possible complications and performance under moderate sedation, but airway trauma can have life-threatening consequences as well. A recent study of Mührmann et al. (2018) emphasizes considering both time and wall contact for an adequate performance estimation, as participants made more mistakes the more time they needed for completing a task. Bronchoscopy trainee selection based only on time performance would wrongly exclude those who perform slow but without mistakes, whereas fast but imprecise trainees would be allowed to operate on real patients.

4.2.3 Task discrimination. Another limitation of this study is that the Basic Scope Manipulation and Step-by-Step Diagnostic Manoeuvres tasks might be too easy to be able to discriminate between levels of psychomotor ability. For both tasks, all participants were able to reach maximum performance within low amounts of trials. A lack of difficulty leads to low task discrimination since it does not differentiate accurately between performance levels of participants. During the first task individuals had to follow a blue ball, which aimed for developing hand-eye coordination. During the Step-by-Step Diagnostic Manoeuvres task participants had to cover a relatively short distance by the bronchoscope before obtaining tissue samples. In addition, as we discussed before, we believe that the third task depends more on cognitive resources rather than on psychomotor ability. Thus, we question the ability of the first and third task to differentiate individuals who are talented enough for surgical trainees from those who are less suitable (yet).

4.3 Implications for Future Research

First of all, we encourage medical education institutes to implement and further develop performance-based selection and -training since individuals differ at bronchoscopic skill acquisition. Individuals show differences in the maximum performance they'll reach and the speed of bronchoscopic skill acquisition. Estimation of individual learning curves using repeated performance measures seems to be a promising method to predict individual's capacity, which enables individual-tailored surgery training to increase educational efficiency.

In the future, it will become crucial to explore the transferability of bronchoscopy simulator tasks to the real surgery room. It is not only relevant to test whether simulator-acquired skills are transferable to real life MIS, but also to examine whether simulator performance predicts real surgical performance. Besides technical psychomotor skills, surgeons also need to master working under pressure, cooperating within a team and making major decisions. Thus, on the long run, the relationship between basic simulator tasks and whole bronchoscopic procedures could be explored. Novices with a medical background, such as surgical trainees without prior endoscopic experiences but with a solid theoretical background would be suitable participants. However, since the prediction of bronchoscopic performance is still in its infancy, a stepwise approach is recommended to tackle this complex issue gradually.

We would recommend to replicate the current study with a greater sample size to ensure certainty and to implement some improvements. The first task can be used to get used to the simulator, for instance by performing a few exercise trials. Afterwards, the second should be performed similar to the current study. As findings of the third task are highly uncertain, the Step-by-Step Diagnostic Manoeuvres should be repeated as well. Here, instructions prior to the task should be improved. In general, we recommend to include both

time-on-task and wall contact as performance measures for all task to accurately establish individual capacity.

An useful extension to the current study would be to include expert performance as cut off scores. Currently, we are able to identify individual differences in maximum performance by means of individual learning curves, but we do not know yet which participants can be seen as talented. However, as the ultimate goal is to identify talented future bronchoscopists, it is crucial to establish reliable performance-based assessment guidelines. If a simulator task is proven as valid, the next step is to include a performance boundary to distinguish between suitable and unsuitable candidates. A study of Groenier, Schmettow and Huijser (2017) suggested that it might be possible to identify trainees who will struggle learning laparoscopy using simulator-based assessment. Individuals performing below or around an estimated cut-off score were permitted to continue surgery training, whereas two participants performing above the cut-off score will most likely not achieve proficiency. By applying an evidence-based cut-off score to the current study, it might also be possible to enable sufficient performing candidates to continue surgical training, whereas lower performing individuals would be rejected.

4.4 Implications for Medical Education

In terms of surgical performance assessment or -selection within medical education, current findings enable careful suggestions for implementation. Since selection of future surgeons has enormous consequences to student's future as well as to patients safety, we strongly recommend caution with interpreting our highly uncertain findings and emphasize the need for more research towards using basic bronchoscopy tasks as selection tools before implementing these. Nevertheless, we will present some suggestions for MIS-education in the next section.

As described earlier, we encourage medical education institutes to implement and further develop performance-based selection and -training, since individuals differ at bronchoscopic skill acquisition. However, we cannot say with certainty whether or not the three simulator tasks are a valid psychometric tool for predicting future basic bronchoscopic performance. Since the Basic Scope Navigation task revealed not to be an optimal predictor of simulator performance, we tend towards using this task as an exercise to get used to performing tasks on the BRONCH Mentor™ by means of a few exercise trials. In contrast, the Guided Anatomical Navigation task seems to be the most promising task as predictor of future surgical performance. This task enables estimating individual learning curves and differentiates between performance levels. Since our data suggest that initial performance might predict maximum performance, this needs further exploration. Next, on the first glance the Step-by-Step Diagnostic Manoeuvres task shows consistent results, but lacks predictive ability. However, due to extremely high uncertainty of the second and third task, more research is required on exploring the validity of these tasks as predictor of future performance. In general, to gain an accurate prediction of individual's surgical capacity, it is recommended to consider both time and accuracy as performance measures.

4.5 Conclusion

This study provides new insights about the applicability of individual learning curves to explore the acquisition of skills required for minimally invasive surgery. Based on our findings we can conclude that individuals differ at acquiring basic bronchoscopic skills. This, again, emphasizes the need for performance-based selection and assessment of future surgeons. Hence, due to high uncertainty, we were not able to conclude whether the three basic simulator tasks used in the current study are valid psychometric tools to predict basic bronchoscopic performance. As simulation-based prediction of bronchoscopic skills is still in

its infancy, further research is required to determine which basic tasks and skills underly real-life bronchoscopic procedures.

5 References

Arendt, A., Schmettow, M., & Groenier, M. (Master thesis, 2017). Towards reliable and valid prediction of MIS-performance with basic laparoscopic tasks in the LapSim and low-fi dexterity tasks. University of Twente. Retrieved from http://essay.utwente.nl/73211/1/arendt_MA_bms.pdf

Alvand, A., Auplish, S., Khan, T., Gill, H. S., & Rees, J. L. (2011). Identifying orthopaedic surgeons of the future. *Journal of Bone &Amp; Joint Surgery, British Volume*, *93–B*(12), 1586 LP-1591.

Breedveld, P., Stassen, H. G., Meijer, D. W., & Stassen, L. P. S. (1999). Theoretical background and conceptual solution for depth perception and eye-hand coordination problems in laparoscopic surgery. *Minimally Invasive Therapy & Allied Technologies*, *8*(4), 227–234. doi: 10.3109/13645709909153166

Bürkner, P.-C. (2018). brms: Bayesian Regression Models using “Stan.” Retrieved from https://cran.r-project.org/package=brms

Dalal, P. G., Dalal, G. B., Pott, L., Bezinover, D., Prozesky, J., & Bosseau Murray, W. (2011). Learning curves of novice anesthesiology residents performing simulated fibreoptic upper airway endoscopy. *Canadian Journal of Anesthesia/Journal Canadien d’anesthésie*, *58*(9), 802. doi: 10.1007/s12630-011-9542-2

Epstein, R. M., & Hundert, E. M. (2002). Defining and assessing professional competence. *JAMA*, *287*(2), 226–235. doi: 10.1001/jama.287.2.226

Ernst, A., & Herth, F. J. F. (2013). *Principles and practice of interventional pulmonology*. *Principles and Practice of Interventional Pulmonology*. doi: 10.1007/978-1-4614-4292-9

Ernst, A., Silvestri, G. A., & Johnstone, D. (2003). Interventional pulmonary procedures:

Guidelines from the American College of Chest Physicians. *Chest*, *123*(5), 1693–1717. doi: 10.1378/chest.123.5.1693

Ernst, A., Wahidi, M. M., Read, C. A., Buckley, J. D., Addrizzo-Harris, D. J., Shah, P. L., … Silvestri, G. A. (2015). Adult Bronchoscopy Training. *CHEST*, *148*(2), 321–332. doi: 10.1378/chest.14-0678

Fielding, D. I., Maldonado, F., & Murgu, S. (2014). Achieving competency in bronchoscopy: Challenges and opportunities. *Respirology*, *19*(4), 472–482. doi: 10.1111/resp.12279

Fuchs, K. H. (2002). Minimally invasive surgery. *Endoscopy*, *34*(2), 154–159.

Fitts PM, Posner MI. Human Performance. Brooks/Cole Pub. Co; Belmont, CA: 1967.

Gallagher, A. G., Cowie, R., Crothers, I., Jordan-Black, J.-A., & Satava, R. M. (2003). PicSOr: an objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopopic performance. *Surgical Endoscopy and Other Interventional Techniques*, *17*(9), 1468–1471.

Gallagher, A. G., Leonard, G., & Traynor, O. J. (2009). Role and feasibility of psychomotor and dexterity testing in selection for surgical training. *ANZ Journal of Surgery*, *79*(3), 108–113. doi: 10.1111/j.1445-2197.2008.04824.x

Gallagher, A. G., Smith, C. D., Bowers, S. P., Seymour, N. E., Pearson, A., McNatt, S., … Satava, R. M. (2003). Psychomotor skills assessment in practicing surgeons experienced in performing advanced laparoscopic procedures. *Journal of the American College of Surgeons*, *197*(3), 479–488. doi: 10.1016/S1072-7515(03)00535-0

Grantcharov, T. P. (2008). Is virtual reality simulation an effective training method in surgery? *Nature Reviews Gastroenterology & Hepatology*, *5*(5), 232.

Grantcharov, T. P., & Funch-Jensen, P. (2009). Can everyone achieve proficiency with the laparoscopic technique? Learning curve patterns in technical skills acquisition. *The American Journal of Surgery*, *197*(4), 447–449. doi:

https://doi.org/10.1016/j.amjsurg.2008.01.024

Groenier, M., Schraagen, J. M. C., Miedema, H. A. T., & Broeders, I. A. J. M. (2014). The role of cognitive abilities in laparoscopic simulator training. *Advances in Health Sciences Education*, *19*(2), 203–217. doi: 10.1007/s10459-013-9455-7

H, K. K., J, D. M., Christiaan, H., M, V. N. J. G., O, C. H. H., & N, P. J. E. (2016). The predictive value of aptitude assessment in laparoscopic surgery: a meta-analysis. *Medical Education*, *50*(4), 409–427. doi: 10.1111/medu.12945

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review*, *7*(2), 185–207. doi: 10.3758/BF03212979

Herth, F. J. F., Shah, P. L., & Gompelmann, D. (2017). Interventional Pulmonology: Retrieved from https://books.google.nl/books?id=Qg9BDwAAQBAJ

Hofstad, E. F., Våpenstad, C., Chmarra, M. K., Langø, T., Kuhry, E., & Mårvik, R. (2013). A study of psychomotor skills in minimally invasive surgery: what differentiates expert and nonexpert performance. *Surgical Endoscopy*, *27*(3), 854–863. doi: 10.1007/s00464-012-2524-9

Jabbour, N., Reihsen, T., Sweet, R. M., & Sidman, J. D. (2011). Psychomotor skills training in pediatric airway endoscopy simulation. *Otolaryngology--Head and Neck Surgery : Official Journal of American Academy of Otolaryngology-Head and Neck Surgery*, *145*(1), 43–50. doi: 10.1177/0194599811403379

Kaufman, H. H., Wiegand, R. L., & Tunick, R. H. (1987). Teaching surgeons to operate—Principles of psychomotor skills training. *Acta Neurochirurgica*, *87*(1), 1–7. doi: 10.1007/BF02076007

Konge, L., Arendrup, H., von Buchwald, C., & Ringsted, C. (2011). Using Performance in Multiple Simulated Scenarios to Assess Bronchoscopy Skills. *Respiration*, *81*(6), 483–

490.

Konge, L., Clementsen, P., Larsen, K. R., Arendrup, H., Buchwald, C., & Ringsted, C. (2012). Establishing Pass/Fail Criteria for Bronchoscopy Performance. *Respiration*, *83*(2), 140–146.

Lee, K., Junghans, B. M., Ryan, M., Khuu, S., & Suttle, C. M. (2014). Development of a novel approach to the assessment of eye–hand coordination. *Journal of Neuroscience Methods*, *228*, 50–56. doi: https://doi.org/10.1016/j.jneumeth.2014.02.012

Louridas, M., Szasz, P., Fecso, A. B., Zywiel, M. G., Lak, P., Bener, A. B., … Grantcharov, T. P. (2017). Practice does not always make perfect: need for selection curricula in modern surgical training. *Surgical Endoscopy*, *31*(9), 3718–3727. doi: 10.1007/s00464-017-5572-3

Luursema, J.-M., Buzink, S. N., Verwey, W. B., & Jakimowicz, J. J. (2010). Visuo-spatial ability in colonoscopy simulator training. *Advances in Health Sciences Education : Theory and Practice*, *15*(5), 685–694. doi: 10.1007/s10459-010-9230-y

Maan, Z. N., Maan, N. I., Darzi, A. W., & Aggarwal, R. (2012). Systematic review of predictors of surgical performance. *BJS*, *99*(12), 1610–1621. doi: 10.1002/bjs.8893

Madan, A. K., Frantzides, C. T., Park, W. C., Tebbit, C. L., Kumari, N. V. A., & O'Leary, P. J. (2005). Predicting baseline laparoscopic surgery skills. *Surgical Endoscopy*, *19*(1), 101–104. doi: 10.1007/s00464-004-8123-7

National Aeronautics and Space Administration (NASA). (n.d.). NASA TLX. Retrieved May 15, 2018, from https://humansystems.arc.nasa.gov/groups/tlx/

Ost, D., DeRosiers, 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. doi: 10.1164/ajrccm.164.12.2102087

Ouellette, D. R. (2006). The safety of bronchoscopy in a pulmonary fellowship program.

Chest, *130*(4), 1185–1190. doi: 10.1378/chest.130.4.1185

Pache, B., Hubner, M., Jurt, J., Demartines, N., & Grass, F. (2017). Minimally invasive surgery and enhanced recovery after surgery: The ideal combination? *Journal of Surgical Oncology*, *116*(5), 613–616. doi: 10.1002/jso.24787

Papachristofi, O., Jenkins, D., & Sharples, L. D. (2016). Assessment of learning curves in complex surgical interventions: a consecutive case-series study. *Trials*, *17*, 266. doi: 10.1186/s13063-016-1383-4

Pastis, N. J., Vanderbilt, A. A., Tanner, N. T., Silvestri, G. A., Huggins, J. T., Svigals, Z., & Shepherd, R. W. (2014). Construct Validity of the Simbionix Bronch Mentor Simulator for Essential Bronchoscopic Skills. *Journal of Bronchology & Interventional Pulmonology*, *21*(4).

R Core Team. (2018). R: A Language and Environment for Statistical Computing. Retrieved from https://www.r-project.org/

Sadideen, H., Hamaoui, K., Saadeddin, M., & Kneebone, R. (2012). Simulators and the simulation environment: Getting the balance right in simulation-based surgical education. *International Journal of Surgery*, *10*(9), 458–462. doi: https://doi.org/10.1016/j.ijsu.2012.08.010

Samia, H., Khan, S., Lawrence, J., & Delaney, C. P. (2013). Simulation and Its Role in Training. *Clinics in Colon and Rectal Surgery*, *26*(1), 47–55. doi: 10.1055/s-0033-1333661

Schijven, M. P., & Jakimowicz, J. (2004). The learning curve on the Xitact LS 500 laparoscopy simulator: profiles of performance. *Surgical Endoscopy*, *18*(1), 121–127. doi: 10.1007/s00464-003-9040-x

Schmettow, M. (2018a). ARY: Learning curve functions and formulas. Retrieved from https://rdrr.io/github/schmettow/asymptote/man/ARY.html

Schmettow, M. (2018b). *New Statistics for the Design Researcher - A Bayesian course in tidy R*.

Silvennoinen, M., Mecklin, J. P., Saariluoma, P., & Antikainen, T. (2009). Expertise and skill in minimally invasive surgery. *Scandinavian Journal of Surgery*.

Simbionix. (2018a). Essential Bronchoscopy [video]. Retrieved from http://simbionix.com/simulators/bronch-mentor/

Simbionix. (2018b). Essential Bronchoscopy Module. Retrieved from http://simbionix.com/simulators/bronch-mentor/bronch-library-of-modules/bronch-bronchoscopy/

Simoff, M. J. (2018). Learning to Look Through the Bronchoscope. *CHEST*, *148*(2), 301–303. doi: 10.1378/chest.15-0551

Stather, D. R., MacEachern, P., Chee, A., Dumoulin, E., & Tremblay, A. (2013). Trainee Impact on Procedural Complications: An Analysis of 967 Consecutive Flexible Bronchoscopy Procedures in an Interventional Pulmonology Practice. *Respiration*, *85*(5), 422–428. doi: 10.1159/000346650

Stefanidis, D., Korndorffer Jr, J. R., Black, F. W., Dunne, J. B., Sierra, R., Touchard, C. L., … Scott, D. J. (2006). Psychomotor testing predicts rate of skill acquisition for proficiency-based laparoscopic skills training. *Surgery*, *140*(2), 252–262. doi: 10.1016/j.surg.2006.04.002

Stommel, M., & Wills, C. (2004). *Clinical research: Concepts and principles for advanced practice nurses*. Lippincott Williams & Wilkins.

Tsui, C., Klein, R., & Garabrant, M. (2013). Minimally invasive surgery: national trends in adoption and future directions for hospital strategy. *Surgical Endoscopy*, *27*(7), 2253–2257. doi: 10.1007/s00464-013-2973-9

Wahidi, M. M., Silvestri, G. A., Coakley, R. D., Ferguson, J. S., Shepherd, R. W., Moses, L., … Downie, G. H. (2010). A prospective multicenter study of competency metrics and educational interventions in the learning of bronchoscopy among new pulmonary fellows. *Chest*, *137*(5), 1040–1049. doi: 10.1378/chest.09-1234

Wanzel, K. R., Ward, M., & Reznick, R. K. (2002). Teaching the surgical craft: From selection to certification. *Current Problems in Surgery*, *39*(6), 573–659.

Wentink, M., Breedveld, P., Meijer, D. W., & Stassen, H. G. (2000). Endoscopic camera rotation: a conceptual solution to improve hand-eye coordination in minimally-invasive surgery. *Minimally Invasive Therapy & Allied Technologies*, *9*(2), 125–131. doi: 10.3109/13645700009063059

White, C., Rodger, M., & Tang, T. (2016). Current understanding of learning psychomotor skills and the impact on teaching laparoscopic surgical skills. *The Obstetrician & Gynaecologist*, *18*(1), 53–63. doi: 10.1111/tog.12255

Appendix A. Materials

Figure 12. Handle to control the bronchoscope.

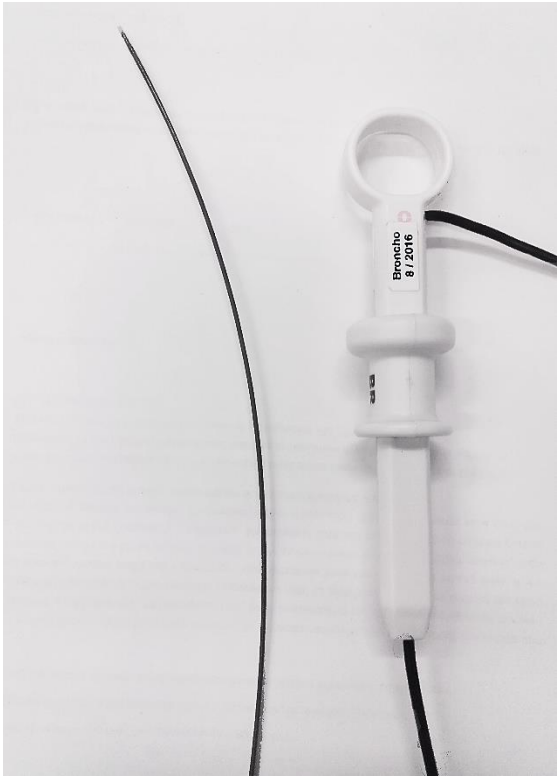


Figure 13. Sampling tool to insert in working channel

Appendix B. Instructions

1.1 General Introduction

This study consists of two simulator tasks. Before you start with a task, you will receive instructions on the specific task. First, you will be requested to fill out an informed consent form.

→ *Please provide the informed consent form to the participant and check whether the participant signs the form.*

After that, you are requested to fill out a demographics questionnaire with questions regarding your age, nationality and so on. Next, you will perform two tasks on the simulator. You have to repeat both tasks 20 times. Please do not hesitate to ask questions. Do you have any questions so far?

→ *Please provide the demographics questionnaire to the participant. Fill in participant's participant number, task number and session number.*

1.2 Explanation Tasks

Today you will practice two basic bronchoscopic tasks. Bronchoscopy is a procedure to examine the airways for abnormalities, such as infections, tumors or obstruction, by inserting a bronchoscope through the nose or mouths.

First, you will view two instruction videos about the use of the simulator and how to handle the scope. Instructions about the specific task will be presented on the screen before you start with the task.

→ *Please show the first two instruction videos to the participant.*

After watching the video instruction, you will start with the first basic scope handling task. You complete the first task, then the second task. After each trial, you have to fill out a short questionnaire. Repeat each task 20 times. Please try to be as accurate and quick as possible. Do not get discouraged from low scores, professionals train many years on these difficult tasks. We will mainly observe possible progress, less focus is on your actual performance.

1.2.1 Instruction Video Task 1 and 2 (MentorLearn, n.d.)

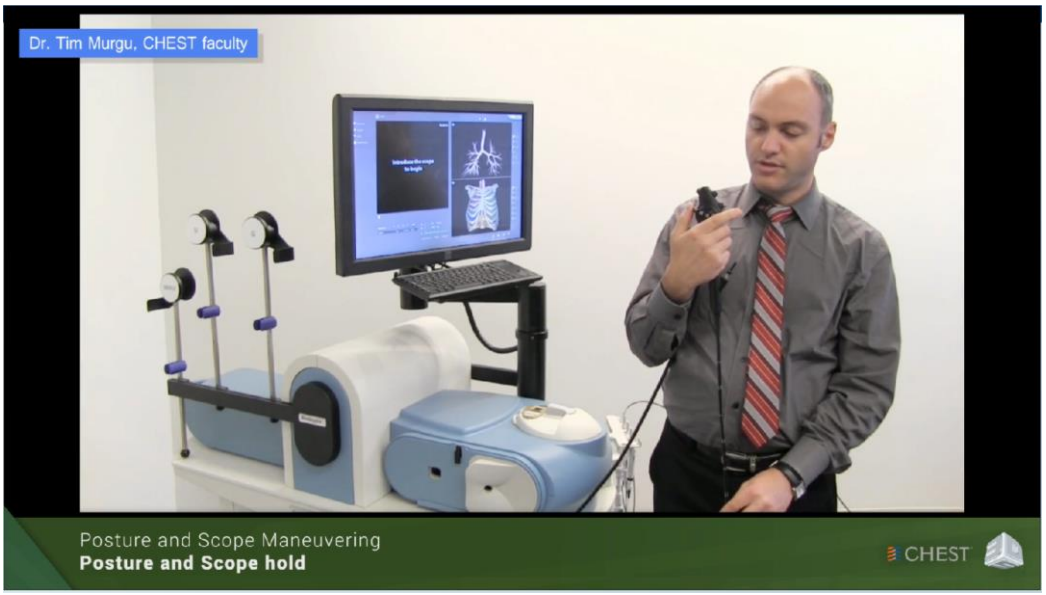


Figure 14 Instruction video bronchoscopy (Simbionix, n.d.)

1.2.2 Instruction Scope

How to use the bronchoscope:

- Control unit's lever: to move the bronchoscope's tip up and down
- Turn arm AND move the handle to navigate the scope
- Do NOT bend scope to a very small angle, this will cause damage to the fiberoptic fibres!
- Keep insertion tube straight when inserting the scope into the mouth.

1.2.3 Task 1. Basic Scope Manipulation (Simbionix, 2018c)

Learn to navigate the bronchoscope in a cyber environment to further develop hand-eye coordination.

Introduce the scope through the mouth until reaching a "Start" sign.

Navigate the scope in a narrowing industrial lumen, following a guiding light.

Keep insertion tube straight and use the control unit's roll and lever to keep scope's tip in mid-lumen and avoid wall contact.

Your score will drop with each wall contact based on lumen's width. The path is randomly selected by the software each time you start.

Good Luck!

1.2.4 Task 2. Guided Anatomical Navigation (Simbionix, 2018c)

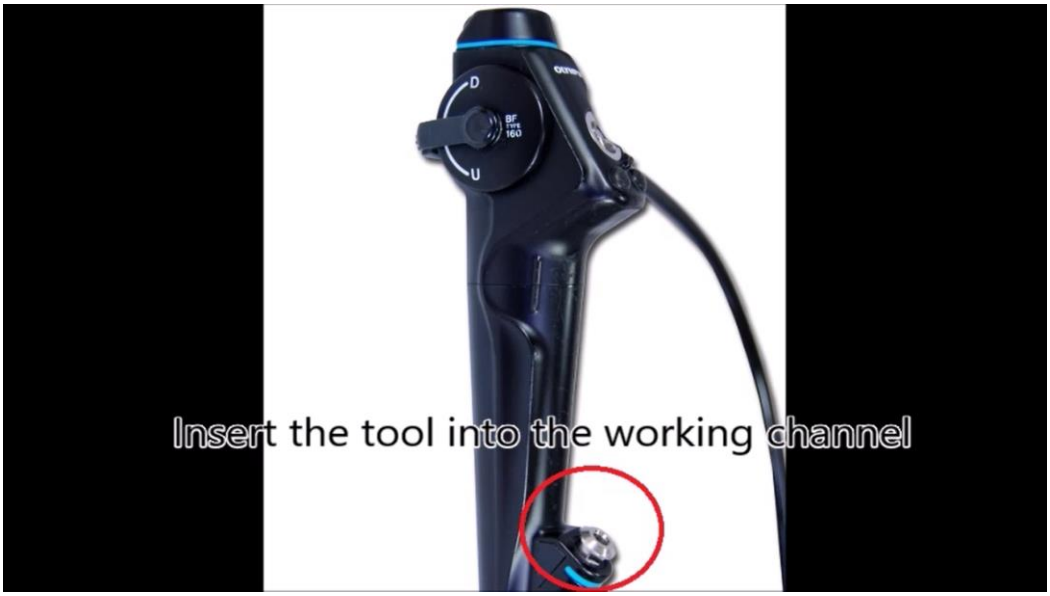
Learn and practice correct scope navigation within the anatomical environment.

Perform a complete airway inspection using directional guidance: match the scope's 'viewfinder' to a light bulb figure, which indicates the optimal scope roll and flex/extent for each anatomical carina.

Audio visual indications will show whether the maneuver was performed satisfactorily.

This task was developed in conjunction with and endorsed by the American Association for Bronchology and Interventional Pulmonology – AABIP"

1.2.5 Instruction Video Task 3



1.2.6 Task 3. Step-By-Step Diagnostic Maneuvers (Simbionix, 2018c)

Acquire the skills needed to perform a forceps, brush and transbronchial needle biopsy. Follow the step by step instructions to obtain tissue sample from a defined region and get immediate feedback about the quality and the efficacy of the maneuver. This task was developed in conjunction with and endorsed by the American Association for Bronchology and Interventional Pulmonology – AABIP

Appendix C. Informed Consent

Informed Consent

Title Research: Learning bronchoscopy on a simulator

Doctor(s) Directing Research: Dr. Martin Schmettow, Dr. Marleen Groenier

Undergraduate students conducting experiments: Marlise Westerhof, Luise Warnke

'I hereby declare that I have been informed in a manner which is clear to me about the nature and method of the research. My questions have been answered to my satisfaction. I agree of my own free will to participate in this research. I reserve the right to withdraw this consent without the need to give any reason and I am aware that I may withdraw from the experiment at any time. If my research results are to be used in scientific publications or made public in any other manner, then they will be made completely anonymous. My personal data will not be disclosed to third parties without my express permission. If I request further information about the research, now or in the future, I may contact Marlise Westerhof

(m.w.westerhof@student.utwente.nl).

If you have any complaints about this research, please direct them to the secretary of the Ethics Committee of the Faculty of Behavioural Sciences at the University of Twente, Drs. L. Kamphuis-Blikman P.O. Box 217, 7500 AE Enschede (NL), telephone: +31 (0)53 489 3399; email: l.j.m.blikman@utwente.nl).

Signed in duplicate:

…… …… ……

Name subject Signature Date

I have provided explanatory notes about the research. I declare myself willing to answer to the best of my ability any questions which may still arise about the research.'

…… …… ……

Name researcher Signature Date

Appendix D. Figures

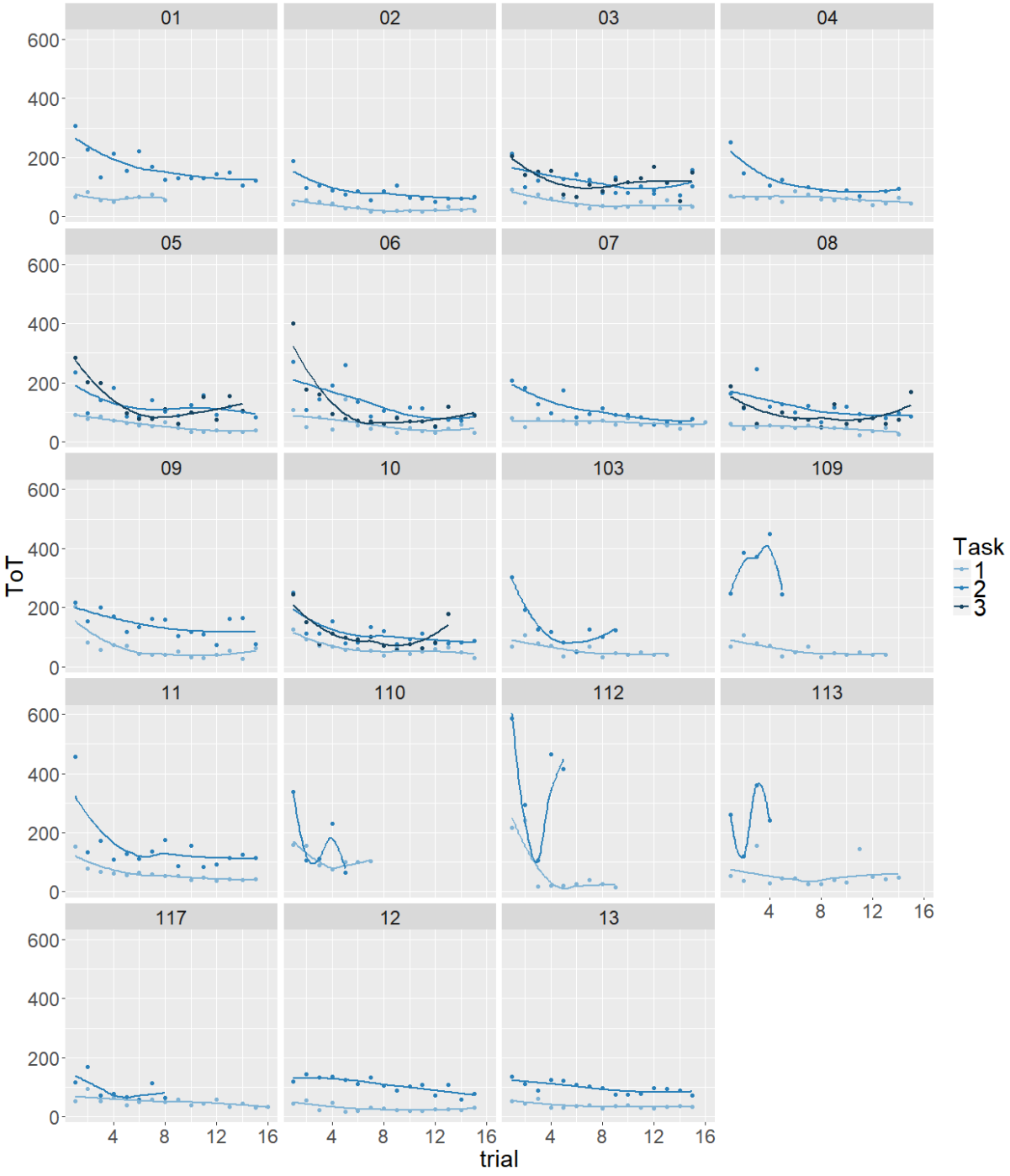


Figure 15. Raw time-on-task for the three simulator tasks per individual. Time-on-task (in seconds) on the y-axis, amount of trials on the x-axis (low values on the y-axis indicate superior performance).

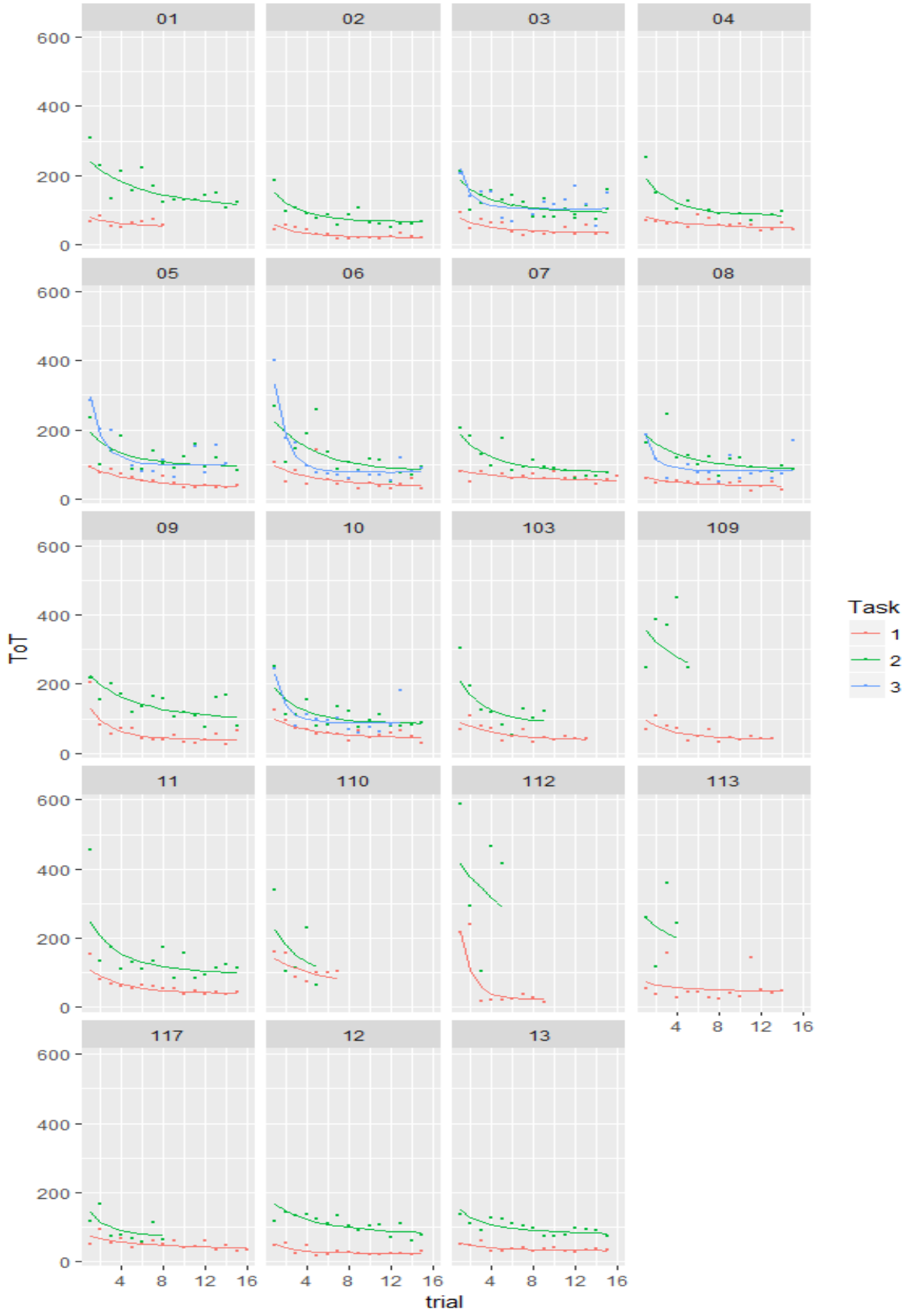


Figure 16. Estimated individual learning curves for maximum performance on time-on-task. Low values on the y-axis indicate superior performance.

Appendix E. R-Syntax

Data analysis

Author: Martin Schmettow

Date: 20 July, 2018

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

purp.data = F
purp.mcmc = F

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(stringr)
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.

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:brms':
##
##     fixef, ranef

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

load("MW18.Rda")
```

Data preparation

```
sim_task1 <-
  c(
  dir(path = "raw_data/MW/",
      pattern = "^Participant\\d{2}_Participant\\d{2}_Essential Bronchoscop
y_columns.csv",
      full.names = T),
  dir(path = "raw_data/AK/",
      pattern = "^Participant\\d{3}_Participant\\d{3}_Essential Bronchoscop
y_columns.csv",
     full.names = T,
     recursive = T)
  ) %>% print()

##  [1] "raw_data/MW/Participant01_Participant01_Essential Bronchoscopy_col
umns.csv"
##  [2] "raw_data/MW/Participant02_Participant02_Essential Bronchoscopy_col
umns.csv"
##  [3] "raw_data/MW/Participant03_Participant03_Essential Bronchoscopy_col
umns.csv"
##  [4] "raw_data/MW/Participant04_Participant04_Essential Bronchoscopy_col
umns.csv"
##  [5] "raw_data/MW/Participant05_Participant05_Essential Bronchoscopy_col
umns.csv"
##  [6] "raw_data/MW/Participant06_Participant06_Essential Bronchoscopy_col
umns.csv"
##  [7] "raw_data/MW/Participant07_Participant07_Essential Bronchoscopy_col
umns.csv"
##  [8] "raw_data/MW/Participant08_Participant08_Essential Bronchoscopy_col
umns.csv"
##  [9] "raw_data/MW/Participant09_Participant09_Essential Bronchoscopy_col
umns.csv"
## [10] "raw_data/MW/Participant10_Participant10_Essential Bronchoscopy_col
umns.csv"
## [11] "raw_data/MW/Participant11_Participant11_Essential Bronchoscopy_col
umns.csv"
## [12] "raw_data/MW/Participant12_Participant12_Essential Bronchoscopy_col
umns.csv"
## [13] "raw_data/MW/Participant13_Participant13_Essential Bronchoscopy_col
umns.csv"
## [14] "raw_data/AK//Participant103/Participant103_Participant103_Essentia
l Bronchoscopy_columns.csv"
## [15] "raw_data/AK//Participant109/Participant109_Participant109_Essentia
l Bronchoscopy_columns.csv"
```

```
## [16] "raw_data/AK//Participant110/Participant110_Participant110_Essentia
l Bronchoscopy_columns.csv"
## [17] "raw_data/AK//Participant112/Participant112_Participant112_Essentia
l Bronchoscopy_columns.csv"
## [18] "raw_data/AK//Participant113/Participant113_Participant113_Essentia
l Bronchoscopy_columns.csv"
## [19] "raw_data/AK//Participant117/Participant117_Participant117_Essentia
l Bronchoscopy_columns.csv"
```

```
sim_task23 <-
  c(
  dir(path = "raw_data/MW/",
      pattern = "^Participant\\d{2}_Participant\\d{2}_Essential Bronchoscop
y_rows.csv",
      full.names = T),
  dir(path = "raw_data/AK/",
      pattern = "^Participant\\d{3}_Participant\\d{3}_Essential Bronchoscop
y_rows.csv",
     full.names = T,
     recursive = T)
  ) %>% print()
```

```
##  [1] "raw_data/MW/Participant01_Participant01_Essential Bronchoscopy_row
s.csv"
##  [2] "raw_data/MW/Participant02_Participant02_Essential Bronchoscopy_row
s.csv"
##  [3] "raw_data/MW/Participant03_Participant03_Essential Bronchoscopy_row
s.csv"
##  [4] "raw_data/MW/Participant04_Participant04_Essential Bronchoscopy_row
s.csv"
##  [5] "raw_data/MW/Participant05_Participant05_Essential Bronchoscopy_row
s.csv"
##  [6] "raw_data/MW/Participant06_Participant06_Essential Bronchoscopy_row
s.csv"
##  [7] "raw_data/MW/Participant07_Participant07_Essential Bronchoscopy_row
s.csv"
##  [8] "raw_data/MW/Participant08_Participant08_Essential Bronchoscopy_row
s.csv"
##  [9] "raw_data/MW/Participant09_Participant09_Essential Bronchoscopy_row
s.csv"
## [10] "raw_data/MW/Participant10_Participant10_Essential Bronchoscopy_row
s.csv"
## [11] "raw_data/MW/Participant11_Participant11_Essential Bronchoscopy_row
s.csv"
## [12] "raw_data/MW/Participant12_Participant12_Essential Bronchoscopy_row
s.csv"
## [13] "raw_data/MW/Participant13_Participant13_Essential Bronchoscopy_row
s.csv"
## [14] "raw_data/AK//Participant103/Participant103_Participant103_Essentia
l Bronchoscopy_rows.csv"
## [15] "raw_data/AK//Participant109/Participant109_Participant109_Essentia
l Bronchoscopy_rows.csv"
## [16] "raw_data/AK//Participant110/Participant110_Participant110_Essentia
l Bronchoscopy_rows.csv"
```

```
## [17] "raw_data/AK//Participant112/Participant112_Participant112_Essentia
l Bronchoscopy_rows.csv"
## [18] "raw_data/AK//Participant113/Participant113_Participant113_Essentia
l Bronchoscopy_rows.csv"
## [19] "raw_data/AK//Participant117/Participant117_Participant117_Essentia
l Bronchoscopy_rows.csv"
```

```
time_task23 <-
  c(
    dir(path = "raw_data/MW/",
        pattern = "Participant\\d{2}_task[23]_time.xlsx",
        full.names = T),
    dir(path = "raw_data/AK/",
        pattern = "Participant\\d{3}_task2_time.xls",
        full.names = T,
        recursive = T)
  ) %>% print()
```

```
##  [1] "raw_data/MW/Participant01_task2_time.xlsx"
##  [2] "raw_data/MW/Participant02_task2_time.xlsx"
##  [3] "raw_data/MW/Participant03_task2_time.xlsx"
##  [4] "raw_data/MW/Participant03_task3_time.xlsx"
##  [5] "raw_data/MW/Participant04_task2_time.xlsx"
##  [6] "raw_data/MW/Participant05_task2_time.xlsx"
##  [7] "raw_data/MW/Participant05_task3_time.xlsx"
##  [8] "raw_data/MW/Participant06_task2_time.xlsx"
##  [9] "raw_data/MW/Participant06_task3_time.xlsx"
## [10] "raw_data/MW/Participant07_task2_time.xlsx"
## [11] "raw_data/MW/Participant08_task2_time.xlsx"
## [12] "raw_data/MW/Participant08_task3_time.xlsx"
## [13] "raw_data/MW/Participant09_task2_time.xlsx"
## [14] "raw_data/MW/Participant10_task2_time.xlsx"
## [15] "raw_data/MW/Participant10_task3_time.xlsx"
## [16] "raw_data/MW/Participant11_task2_time.xlsx"
## [17] "raw_data/MW/Participant12_task2_time.xlsx"
## [18] "raw_data/MW/Participant13_task2_time.xlsx"
## [19] "raw_data/AK//Participant103/Participant103_task2_time.xls"
## [20] "raw_data/AK//Participant109/Participant109_task2_time.xls"
## [21] "raw_data/AK//Participant110/Participant110_task2_time.xls"
## [22] "raw_data/AK//Participant112/Participant112_task2_time.xls"
## [23] "raw_data/AK//Participant113/Participant113_task2_time.xls"
## [24] "raw_data/AK//Participant117/Participant117_task2_time.xls"
```

```
read_sim_task1 <- function(x) {
  out <- read_csv(x)
  colnames(out) <- str_replace_all(colnames(out), "\\s", "")
#  print(colnames(out))
  out %>%
    dplyr::mutate(
      Part = str_extract(LastName, "\\d+"),
      Task = as.integer(1),
      Device = "BronchoSim",
      File = as.character(x),
      trial = as.integer(Repetition - min(Repetition) +1),
      Researcher = if_else(str_detect(x, "AK"), "AK", "MW"),
```

```
      Route = if_else(Researcher == "AK", NA_character_, as.character(Text
5)),
      ToT = if_else(Researcher == "AK", as.character(Text1), as.character(T
ext4)),
      ToT = as.numeric(seconds(hms(ToT))),
      wall_contacts = if_else(Researcher == "AK", NA_integer_, as.integer(N
um3)),
      perc_wall = as.numeric(if_else(Researcher == "AK", NA_real_, as.numer
ic(Num2)))
      )  %>%
  select(Part, Task, trial, Route, ToT, wall_contacts, perc_wall, Device) #
# TEMPLATE
}
read_sim_task1(sim_task1[1])
read_sim_task1(sim_task1[18])



# read_sim_task23 <- function(x){
#   out <- read_csv(x)
#   colnames(out) <- str_replace_all(colnames(out), "\\s", "")
#   out %>%
#     mutate(
#     Part = str_extract(LastName, "\\d+"),
#       Task = as.integer(1),
#       Device = "BronchoSim",
#       File = as.character(x),
#       trial = as.integer(Repetition - min(Repetition) +1),
#       Researcher = if_else(str_detect(x, "AK"), "AK", "MW")) %>%
#     select(Part, Task, trial, Device)
#     #select(Part, Task, trial, Route, ToT, wall_contacts, perc_wall, Devi
ce)
# }
# str(read_sim_task23(sim_task23[1]))
# str(read_sim_task23(sim_task23[18]))
#



#
#
#
#
# Sim_task23 <-
#   set_names(sim_task23) %>%
#   map_df(read_sim_task23)


## HERE


read_time_task23 <- function(x) {
  out <- read_excel(x)
  print(colnames(out))
  if(! ("Time_Task3_Total" %in% colnames(out))) out$Time_Task3_Total <- NA
```

```
  if(! ("TimeOnTask" %in% colnames(out))) out$TimeOnTask <- NA
  out %>%
    mutate(Part = str_extract(as.character(Participant), "\\d+"),
           Device = "BronchoSim",
           trial = as.integer(Repetition),
           wall_contacts = NA_integer_,
           perc_wall =  NA_real_,
           Route = NA_character_,
           ToT = as.numeric(if_else(Task == 2, as.character(TimeOnTask), a
s.character(Time_Task3_Total)))) %>%
    select(Part, Task, trial, Route, ToT, wall_contacts, perc_wall, Device)
}
read_time_task23(time_task23[1])
read_time_task23(time_task23[4])
read_time_task23(time_task23[16])


Sim_task1 <-
  set_names(sim_task1) %>%
  map_df(read_sim_task1) %>%
  print()

Time_task23 <-
  set_names(time_task23) %>%
  map_df(read_time_task23) %>%
  print()

MW18 <-
  bind_rows(Sim_task1, Time_task23) %>%
  mutate(Task = as.factor(Task)) %>%
  filter(!is.na(ToT)) %>%
  mascutils::as_tbl_obs() %>%
  print()

summary(MW18)

save(MW18, file = "MW18.Rda")
```

Wall contact data

```
read_wall_1 <-
  function(x){
    read_excel(x) %>%
      select(Part = `Last Name`, `Case Number`, Repetition, )
  }
```

Data exploration

```
load("MW18.Rda")
```

Descriptives

Number of observations

```
MW18 %>%
  group_by(Part, Task) %>%
  summarize(N_trials = n()) %>%
  ungroup() %>%
  group_by(Task) %>%
  summarize(N_Part = n(),
            min(N_trials), median(N_trials), max(N_trials)) %>%
  knitr::kable()
```

Task	N_Part	min(N_trials)	median(N_trials)	max(N_trials)
1	19	7	15	16
2	19	4	15	20
3	5	13	14	15

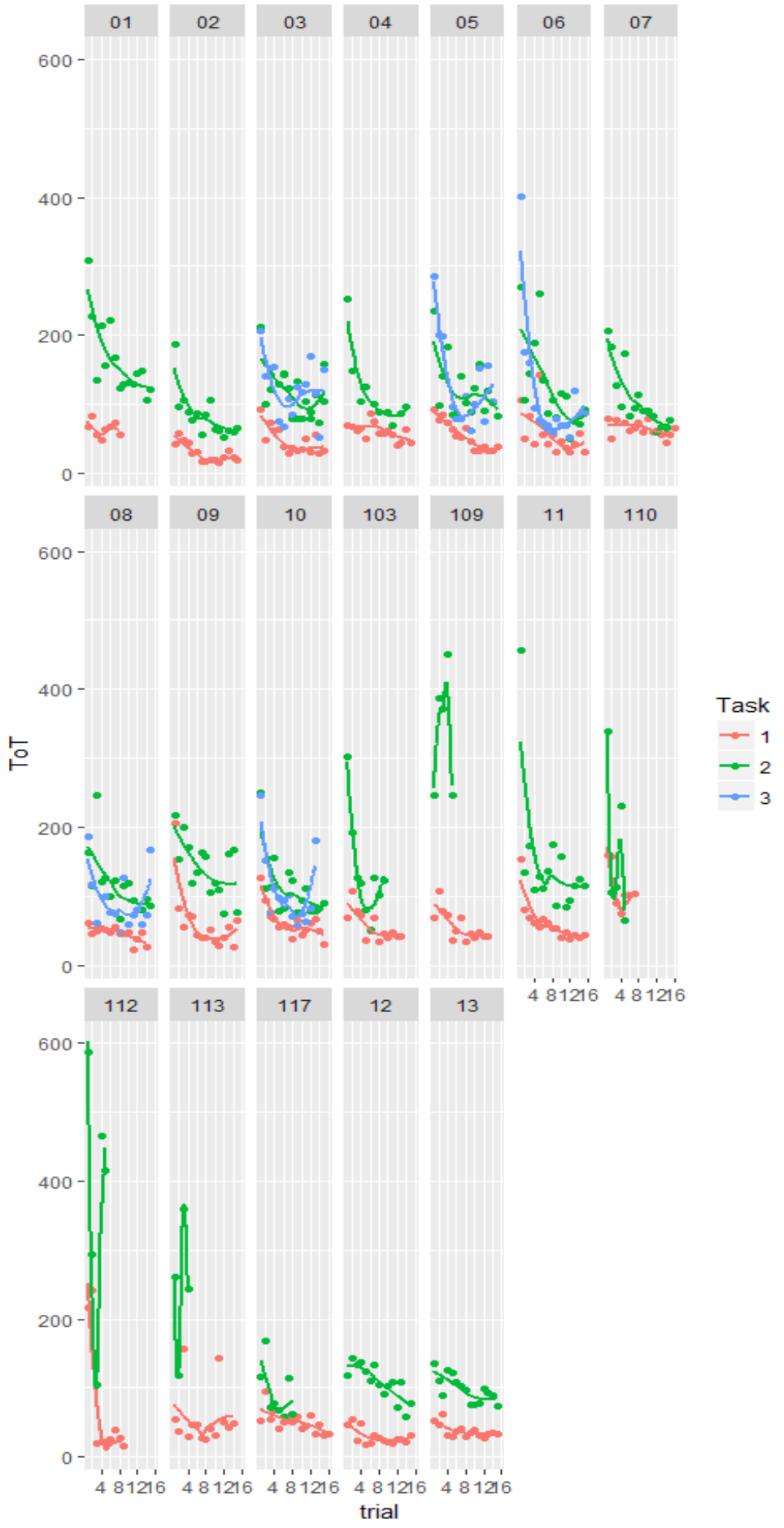
Plots

```
load("MW18.Rda")

MW18
```

Data set: showing 8 of 560 observations

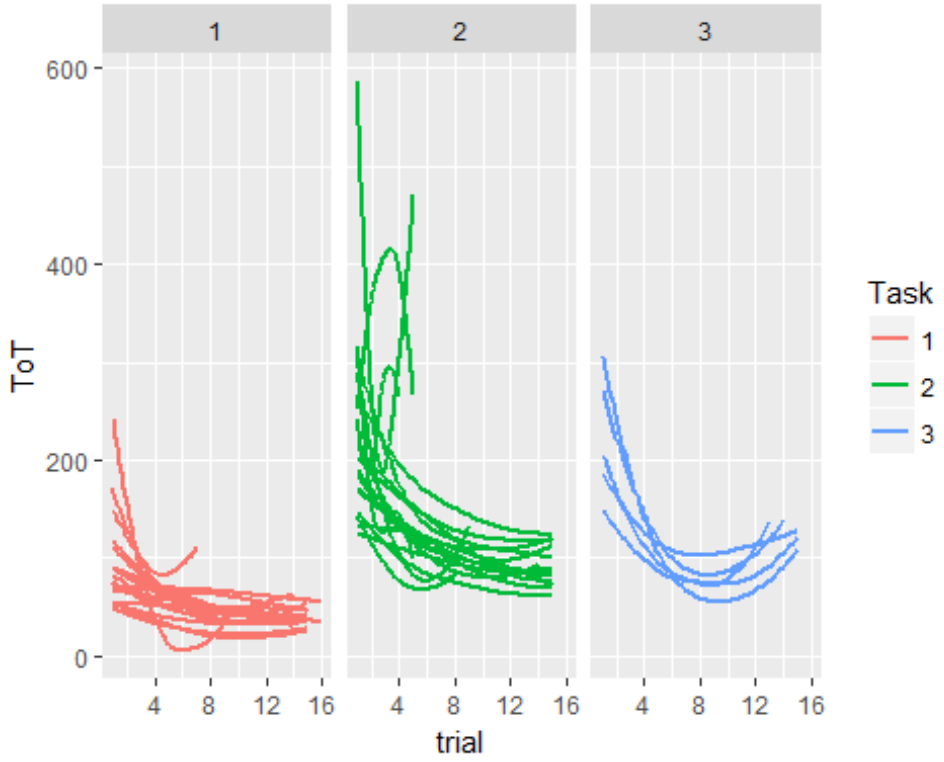
Obs	Part	Task	trial	Route	ToT	wall_contacts	perc_wall	Device
274	01	2	15	NA	121.80	NA	NA	BronchoSim
94	07	1	12	6	62.00	2	1	BronchoSim
96	07	1	14	6	44.00	2	8	BronchoSim
406	07	2	14	NA	65.73	NA	NA	BronchoSim
411	08	2	4	NA	120.00	NA	NA	BronchoSim
138	10	1	11	6	52.00	0	0	BronchoSim
204	109	1	4	NA	72.00	NA	NA	BronchoSim
553	117	2	1	NA	117.00	NA	NA	BronchoSim

```
MW18 %>%
  ggplot(aes(x = trial, color = Task, y = ToT)) +
  facet_wrap(~Part, ncol = 7) +
  geom_point() +
  geom_smooth(se = F, span = 1)

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

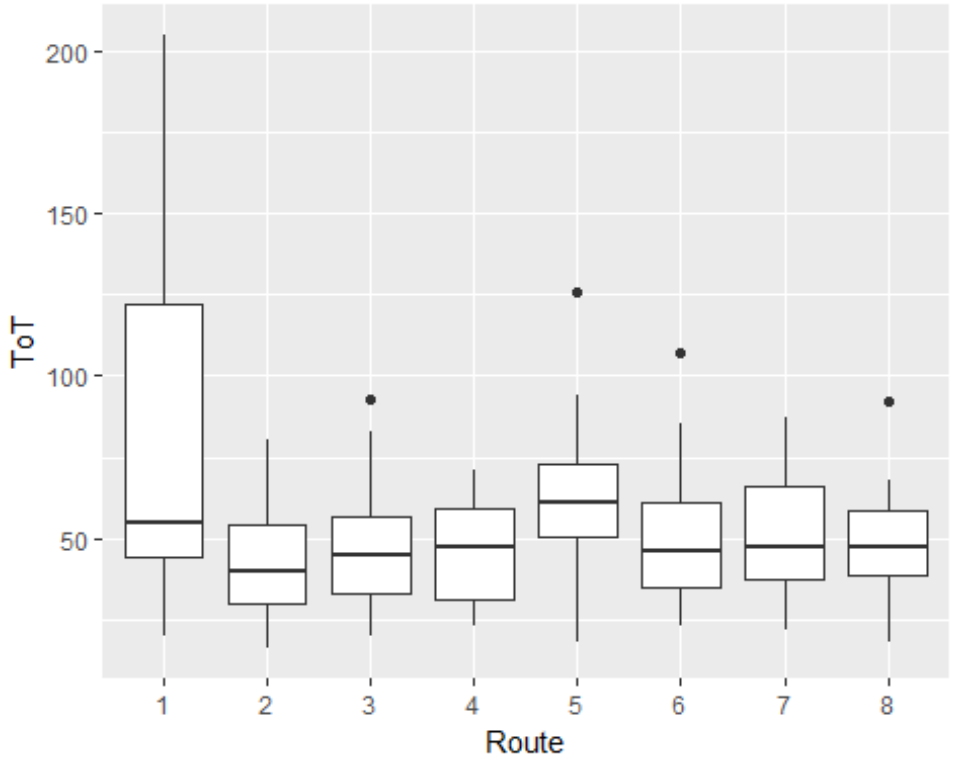

01
02
03
04
05
06
07
08
09
10
103
109
11
110
112
113
117
12
13
600
400
200
0
ToT
4 8 12 16
trial
Task
1
2
3

```
MW18 %>%
  ggplot(aes(x = trial, color = Task, y = ToT)) +
  facet_wrap(~Task, ncol = 3) +
  geom_smooth(se = F, aes(group = Part), span = 1.5)
```

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




```
MW18 %>%
  filter(!is.na(Route)) %>%
  ggplot(aes(x = Route, y = ToT)) +
  geom_boxplot()
```



Time on task

Setting up the LARY model:

```
lazyeval::f_lhs(LARY) <- quote(ToT)
LARY

## ToT ~ exp(ampl) * exp(-exp(rate) * trial) + exp(asym)
## <environment: namespace:asymptote>

# Random effects and correlations
F_ef_lary_1 <- list(
  formula(ampl ~ 0 + Task + (0 + Task|corr1|Part)),
  formula(rate ~ 0 + Task + (0 + Task|corr2|Part)),
  formula(asym ~ 0 + Task + (0 + Task|corr3|Part)))

# INcluding difficulty of Route (asym and ampl only)
F_ef_lary_2 <- list(
  formula(ampl ~ 0 + Task + Route + (0 + Task|corr1|Part)),
  formula(rate ~ 0 + Task + (0 + Task|corr1|Part)),
  formula(asym ~ 0 + Task + Route + (0 + Task|corr1|Part)))

# log scale weak priors
F_pr_lary_1 <- c(set_prior("normal(1, 5)", nlpar = "ampl"),
                 set_prior("normal(-1, 5)", nlpar = "rate"),
                 set_prior("normal(0.5, 5)", nlpar = "asym"))
```

M_2 LARY gamma

```
M_2 <-
  brm(bf(LARY,
         flist = F_ef_lary_1, nl = TRUE),
      prior = F_pr_lary_1,
      family = Gamma(link = identity),
      data = MW18,
      iter = 0)

M_2 <-
  brm(fit = M_2,
      data = MW18,
      iter = 3000, warmup = 2000, chains = 5,
      init = "0",
      control = list(adapt_delta = 0.999,
                     max_treedepth = 12))

save(M_2, file = "M_2.Rda")

# load("M_1.Rda")
load("M_2.Rda")
M_2
```

```
## Warning: There were 1 divergent transitions after warmup. Increasing ada
pt_delta above 0.999 may help.
## See http://mc-stan.org/misc/warnings.html#divergent-transitions-after-wa
rmup
```

```
##  Family: gamma
##   Links: mu = identity; shape = identity
## Formula: ToT ~ exp(ampl) * exp(-exp(rate) * trial) + exp(asym)
##          ampl ~ 0 + Task + (0 + Task | corr1 | Part)
##          rate ~ 0 + Task + (0 + Task | corr2 | Part)
##          asym ~ 0 + Task + (0 + Task | corr3 | Part)
##    Data: MW18 (Number of observations: 560)
## Samples: 5 chains, each with iter = 3000; warmup = 2000; thin = 1;
##          total post-warmup samples = 5000
##
## Group-Level Effects:
## ~Part (Number of levels: 19)
##                           Estimate Est.Error l-95% CI u-95% CI Eff.Samp
le
## sd(ampl_Task1)                0.81      0.23     0.44     1.33       20
84
## sd(ampl_Task2)                0.45      0.16     0.14     0.78       14
18
## sd(ampl_Task3)                0.81      0.92     0.04     2.91       22
64
## sd(rate_Task1)                0.66      0.25     0.27     1.25       15
44
## sd(rate_Task2)                0.63      0.27     0.12     1.23       11
45
## sd(rate_Task3)                0.47      0.52     0.01     1.78       28
```

```
16
## sd(asym_Task1)                 0.32      0.08     0.19     0.51        19
62
## sd(asym_Task2)                 0.21      0.12     0.04     0.51         6
39
## sd(asym_Task3)                 0.26      0.22     0.04     0.81        20
33
## cor(ampl_Task1,ampl_Task2)     0.53      0.28    -0.13     0.94        28
39
## cor(ampl_Task1,ampl_Task3)     0.19      0.51    -0.82     0.94        50
00
## cor(ampl_Task2,ampl_Task3)     0.18      0.49    -0.81     0.92        50
00
## cor(rate_Task1,rate_Task2)    -0.36      0.37    -0.92     0.47        13
24
## cor(rate_Task1,rate_Task3)     0.07      0.52    -0.88     0.91        50
00
## cor(rate_Task2,rate_Task3)     0.01      0.51    -0.88     0.90        50
00
## cor(asym_Task1,asym_Task2)     0.29      0.37    -0.48     0.90        16
24
## cor(asym_Task1,asym_Task3)    -0.04      0.51    -0.90     0.85        35
83
## cor(asym_Task2,asym_Task3)     0.24      0.49    -0.79     0.95        29
00
##                             Rhat
## sd(ampl_Task1)              1.00
## sd(ampl_Task2)              1.00
## sd(ampl_Task3)              1.00
## sd(rate_Task1)              1.00
## sd(rate_Task2)              1.00
## sd(rate_Task3)              1.00
## sd(asym_Task1)              1.00
## sd(asym_Task2)              1.00
## sd(asym_Task3)              1.00
## cor(ampl_Task1,ampl_Task2) 1.00
## cor(ampl_Task1,ampl_Task3) 1.00
## cor(ampl_Task2,ampl_Task3) 1.00
## cor(rate_Task1,rate_Task2) 1.00
## cor(rate_Task1,rate_Task3) 1.00
## cor(rate_Task2,rate_Task3) 1.00
## cor(asym_Task1,asym_Task2) 1.00
## cor(asym_Task1,asym_Task3) 1.00
## cor(asym_Task2,asym_Task3) 1.00
##
## Population-Level Effects:
##            Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
## ampl_Task1     4.19      0.23     3.69     4.61       2386 1.00
## ampl_Task2     5.11      0.18     4.75     5.45       2564 1.00
## ampl_Task3     6.04      0.59     4.86     7.11       2435 1.00
## rate_Task1    -1.26      0.33    -1.96    -0.68       1407 1.00
## rate_Task2    -1.35      0.30    -1.98    -0.78       2525 1.00
## rate_Task3    -0.04      0.36    -0.72     0.66       3862 1.00
## asym_Task1     3.59      0.12     3.34     3.82       1264 1.00
```

```
## asym_Task2     4.47      0.11     4.24     4.68       1554 1.00
## asym_Task3     4.53      0.13     4.26     4.79       2348 1.00
##
## Family Specific Parameters:
##       Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
## shape    12.21      0.78    10.76    13.80       5000 1.00
##
## Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
## is a crude measure of effective sample size, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
PP_2 <- post_pred(M_2)
P_2 <- posterior(M_2)
```

ARY parameters by task:

```
fixef(P_2)
```

Estimates with 95% credibility limits

nonlin	fixef	center	lower	upper
ampl	Task1	4.1943808	3.6947096	4.6142470
ampl	Task2	5.1095258	4.7509152	5.4521104
ampl	Task3	6.0573316	4.8625955	7.1077443
rate	Task1	-1.2435748	-1.9646393	-0.6779146
rate	Task2	-1.3437421	-1.9828356	-0.7752111
rate	Task3	-0.0424568	-0.7158604	0.6607809
asym	Task1	3.5984714	3.3381816	3.8178155
asym	Task2	4.4719664	4.2363012	4.6787783
asym	Task3	4.5278499	4.2584186	4.7850494

Individual differences as standard deviations by task and ARY parameters:

```
grpef(P_2)
```

Estimates with 95% credibility limits

nonlin	fixef	center	lower	upper
ampl	Task1	0.7746708	0.4436903	1.3254961
ampl	Task2	0.4352048	0.1447338	0.7776277
ampl	Task3	0.5950534	0.0362137	2.9078687
rate	Task1	0.6322952	0.2726400	1.2487957
rate	Task2	0.6184308	0.1166074	1.2326481
rate	Task3	0.3303223	0.0142953	1.7823352
asym	Task1	0.3149380	0.1852992	0.5107994
asym	Task2	0.1864147	0.0377349	0.5139578
asym	Task3	0.2072504	0.0438200	0.8087418

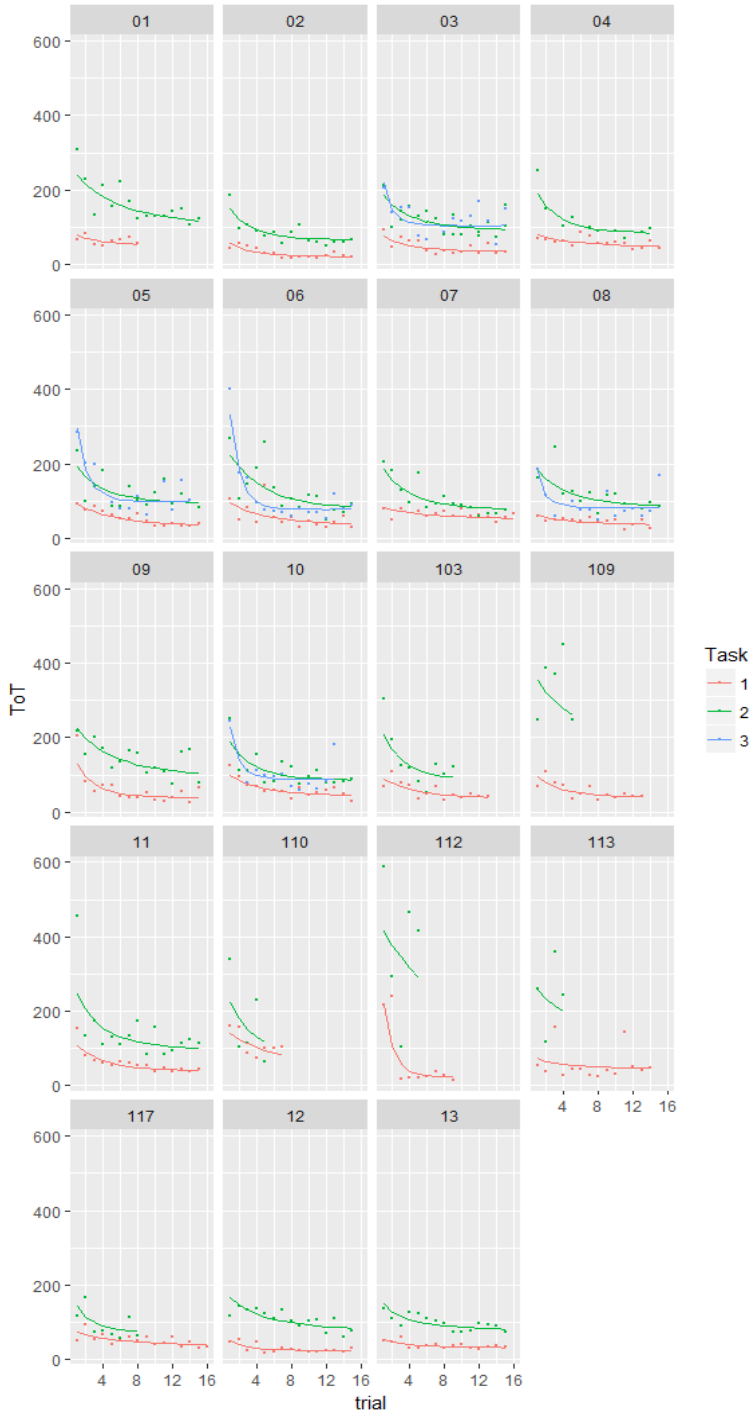
Correlations between tasks

```
P_2 %>%
  filter(type == "cor") %>%
  group_by(parameter) %>%
  summarize(center = median(value),
            lower = quantile(value, .025),
            upper = quantile(value, .975)) %>%
  separate(parameter, into = c("type", "level", "nonlin", "Cor_1", "X", "Co
r_2")) %>%
  select(nonlin, Cor_1, Cor_2, center, lower, upper) %>%
  knitr::kable()
```

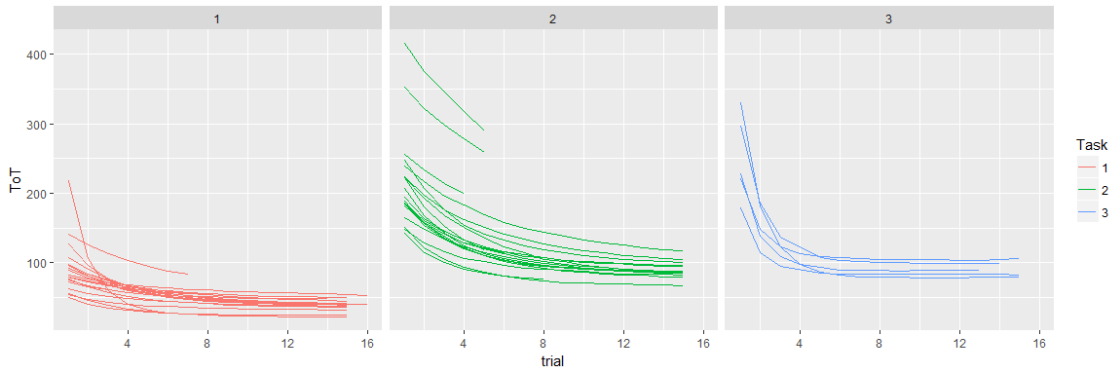
nonlin	Cor_1	Cor_2	center	lower	upper
ampl	Task1	Task2	0.5811647	-0.1277042	0.9379199
ampl	Task1	Task3	0.2485072	-0.8191381	0.9445333
ampl	Task2	Task3	0.2417580	-0.8086615	0.9227551
asym	Task1	Task2	0.3178753	-0.4804567	0.9013017
asym	Task1	Task3	-0.0408820	-0.8978128	0.8547597
asym	Task2	Task3	0.3051214	-0.7888903	0.9477258
rate	Task1	Task2	-0.4068122	-0.9150813	0.4688237
rate	Task1	Task3	0.1090322	-0.8797787	0.9093623
rate	Task2	Task3	0.0158613	-0.8832989	0.8977062

Estimated curves

```
T_pred_2 <-
  MW18 %>%
  filter(!is.na(ToT)) %>%
  bind_cols(predict(PP_2)) %>%
  mutate(resid = ToT - center)

T_pred_2 %>%
  ggplot(aes(x = trial, y = ToT, color = Task)) +
  facet_wrap(~Part, ncol = 4) +
  geom_point(size = .5) +
  geom_line(aes(y = center))
```




```
T_pred_2 %>%
  ggplot(aes(x = trial, y = ToT, color = Task)) +
  facet_grid(~Task) +
  geom_line(aes(y = center, group = Part))
```



Wall contacts

Only MW task 1.

```
MW18_1 <-
  MW18 %>%
  filter(!is.na(wall_contacts))

lazyeval::f_lhs(LARY) <- quote(wall_contacts)
LARY

## wall_contacts ~ exp(ampl) * exp(-exp(rate) * trial) + exp(asym)
## <environment: namespace:asymptote>

# INcluding difficulty of Route (asym and ampl only)
F_ef_lary_3 <- list(
  formula(ampl ~ (1|Route) + (1|Part)),
  formula(rate ~ (1|Part)),
  formula(asym ~ (1|Route) + (1|Part)))

M_9 <-
  brm(bf(LARY,
         flist = F_ef_lary_3, nl = TRUE),
      prior = F_pr_lary_1,
      family = negbinomial(link = "identity"),
      data = MW18_1,
      iter = 0)

M_9 <-
  brm(fit = M_9,
      data = MW18_1,
      iter = 7000, warmup = 6000, chains = 5,
      init = "0",
      control = list(adapt_delta = 0.999,
                     max_treedepth = 14))

save(M_9, file = "M_9.Rda")
```

Results

```
load("MW18.Rda")
load("M_9.Rda")
M_9
```

```
## Warning: There were 40 divergent transitions after warmup. Increasing ad
apt_delta above 0.999 may help.
## See http://mc-stan.org/misc/warnings.html#divergent-transitions-after-wa
rmup
```

```
##  Family: negbinomial
##   Links: mu = identity; shape = identity
## Formula: wall_contacts ~ exp(ampl) * exp(-exp(rate) * trial) + exp(asym)

##          ampl ~ (1 | Route) + (1 | Part)
##          rate ~ (1 | Part)
##          asym ~ (1 | Route) + (1 | Part)
##    Data: MW18_1 (Number of observations: 187)
## Samples: 5 chains, each with iter = 7000; warmup = 6000; thin = 1;
##          total post-warmup samples = 5000
##
## Group-Level Effects:
## ~Part (Number of levels: 13)
##                     Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
## sd(ampl_Intercept)     4.49      7.12     0.08    23.20        459 1.02
## sd(rate_Intercept)     4.14      6.91     0.08    21.44       1460 1.00
## sd(asym_Intercept)     0.54      1.01     0.08     3.43        289 1.03
##
## ~Route (Number of levels: 8)
##                     Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
## sd(ampl_Intercept)     3.98      5.91     0.13    20.08        646 1.01
## sd(asym_Intercept)     0.80      1.50     0.18     4.04        351 1.02
##
## Population-Level Effects:
##                Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
## ampl_Intercept    -1.04      4.14   -10.24     7.28       1043 1.00
## rate_Intercept     1.10      4.96    -8.39    10.83        118 1.04
## asym_Intercept     0.04      2.70    -9.37     1.39         33 1.08
##
## Family Specific Parameters:
##       Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
## shape    93.68     75.16    18.90   300.92       5000 1.00
##
## Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
## is a crude measure of effective sample size, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
PP_9 <- post_pred(M_9)
P_9 <- bayr::posterior(M_9)
```

Effects

```
fixef(P_9)
```

Estimates with 95% credibility limits

nonlin	center	lower	upper
ampl	-0.1824841	-10.237932	7.282350
rate	0.8006849	-8.387019	10.831161
asym	0.8926386	-9.373423	1.394162

```
ranef(P_9)
```

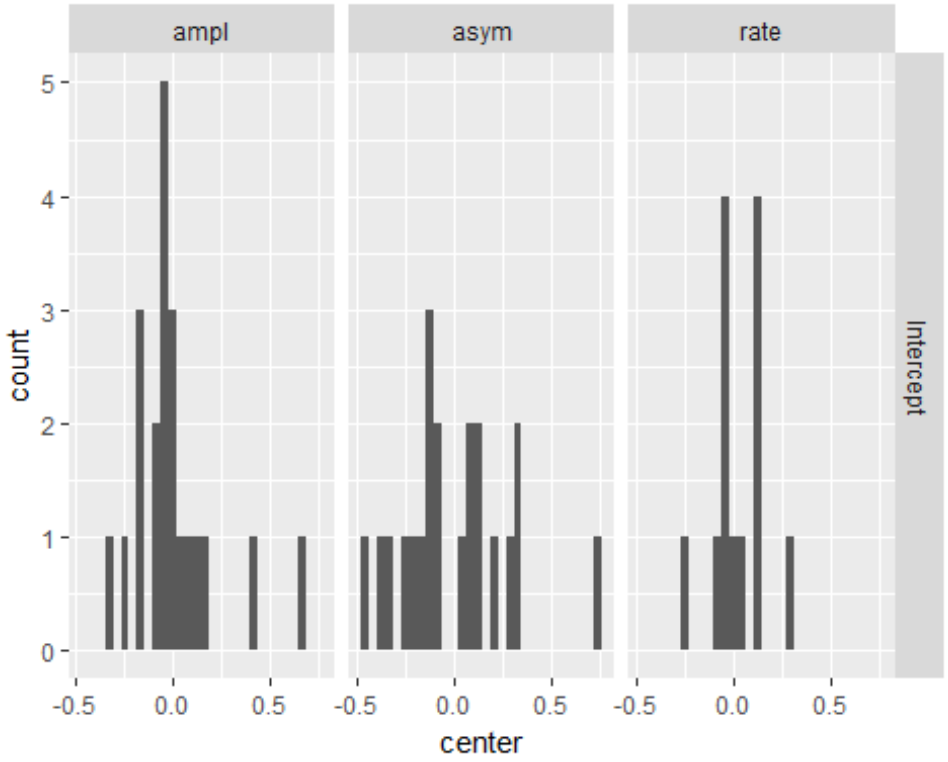
Estimates with 95% credibility limits

nonlin	re_factor	re_entity	center	lower	upper
ampl	Part	01	-0.0566498	-14.036144	15.1313308
ampl	Part	02	-0.0846545	-16.935435	12.5575751
ampl	Part	03	0.1167191	-15.947684	13.8439650
ampl	Part	04	-0.2401105	-16.688141	14.0824638
ampl	Part	05	-0.0364995	-17.307063	13.2686907
ampl	Part	06	0.4250249	-14.244032	15.7143520
ampl	Part	07	0.0151634	-18.020819	14.9717433
ampl	Part	08	-0.0597071	-15.843575	13.5402749
ampl	Part	09	0.1514147	-15.917277	13.8141085
ampl	Part	10	-0.0170094	-18.685505	13.8250378
ampl	Part	11	-0.1033819	-16.892387	12.8325827
ampl	Part	12	-0.0093607	-15.871976	14.5011475
ampl	Part	13	-0.1539399	-15.582159	14.8929855
ampl	Route	1	0.6730245	-14.417557	10.5551243
ampl	Route	2	-0.0284467	-15.269412	9.9487764
ampl	Route	3	-0.3447310	-14.021446	9.8437629
ampl	Route	4	-0.1571457	-14.771124	11.1173688
ampl	Route	5	0.0275770	-13.874029	10.4371165
ampl	Route	6	-0.0576554	-11.861099	12.4032597
ampl	Route	7	0.0717612	-14.563880	10.1191907
ampl	Route	8	-0.1607724	-15.550736	10.2770832
rate	Part	01	0.1065363	-11.394661	13.4468901
rate	Part	02	0.1138518	-11.123216	15.0732223
rate	Part	03	-0.0525814	-12.081566	13.9099210
rate	Part	04	0.3108274	-9.108650	15.3207403
rate	Part	05	0.0603058	-11.938251	14.7582919
rate	Part	06	-0.2580708	-14.048107	10.4328893
rate	Part	07	-0.0496702	-12.503352	15.3179828
rate	Part	08	0.0137353	-11.005440	13.2733856
rate	Part	09	-0.0670921	-12.139208	13.4847741
rate	Part	10	-0.0297927	-12.571042	13.8352183

rate	Part	11	0.1240276	-10.914212	13.0128373
rate	Part	12	-0.0260685	-11.778488	13.6417917
rate	Part	13	0.1397181	-11.361845	15.4831365
asym	Part	01	-0.1307191	-1.323643	0.5949877
asym	Part	02	-0.1111998	-1.028139	0.8273757
asym	Part	03	0.3150331	-1.011366	1.2856180
asym	Part	04	-0.4622728	-1.490668	0.6044488
asym	Part	05	-0.0634973	-1.175250	0.9078342
asym	Part	06	0.3183323	-1.024252	1.2405218
asym	Part	07	0.1420462	-1.141776	0.9493688
asym	Part	08	0.0212297	-1.300549	0.9405796
asym	Part	09	0.2161349	-1.087014	0.9994735
asym	Part	10	0.0968759	-1.186679	0.7868770
asym	Part	11	-0.1465149	-1.066409	0.7933707
asym	Part	12	0.0741776	-1.167493	0.7419694
asym	Part	13	-0.2578311	-1.478133	0.5536772
asym	Route	1	0.7398130	-1.748289	1.7894998
asym	Route	2	-0.1387354	-1.668497	0.9207195
asym	Route	3	-0.3120050	-2.024660	0.9115068
asym	Route	4	-0.1983074	-2.048098	0.9382771
asym	Route	5	0.1238265	-1.442983	1.1220045
asym	Route	6	-0.3579690	-1.730914	1.0438951
asym	Route	7	0.2872138	-1.447545	1.3155564
asym	Route	8	-0.0952104	-1.740833	1.2592476

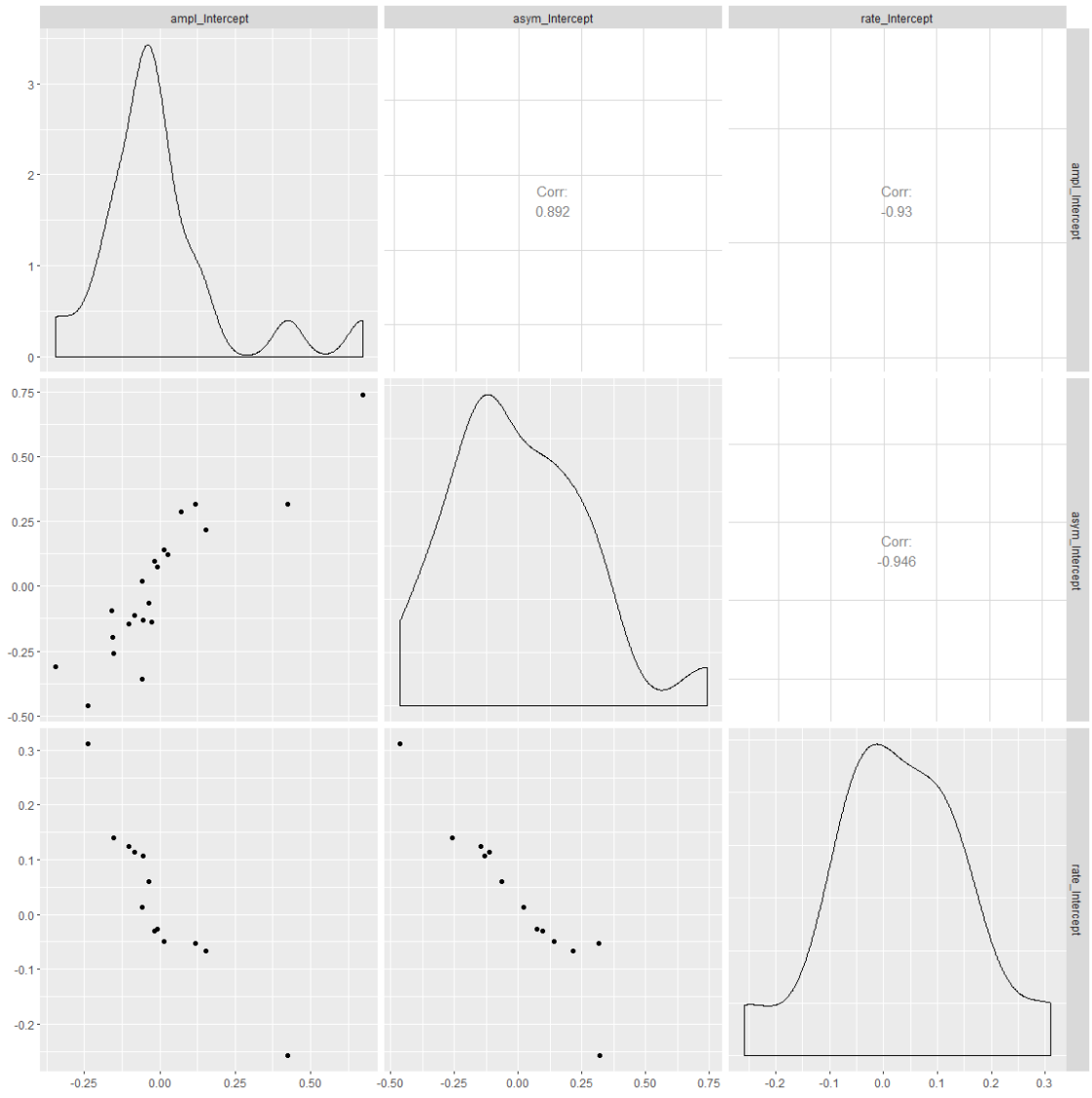
```
ranef(P_9) %>%
  rename(Task = fixef) %>%
  ggplot(aes(x = center)) +
  facet_grid(Task~nonlin) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Correlations

```
T_ranef_wide <-
  ranef(P_9) %>%
  filter(nonlin != "offset") %>%
  mutate(parameter = str_c(nonlin, fixef, sep = "_")) %>%
  select(parameter, Part = re_entity, center) %>%
  spread(key = parameter, value = center)

T_ranef_wide %>%
  select(-Part) %>%
  GGally::ggpairs()

## Warning in (function (data, mapping, alignPercent = 0.6, method =
## "pearson", : Removed 8 rows containing missing values

## Warning in (function (data, mapping, alignPercent = 0.6, method =
## "pearson", : Removed 8 rows containing missing values

## Warning: Removed 8 rows containing missing values (geom_point).

## Warning: Removed 8 rows containing missing values (geom_point).

## Warning: Removed 8 rows containing non-finite values (stat_density).
```


ampl_Intercept
asym_Intercept
rate_Intercept
Corr:
0.892
Corr:
-0.93
Corr:
-0.946
ampl_Intercept
asym_Intercept
rate_Intercept