Towards a novel approach to applied research: The role of motor sequence learning in the process of mastering complex motor procedures

Lida Z. David

Master Thesis

Department of Cognitive Psychology and Ergonomics
Faculty of Behavioral, Management and Social Sciences (BMS)

April, 2018

1st supervisor: dr. Martin Schmettow
2nd supervisor: dr. Marleen Groenier

UNIVERSITY OF TWENTE.

Acknowledgment

I would like to thank my supervisor, dr. Martin Schmettow, without whom this thesis would not have been possible. His immersive knowledge on programming and statistics, as well as his bright ideas and willingness to share and explain those to me, were invaluable to the composition of this paper.

I would also like to thank dr. Marleen Groenier, for her wide knowledge around laparoscopic research and her valuable feedback that helped me develop, structure, and present my thoughts cohesively throughout this paper. I am also thankful for the opportunity she gave me to perform some laparoscopic simulator tasks and experience myself the demands of laparoscopy (my performance on which proved I am a far better writer than I would ever be a surgeon – hopefully).

Lastly, I would like to express my gratitude to my family for their treasured and unconditional support throughout this thesis, and especially to my grandparents, for digitally accompanying me throughout these long study hours.

Abstract

Background: Research strives for the development of an optimised laparoscopic training process to help trainees acquire all skills necessary for surgery. Laparoscopic simulator tasks pose a promising alternative to training procedures. Yet the repetitive nature of practice using such tasks runs the risk of exercising only motor sequence learning (MSL); that is, the automation of a specific sequence of movements (Verwey, 1996). Sufficient task variability needs to be ensured, to allow acquisition of holistic, generalizable skills.

Methods: Our research attempted to shed light into the role of MSL on performance, to better understand its implications to skill acquisition, if it were promoted in laparoscopic tasks. We thus introduced the *varied-fixed learning (VFL)* paradigm, consisting of dexterity tasks with a varied and a fixed part, the former allowing participants to practice holistic skill acquisition blocking the practice of MSL, while the latter allowing them to practice MSL. Both parts were modelled on a two-part exponential learning curve, assessing the learning parameters amplitude, rate, and asymptote for both parts. This paradigm enabled us to capture the proportion of MSL to the general motor learning process involved in mastering a new complex procedure, which was estimated from the difference between the two amplitudes.

Results: Findings suggested little involvement of MSL to general motor learning. The proportion of MSL was not very pronounced in neither of the two tasks (mirror-drawing = 0.10, *95% CI* [0.04; 0.22]; clips-and-string = 0.05, *95% CI* [0.00; 0.55]), and uncertainty was high in both. However, MSL's proportion should not be overlooked, since this was only a pilot study with highly uncertain results, implying the possibility of a more pronounced proportion of MSL.

Conclusions: Findings suggest that MSL has a small, but still detectable role in improving performance on complex motor procedures. This underlines the importance of variability in simulator tasks, to ensure surgeons master transferable skills in training. Our research affirmed the feasibility of the VFL paradigm and learning curve model to investigate learning, which offers a promising alternative to research in applied settings such as laparoscopy. Having ensured its usefulness, more exhaustive investigation is needed to reach final conclusions on the MSL/general motor learning ratio.

Keywords: laparoscopy, motor sequence learning, learning curves, dexterity tasks.

Table of Contents

1. Introduction

Laparoscopic skill acquisition has received great attention within the field of medical education. However, there is still considerable ambiguity regarding the processes involved in learning and mastering new skills required for this type of surgery, while training conditions for optimised skill acquisition remain unknown (Cook et al., 2012). Advances in technology have enabled the development of laparoscopic simulators with immense possibilities, since they enable trainees to surpass the classical learning routines of watching procedures, attending lectures, and practicing on actual patients (Nácul, Cavazzola, & Melo, 2015), into performing simulated surgery without risking human lives. Yet, such simulators are currently not incorporated in training procedures. There are still great individual differences observed amongst learning using laparoscopic simulators, with some individuals even failing to acquire all skills necessary for surgery (Grantcharov & Fuch-Jensen, 2009). The medical education society thus faces the challenge of understanding the process of acquiring the complex skills required for laparoscopic surgery. Only then can one develop an optimal way of assessing and adequately training surgeons. Establishing an optimised training process requires the development of effective simulator tasks, whose complexity and variability would foster learning. Thus, this is paper aims to examine the concept and processes comprising learning and how these can be affected by task nature, in an attempt to inform the development of optimised laparoscopic tasks.

In the following sections we first introduce the reader to the *resemblance spectrum*, which depicts the shared similarities of simulator tasks to low-fidelity, complex motor tasks. This forms the groundwork for understanding the importance of researching learning under conditions that closely resemble a real learning environment. We then propose a new framework for learning, derived from observations made during the execution of such tasks. Based on this framework, a novel experimental paradigm and a model for the analysis of learning curves are constructed, both used in our pilot study, aiming to understand learning and encourage experts to move towards research that is more suitable for applied settings.

1.1. The resemblance spectrum

The resemblance spectrum highlights the importance for creation of more realistic experimental conditions when conducting psychological research for applied settings. It places possible test suites on a continuum with two extremes ranging from not resembling laparoscopy, to complete resemblance (see figure 1). Those that are further away from complete resemblance lack features essential to real surgery, and are thus not as predictive of actual laparoscopic performance. For example, early psychometric studies developed surgeon selection tests similar to those used in other skilled professions (Gallagher & Smith, 2003), to assess certain innate abilities considered related to laparoscopic performance, such as visuo-spatial, perceptual, and psychomotor abilities. However, these abilities did not always succeed in predicting performance (Groenier et al., 2014; Huijser, 2015). Examination of the resemblance spectrum explains these findings by suggesting that innate abilities lack the complexity that exists in laparoscopic surgery, and are thus not predictive of performance. On the other hand, low-fidelity dexterity tasks pose a closer resemblance to real laparoscopic surgery since they inter-combine cognitive and psychomotor demands along with manual dexterity abilities, all of whom are entailed in real surgery. Unlike simplified experimental conditions that represent just a building block of all skills required for task execution, learning complex motor procedures needed for the execution of low-fidelity dexterity tasks, permits all properties associated to the complex use of cognitive and other abilities to be captured.

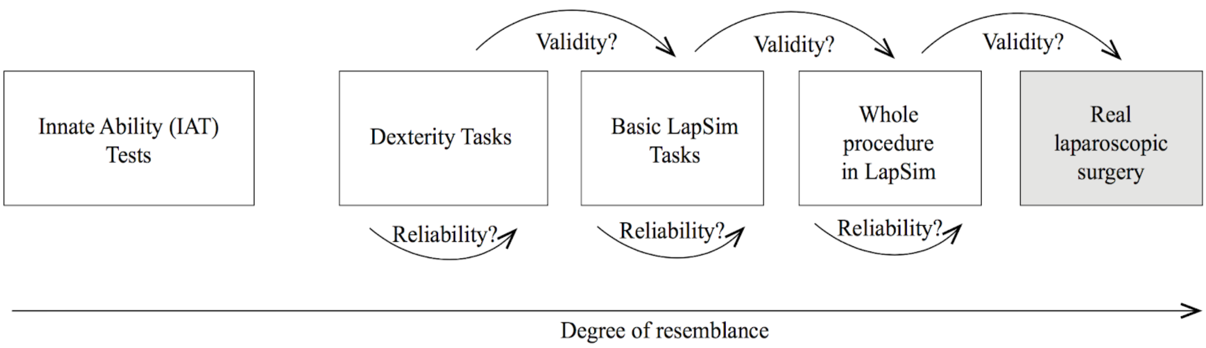


Figure 1. The resemblance spectrum. Test suites placed on a continuum of ascending order (left to right), based on their resemblance to real laparoscopic surgery. The validity of the links between suites that are placed next to each other is stronger than those placed further apart. For example, dexterity tasks have a more valid link to resembling basic tasks than resembling a whole laparoscopic procedure.

Research on the investigation of learning curves for dexterity tasks has made an important observation regarding correlations in learning curves, which can be proven vital to the understanding of the underlying process involved in mastering such tasks. More specifically, studies at the University of Twente investigated participants while performing repeated trials of the same dexterity task, until reaching their maximum performance for that task (Arendt, 2017; Kaschub, 2016). While a strong correlation was initially found in participants' learning curves across tasks (r = 0.73, Kaschub, 2016), this correlation weakened greatly when the repetitive pattern of the same motor sequence required for task execution was eliminated across trials, by introducing greater demand for movement variability (r = 0.33, Arendt, 2017). This significant drop in learning curve correlation revealed that memorisation of a specific motor sequence, a process known as *motor sequence learning* (MSL; Verwey, 1996; Willingham, 2001), could be a main reason for individual differences in task learning. This suggests that MSL may be part of one's general motor learning process, with some tasks promoting its contribution more than others, through being less varied or complex.

The extent to which MSL is involved in learning motor procedures, is a question that needs further investigation for more elaborative answers, since memorisation of the same sequence of movements would offer little assistance during real laparoscopic surgery. While the learning process involved in acquiring a new skill definitely requires some kind of motor automation, it is more complicated than simply memorising a sequence of movements. Simulator tasks that promote MSL would not be successful in training surgeons, since learning how to perform laparoscopic surgery involves not only being able to acquire a demanding set of skills, but also to utilise it across many different settings (Mylopoulos & Woods, 2017). In fact, since patients are different, and their anatomy diverges, it is rarely, if not ever, the case that a surgeon would find oneself performing the exact same procedure, following the exact same steps. Surgeries on different patients would require, at least to some extent, the adaptation of one's technique to the demands of the situation, even if the type of surgery is the same (Moulton, Regehr, Lingard, Merritt, & MacRae, 2010). Understanding the extent to which MSL is involved in learning is important to determine the degree to which simulator tasks should vary, so that simulators would promote a holistic skill acquisition, with minimal focus on MSL. Before going into details on how our study aims to *explore the proportion of MSL to general motor learning*, it is important to discuss what entails learning, and explore the difference between holistic skill acquisition and MSL. Only then can our reader fully understand why practicing MSL would not be beneficial in laparoscopic training.

1.2. Exploring learning

Different authors use different terminologies for terms surrounding skill acquisition and learning, which are mostly abstract and therefore make it difficult to distinguish between them. We thus go on to explore such terms and provide an unambiguous distinction between them.

1.2.1. General motor Learning

According to Fitts and Posner (1967)'s theory, one of the most cited theories around learning, three main stages are involved in learning: namely a cognitive, an associative, and an autonomous stage. The cognitive stage includes identification and decoding of information at hand, the associative involves formation of links between tied aspects of information, and the autonomous entails task execution with minimal input of conscious processing. Within these stages, a set of internal cognitive, perceptual, motor, and perceptual-motor processes that are related to practice and experience, enable an individual to move through the three stages and achieve skilful performance (Schmidt, 1975). Schmidt's theory of motor learning states that formation of different representations in memory regarding the relations between the situation at hand, one's performance and sensations involved, as well as the results obtained, assist in the achievement of skill acquisition. While the above establish a fundamental and prominent picture of how motor learning occurs and what it entails, no answer is provided into what comprises a successful learning process from this description alone.

1.2.2. Holistic skill acquisition

We define successful learning using the term *holistic skill acquisition*, which entails acquiring a new skill that, not only enables successful task performance, but is also *transferable*; that is, it enables an individual to adapt one's performance according to the imminent situation (Mylopoulos & Regehr, 2011). As in the general process of motor learning, an individual has to pass through the three main stages, as those are defined in Fitts and Posner (1967), by making use of the cognitive, perceptual, motor, and perceptual-motor processes introduced by Schmidt (1975). Those help in the construction of representations in memory that are largely based on context. Thus, exposure to varying settings while practicing is critical to create additional, flexible, and adjustable representations.

The importance of variability in training is also prominent in other theories such as the Closed Loop Theory (Adams, 1976), which introduces a feedback loop as central to one's learning. More specifically, facing diverge situations, enable one's nerve system to learn how to handle discrepancies between actual and expected outcome from one's actions, and correct movements accordingly. Adam's theory (1976) can be depicted in applied settings of skill acquisition, such as sports performance. Close investigations of observations in performance have concluded that a set of plateaus, dips, and leaps are required for holistic skill acquisition. The above respectively refer to spurious limits (plateaus), development and implementation of new methods in one's technique (dips), and subsequent advances in performance (leaps). These finally lead to an individual's real performance limits; namely one's performance asymptote (Gray & Lindstedt, 2015). These 'milestones' keep the learner active and motivated through the exploration, development, and the implementation of new ways to perform tasks. Exploration and development of different dips in performance are particularly necessary, since they enable the individual to evolve, and become more flexible in discovery and implementation of new techniques, according to the demands of the situation. This active search for ideal methods and approaches is directly related to improvements in performance (Gray & Lindstedt, 2015).

In consideration of the above, Rasmussen (1983) affirmed that a training procedure can only be considered as holistically effective if it includes three behaviour types, referring to skill-, rule-. and knowledge-based behaviour. These behaviour types respectively refer to performing highly automated motor actions without conscious control (skill-based), executing a task based on stored rules or procedures (rule-based), and performing explicitly formulated actions based on general knowledge when faced with abstract or unknown situations (knowledge-based).

1.2.2. Motor sequence learning

The aforementioned exploration of holistic skill acquisition emphasizes the necessity for task variability during practice. We now introduce motor sequence learning (MSL), to make comprehensible why, while MSL can fall within the scope of 'general motor learning', it does not entail a holistic acquisition of skill. Rather, it is atypical of performance in varied situations such as those involved in laparoscopy, and should thus not be practiced in training. MSL refers to the acquisition of the skill to accurately produce a *specific* sequence of movements with as little effort and attentional monitoring as possible (Abrahamse, Ruitenberg, de Kleine, & Verwey, 2013). MSL is believed to involve a combination of cognitive and motor processing that optimizes performance by creating links between a small number of movements and merging them into subunits, called ‘motor chunks’ (Abrahamse et al., 2013; Verwey, 1996; Miller, 1956). Such chunks can even be achieved with the pure use of kinaesthetic feedback when simple motor sequences are involved, and are determinant of an individual’s performance (Pinzon, Vega, Sanchez, & Zheng, 2016).

More specifically, MSL functions within the framework of the dual processor model (DPM), which is comprised of the cognitive processor and the motor processor (Abrahamse et al., 2013; Verwey, 2001). During initial performance of a novel sequential task, information is processed in the cognitive processor, which translates every stimulus into a relevant response through the use of higher cognitive functions. The cognitive processor passes individual responses on to the motor buffer, which is integrated into working memory, and then triggers the motor processor to execute the relevant movements according to the information loaded in the motor buffer.

With repetitive practice, a *sequence skill* is formed, where the aforementioned ‘motor chunks’ are created, and loaded to the motor buffer as single processing steps, leading to improved, more rapid performance. Initially, the cognitive processor remains active in selecting which motor chunk is passed on to the motor buffer, but practice leads to automation of motor chunks involved in a sequence of movements, with the first motor chunk triggering the loading of subsequent chunks; thus engagement of the cognitive processor is minimized (Verwey et al., 2010, 2013). At this point, explicit recognition of the underlying knowledge and abstract rules governing one’s improved performance is less needed or even difficult to access consciously (Maxwell, Masters, & Williams, 2012).

At first glance, the process of holistic skill acquisition seems somewhat similar to MSL. A cognitive, associative, and autonomous is also involved in MSL, through the use of the cognitive processor (cognitive stage), the creation of motor-chunks (associative stage), and the eventual independent functioning of the motor buffer (autonomous stage). However, there are some great differences between the two processes. Even though an individual may acquire a new motor skill while repeatedly performing the exact same sequence of movements, one has only explored, developed and implemented a method that is ideal for this specific condition, and is thus very context-dependent. Studies have shown that performance of tasks involving MSL is affected by context-dependent information, with impairment of performance when new, conflicting information was introduced while executing the task (Ruitenberg, de Kleine, van der Lubbe, & Verwey, 2012). Moreover, repeating the same sequence of movements mainly achieves great automation of the spatial configuration involved in the execution of a specific task, which may not be challenging enough to provoke higher cognitive processes such as active searching and problem-solving. Through the repetition of the same sequence, individuals do not form diverse representations in memory required for holistic skill acquisition (Schmidt, 1975), but rather involve storing specific patterns of movement. Skill acquisition in MSL is thus rather contextualised, and non-transferable.

1.2.3. A framework for general motor learning

We now develop a model for the process of general motor learning to connect all aforementioned notions under a concise framework. The model comprises two components: *holistic skill acquisition* and *MSL*. The former entails strategies that are developed and employed for the execution of a task, and makes great use of the cognitive processor, while the latter depends on the spatial configuration that is acquired through the repetition of the same sequence, and works almost independently of the cognitive processor (Verwey et al., 2010, 2013). Depending on the extent to which a task includes repetition of the same sequence of movements, learning process has a relatively high or low proportion of MSL and holistic skill acquisition. For example, if a task involves the pure repetition of sequences, then MSL comprises most of the learning process with little practice of holistic acquisition of skills.

1.4. Aim of the study

1.4.1. Proportion of MSL to general motor learning

From the above, one can understand that if training task involves practicing MSL, then the learning process lacks a holistic acquisition of skills that would stimulate automation and successful transfer to real surgery. Thus, a task that promotes MSL could lead to improved performance in training, that would, however, not be representative of an individual’s actual capabilities under the varied and demanding conditions of real surgery. The exact role of MSL in task performance is not known, and therefore our aim is to *explore and estimate the proportion of MSL to the general motor learning process* involved in mastering a new task.

No experimental paradigm so far allows the computation of such a proportion while the learner is actively practicing a novel complex task. In this paper, we propose a paradigm for such an analysis, which we name *varied-fixed learning* paradigm[1] (VFL), and is based on the aforementioned notion that a task can promote either holistic skill acquisition or MSL, depending on its variability. The paradigm entails the manipulation of a task's requirement for either a varied or a repetitive set of movements for its execution, thus allowing us to capture the learning component that comprises each part: holistic skill acquisition and MSL. It thus enables the computation of MSL's proportion and unveils its role on task performance. The paradigm is modelled on individual learning curves for analysis. Details of the statistical model used are introduced in paragraph *2.4. Data analysis* of the Method section.

To test our paradigm, we use low-fidelity dexterity tasks that initially include a varied, and then a fixed part with respect to the set of movements required for execution. Dexterity tasks were chosen because they involve practicing complex fine motor procedures and thus provide a promising framework for exploring the proportion of MSL to general motor learning while an individual executes the complex task. Also, by assessing such tasks, one can capture the entire time it takes for that task to be mastered, which is vital since learning is a continuous time-invested procedure, and the learning process of laparoscopic surgery should be investigated as such.

1.4.2. Dexterity tasks as psychometric tools

Even though this is not our main goal, and as long as our study allows us to, we also attempt to provide some insights on the possibility that such tasks could be used as psychometric tools for the assessment of laparoscopic surgeons, since their nature allows the exploration of the

[1] The name varied-fixed learning paradigm derives from the paradigm's nature, which consists of two parts. The first part includes a set of trials, with each trial requiring a varied set of movements for execution, while the second includes a fixed set of trials, all requiring the same repetitive movements for execution.

continuous learning time, instead of focusing on a specific time period as done with cognitive abilities (Groenier et al., 2014). A decomposition of one's cognitive aptitude into sub-categories, as done so far when testing cognitive aptitude, may not capture the emerging abilities that couple their combined use. Since cognitive aptitude tests have not always been successful in predicting performance (Groenier et al., 2014; Huijser, 2015), dexterity tasks could be determined as a better psychometric tool.

2. Methods

2.1. Participants

A sample of 21 participants, with a mean age of 23.6, (SD=2.73, range 20 to 29) were recruited. Of them, 11 were male and 10 female. Handedness differed, with 18 being right handed and 3 left handed. All participants were university students of various degrees (Degree specialisation: 2 architecture, 4 engineering, 7 social sciences, 8 other).

2.2. Materials

Two tasks were devised for the current study, as described below. Both tasks were made to address fine motor skills, such as pinch prehension and finger flexion, as well as cognitive and visual-perceptual skills, such as alternating and sustained attention, spatial awareness and orientation, and hand-eye coordination. Laparoscopic challenges such as the fulcrum effect[2] were taken into account for the construction of the tasks.

2.2.1. Mirror-drawing task

An asymmetrical frame outline extracted from Arendt's (2017) 'mirror-drawing' task, (p. 21) was printed on an A4 piece of paper. The outline of each frame was black, and consisted of two lines with 5mm distance between them. Its mirrored outline was also printed on another A4

[2] The fulcrum effect refers to the direction of motion, which is depicted on the monitor through which a surgeon views the operating environment, as being opposite to the surgeon's horizontal and vertical hand movement (Gallagher, McClure, McGuigan, Ritchie, & Sheehy, 1998). For example, a left-side movement in a true operating environment is projected as a right-side one on the screen. This may cause dissonance between what the surgeon does and what he or she sees, and thus require counter-intuitive movements, making hand-eye coordination more challenging and complex.

piece of paper. One paper was placed on a tabletop, and reflected on a 30x30 cm mirror that arranged to stand right opposite the chair where the participant would sit. A white surface standing 10 cm higher than the paper was used to block the direct vertical view of the paper from the participant, while allowing the hand to draw on the paper, and the reflected image on the mirror to be visible to the participant. For an illustration see figure 2.

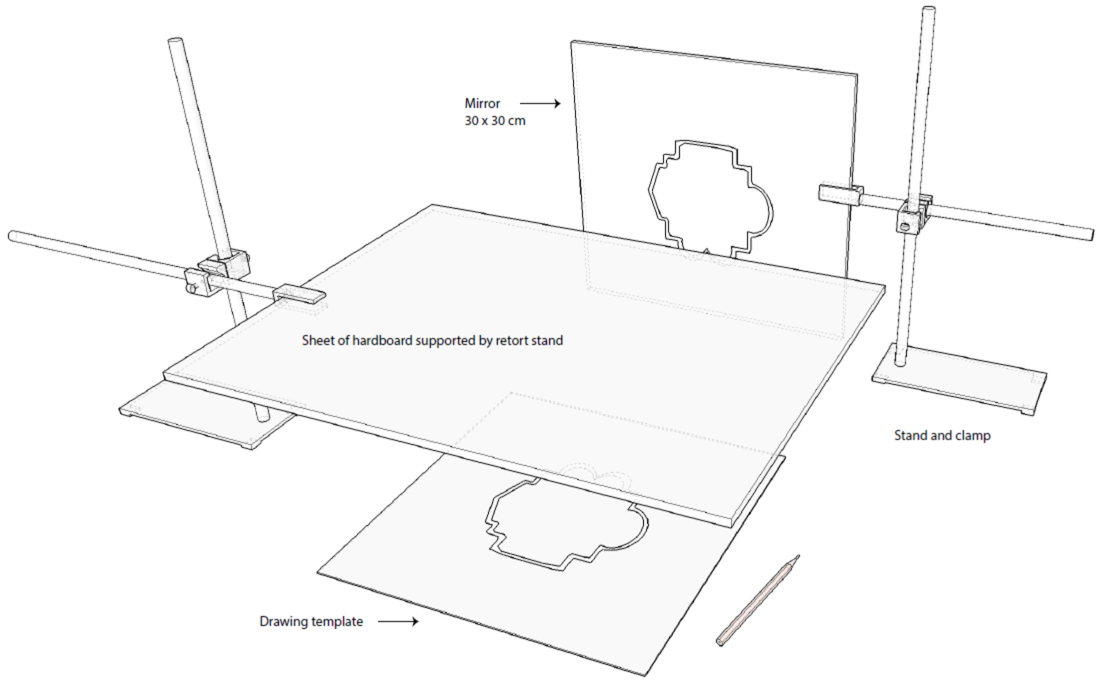


Figure 2. Illustration of the mirror-drawing task. A4 paper with printed frame outline on table. Sheet covering direct view. Mirror opposite participant, reflecting the outline.

2.2.2. Clips-and-string task

29 cable clips were nailed onto a cardboard base of dimensions 26x26cm. Their positioning formed two paths (of 19 clips each) with 7 changes of direction each, and a distance of 3 cm from one clip to its neighbour clips. The paths overlapped but mirrored the other, with one being outlined using a blue, and one using a red marker. 2 cubed boxes (dimensions 6x6x10cm and 7x7x7cm) were made to be placed on two positions on the cardboard. These boxes constituted obstacles that forced the participant to use different hand movements to reach each clip. A string of 50cm length was also provided. The tip of the string was burnt and glued to

create a strong tip that would be passed through the clips more easily. For an illustration see figure 3.

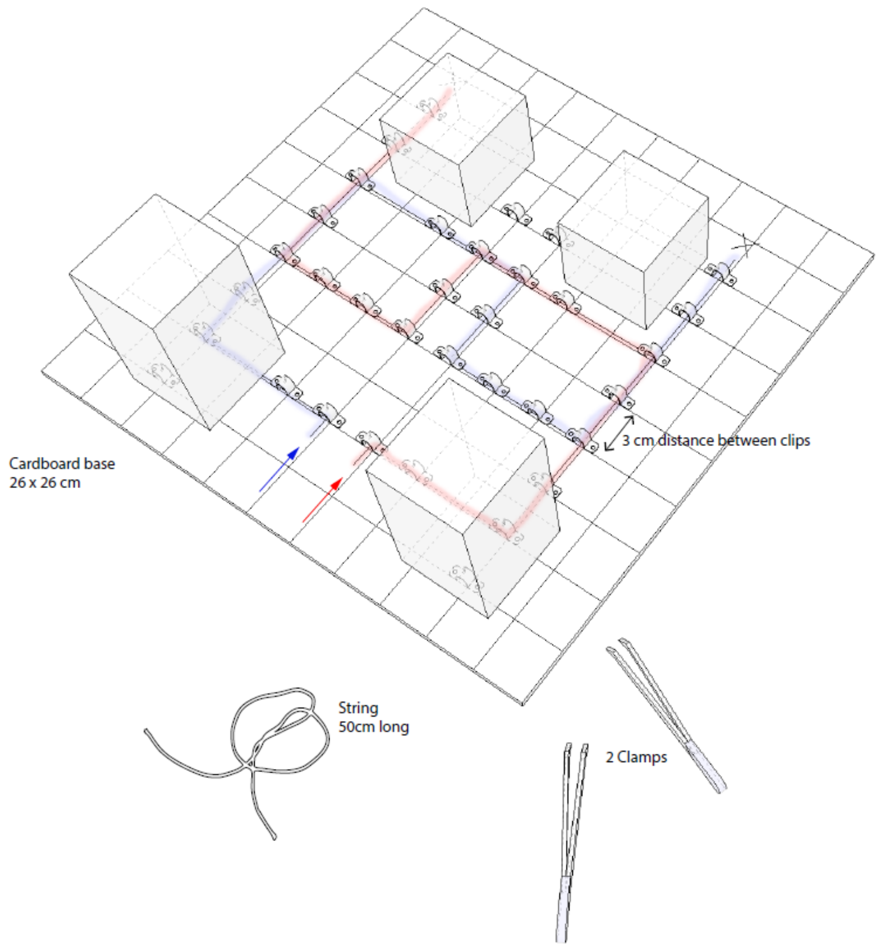


Figure 3. Clips-and-string task. Clips nailed on a cardboard paper, with two differently-coloured paths outlined. Boxes act as obstacles to respective paths. Tweezers and string provided next to apparatus.

2.2.4. Measures and other materials

An informed consent was created to be signed by the participants, stating the conditions of the study and the participant's rights (see Appendix A). Two different set of instructions were written, one for each task (see Appendix B). Also, an excel file was used as an observational form for the experimenter to note the results of the study, as well as any observations throughout, such as motivation level, questions asked, major errors noted, and disruptions. Other materials included a stopwatch.

2.3. Design

A repeated measures design was used for the study, in which all participants carried out all tasks. Each task was developed to accommodate two separate set of trials, in accordance to the VFL paradigm. More specifically, the first set included a variation in the sequence of movements the participant had to execute, which was achieved through the rotation of the apparatus after the completion of each trial. The varied movements prohibited MSL from playing a role in the learning process, and therefore the part was assumed to capture the process of *holistic skill acquisition*. The second part of the tasks included the repetition of the same sequence of movements, now entailing MSL. This part captured the learning process that contained both the skill mastered though holistic skill acquisition, and MSL. Thus, the subtraction of the holistic skill acquisition component from the learning process that entail MSL would produce the pure effect of MSL.

2.4. Procedure

One participant was brought in and sat in front of a table. The experimenter handed the consent form, and read out the instructions of first task that the participant was going to complete. Every participant had to undergo the two tasks the one after the other, but the sequence in which they were completed differed across participants, to ensure there were no order effects. Every task began after any questions were answered and the participant stated that he or she was ready to begin. The two tasks consisted of 26 trials, and each trial was timed (starting the moment the experimented said “go” and stopping as soon as the participant said “done”). The time was noted on the observational form by the experimenter after each trial, along with other observations. A reminder on the trials left was announced by the experimenter halfway through the task, and then again two trials before the final one, to maintain participant motivation.

2.3.1. Mirror-drawing

Using a pencil, the participant had to trace the outline provided, by drawing in-between its two lines, while looking at the mirror instead of looking directly at the paper. This was done to capture the fulcrum effect of laparoscopy, along with other cognitive and ergonomic demands.

The task consisted of two parts, the first one having 16 trials and the second 10. In the first part, the experimenter placed one frame outline on the table, which was rotated 45^{o} before the beginning of every new trial, until the frame completed a rotation of 360^{o}. Then its mirrored frame was presented and rotated again 45^{o} before the beginning of every new trial, until a 360^{o} was again completed. The sequence the frames were presented differed between participants to control for order effects. Once the first part was completed, the last frame rotation was kept on the table, and a vertical line was noted on one part of the frame. This line was used as the start line, and the participant had to continuously trace the same frame for 10 more trials to complete the second part. The experimenter had to pay attention on when the participant passed the start line, to split the seconds in the stopwatch accordingly.

2.3.2. Clips-and-string

In this task, which was also composed of two parts, the participant was instructed to pass the string through the clips using two tweezers (one in each hand). In the first part, the participant had follow one path till the end. After the completion of each trial, the experimenter removed the string from the clips, and rotated the cardboard 45^{o}. The positioning of the boxes was changed to their respective mirrored positions, and the participant was instructed to follow the other path. The same procedure followed, with the experimenter rotating the cardboard 45^{o}, until 16 trials were completed. For the second part, the cardboard was kept in the same position, and the participant had to pass the string through the clips for another 10 trials, following the same path every time.

2.4. Data analysis

A non-linear mixed effects model with an exponential two-part learning curve as a likelihood function was used to run the regression analysis and estimate each individual learning curve. The exponential model of learning, which we called ARARY[3], formed a two-part learning curve that models the two experimental phases of the tasks: *holistic skill acquisition*, and *general motor learning* (for illustration, see figure 4). The first part of the learning curve models

[3] The term ARARY derives from the initials of all five parameters captured in the model, those being *amplitude* (A), *rate* (R), and *asymptote* (Y). Deconstruction of the model, and explanation of each parameter are provided in paragraph *2.4.1. Statistical model for analysis*

Time-on-Task (ToT) when the task's trials varied, while the second part models the one where the trials were repetitive, thus initially capturing the holistic skill acquisition process isolated from MSL, and the combined processes later on. The random component was captured using the Gamma distribution, since the distribution of learning curves is left-skewed and variance decreases as approaching the asymptote, thus the use of a normal distribution would produce a strong bias.

2.4.1. Statistical Model and analysis

Our model considered three aspects of learning: amount of learning within the study (amplitude, δ), learning speed (rate, ρ), and maximum learning capacity (asymptote, ω). These aspects make up a exponential function for learning curves that is preferred to power functions frequently used in the past (Heathcote, Brown, & Mewhort, 2000). Since it was a two-part model, the overall structure consisted of five parameters based on these learning aspects, that determined each individual two-part learning curve: δ_S, δ_M, ρ_S, ρ_M, ω. The number of trial repetition per component t_S and t_M is also modeled. Thus, the performance over trials (ToT) is represented in the following formula:

$$\text{performance} = \omega + \delta_S e^{-\rho_S t_S} + \delta_M e^{-\rho_M t_M}$$

Regression analysis was performed using the package `brsm 2.1` (Brunckner, 2017), and the non-linear functions were built using the dedicated library from the package `asymptote` (Schmettow, 2017). Population-level effects were estimated for analysis.

Our analysis focused on the two amplitudes within each task, assuming that any difference in the ratio δ_S to δ_M would reflect the pure effect of MSL. Population-level parameters were drawn from the posterior distribution of the model, and prediction intervals were estimated. Then we created an ARARY model whose parameters were linearized, running on a log-scale ranging from $-\infty$ to $+\infty$, in order to successfully capture random effects, revealing the variance caused due to individual differences. Those were extracted from the random component of the posterior using the `bayr 0.8.5` package (Schmettow, 2017), thus enabling the investigation of participant-level variation. Intercept random effects were used to capture the general variation in MSL proportion between participants, while the coefficients of slope random effects were used to assess the generalisability of this proportion. Strong dispersion in effects would reveal stronger variation in the proportion of each individual's MSL, making the average MSL proportion less generalizable.

Then, pairwise correlations between task learning curves, along with their estimated 95% credibility intervals were calculated, to examine the internal consistency of such tasks, and explore the possibility of their use as psychometric tests.

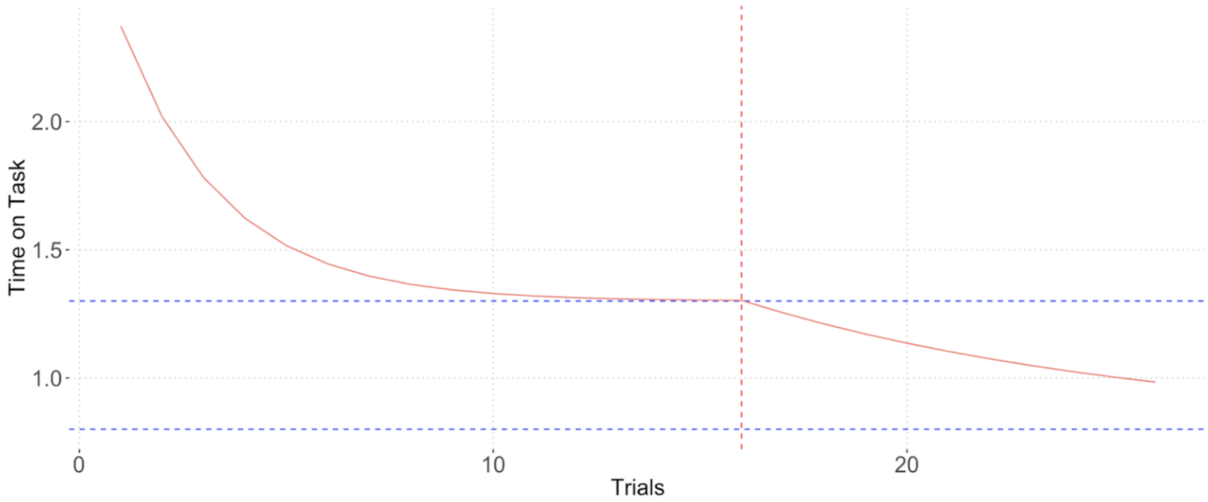


Figure 4. Exponential two-part learning curve model (ARARY). Trials and mapped on the x-axis and ToT is mapped on the y-axis. The two-part learning curve models holistic skill acquisition on the first part, and general motor learning on the second, considering amplitude δ_S and δ_M, rate ρ_S and ρ_M, and asymptote ω.

3. Results

The results of the analysis performed are presented below, in three sub-sections. The first one focuses on a *within-task analysis,* presenting the population-level effects for the learning parameters of each task, while the second focuses on an *across-task analysis,* where pairwise correlations between the learning curves of each participant for the two tasks were computed. In the final section we criticise our model to determine the robustness of the results.

3.1. Population-level effects within tasks

Population-level effects for the five parameters were drawn from the posterior distribution. All four parameters and their 95% credibility limits are reported in Table 1. The exact proportion of MSL to general motor learning, even though not a parameter in itself on our model, was created on the posterior and also presented on the table, at the respective row (labelled MSL) for each task. As seen on the table, the proportion of MSL in the mirror-drawing task was almost twice as large (0.10; *95% CI* [0.04; 0.22]) as the proportion in the clips-and-string task (0.05; 95% CI [0.00; 0.55]), but uncertainty was high in both. While there is large uncertainty, it is fair to suggest that MSL comprises a small proportion of the general motor learning process.

This is also evident by inspecting the δ_{M} to δ_{S} ratio from the amplitude values of each task, as presented at the respective rows in table 1. In the mirror-drawing task, the amplitude in the first part of the learning curve is approximately eight times larger (δ_{S} = 1.81; *95% CI* [1.35; 2.39]) than this of the second part (δ_{M} = 0.21; *95% CI* [0.07; 0.46]), with both values being highly uncertain. For the clips-and-string task, the difference between the two amplitudes is more than twice as large as that of the mirror-drawing, thus yielding a smaller δ_{M} to δ_{S} ratio. Specifically, the amplitude for the holistic skill acquisition part is almost fourteen times larger (δ_{S} = 0.95; *95% CI* [0.72; 1.21]) than the general motor learning amplitude (δ_{M} = 0.95; *95% CI* [0.72; 1.21]). Large uncertainty seems to exist, with credibility limits deviating more than 0.2.

Table 1. Population-level effects for amplitudes and rates per task

Task	Parameter	centre	lower	upper
Mirror-drawing	MSL	0.10	0.04	0.22
	δ_{S}	1.81	1.35	2.39
	δ_{M}	0.21	0.07	0.46
	ρ_{S}	0.47	0.35	0.62
	ρ_{M}	0.06	0.02	0.17
	ω	0.20	0	0.32
clips-and-string	MSL	0.05	0	0.55
	δ_{S}	0.95	0.72	1.21
	δ_{M}	0.05	0	1.17
	ρ_{S}	0.24	0.15	0.42
	ρ_{M}	0.34	0	17.37
	ω	1.09	0	1.23

Estimates with 95% credibility limits

While the proportion of MSL is based on the two amplitudes within each task, some observations on the rates are also worth noting. More specifically, and especially for the mirror-drawing task, the rate of learning seems to be faster in the first part of the learning curve (ρ_S = 0.47; *95% CI* [0.35; 0.62]), than the second (ρ_M = 0.06; *95% CI* [0.02; 0.17]), with notable, but lower uncertainty than the one observed in amplitudes. This suggests that more trials were needed in the second part to reach maximum performance, since improvement in learning was rather slow compared to the first part.

3.2. Participant-level effects within tasks

3.2.1. Individual learning curves

All learning curves generated from both tasks are presented in Figures 3, with the x-axis displaying the trials, while the y-axis displaying the time-on-task (ToT). Overall, the shape of each learning curve seems to differ, stressing the importance to examine individual learning curves rather than an average computed from all of them. A visual inspection of the learning curves for the two tasks follows.

In the mirror-drawing task, not much dissemination of data is observed on an individual level. For the clips-and-string task, there is higher dissemination of data points within each participant learning curve. This might suggest a flaw in the task computed, possibly causing random errors in execution. It should also be noted that there is not visible stabilization to the ToT on the second part of the curves, suggesting that the asymptote was not reached. This is an important observation, since the amplitude of each learning curve is largely dependent on its asymptote. Hence, if the asymptote is not reached, the estimation of the amplitude is rather uncertain. Due to the little repetition of trials, it is likely the case that the true asymptote of each learning curve was not reached for every individual, which suggests more trials are required to reach maximum performance. This suggestion is also supported by the differences in the two rates noted on the population-level, reported on table 1.

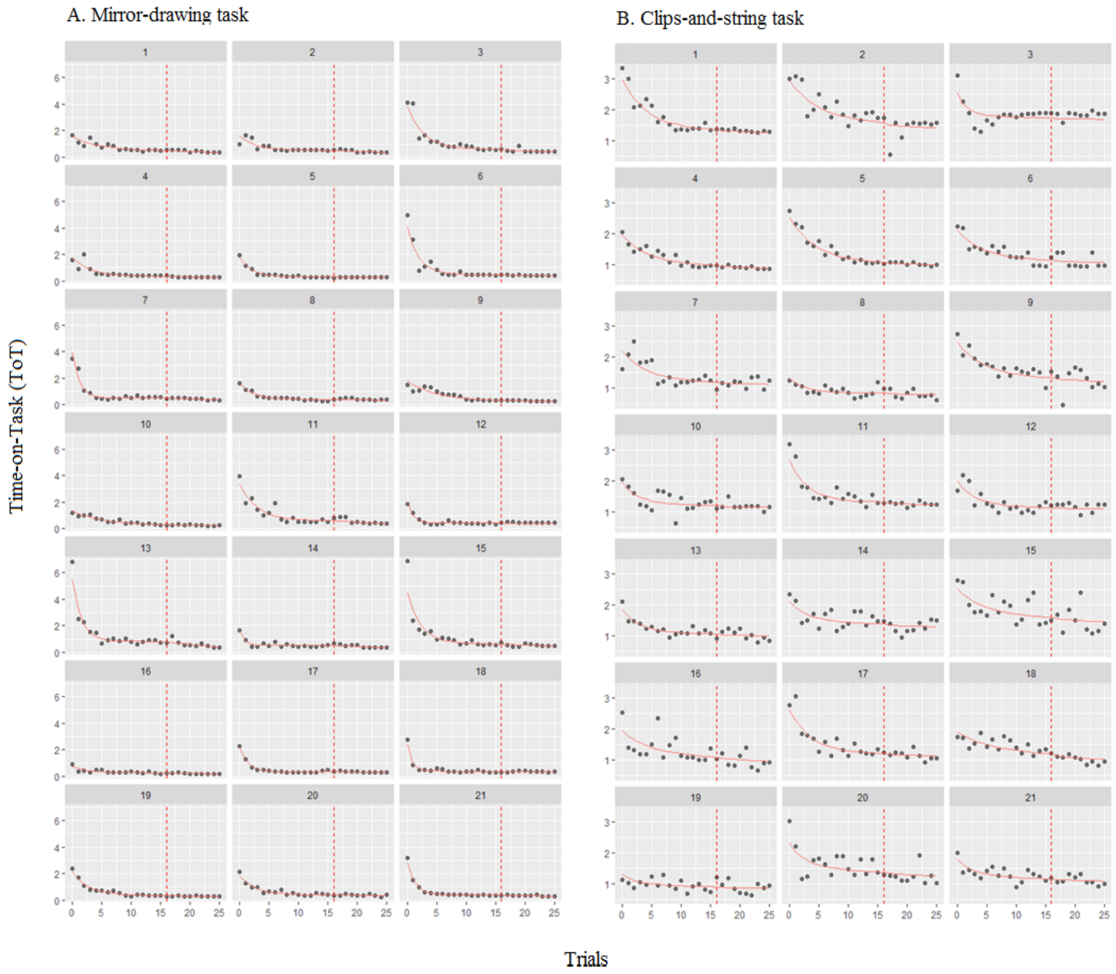


Figure 3. Participant learning Curves (Time-on-Task as a function to task trials) for mirror-drawing (A) and clips-and-string task (B).

The difference in rate between the two tasks on a population-level analysis presented in table 1, is also evident through the visual comparison of the individual curves for the two tasks in figure 3. One can see that the learning rate seems to be lower for the mirror-drawing task, reflecting a more gradual learning process for the clips-and-string task. This difference in rate is in line with the respective calculated population-level effect for rate (see table 1). The difference between the first and second component of the model is minimally, to not-at-all visible in the individual learning curves.

3.2.2. Individual MSL proportions

Having discussed the population mean of the MSL proportion in paragraph 3.1., we go on to investigate whether this mean was representative of all individuals. To capture random effects, an ARARY model was created, in which parameters were linearized to run on the same log-scale ranging from -∞ to +∞, and random effects were extracted from the posterior. A caterpillar plot devised using `ggplot2` package, plotted participant random effects on the x-axis, ordered according to the magnitude of the effect, depicted on the y-axis. The credibility limits were displayed using lines for every effect, with larger lines reflecting higher uncertainty on the MSL proportion within each individual learning curve. As can be seen, there is large variation in the magnitude of each individual's MSL proportion with regards to the population mean, which is also highly uncertain, especially in the mirror drawing task,. This variation from the population mean brings about questionable conclusions concerning the generalizability of MSL proportion. Also, the high uncertainty of the random effects highlights the need for more exhaustive research, since current results are too uncertain to suggest final conclusions.

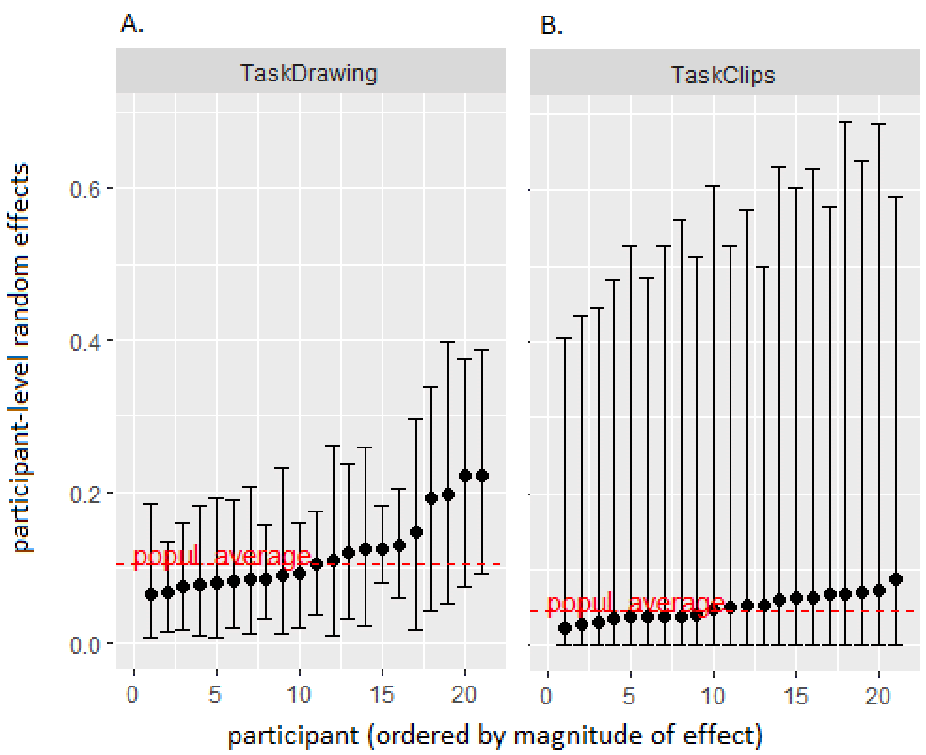


Figure 4. Participant-level random effects for MSL proportion of all participants in mirror-drawing (A) and clips-and-string task (B), plotted according to magnitude of effect, from lowest to highest effect size. Reference line (popul_avegage) is the population mean MSL for each task.

3.2. Correlation of learning parameters across tasks

An across-task analysis was computed using pairwise correlations to estimate population-level effects for the two amplitudes (δ_S, δ_M) and rates (ρ_S, ρ_M). Results are reported in Table 2. A weak positive correlation was found between the asymptotes of the two tasks (r = 0.32), but even this has high uncertainty (*95% CI* [-0.84; 0.97]). This correlation can only suggest the possibility that dexterity tasks could be used as psychometric tests, but no further conclusion can be drawn.

Table 2. Pairwise correlations of population-level parameters

Task		Clips-and-string
Mirror-drawing	δ_S	-0.03 [-0.53; 0.49]
	δ_M	0.13 [-0.93; 0.91]
	ρ_S	-0.06 [-0.76; 0.69]
	ρ_M	-0.02 [-0.95; 0.95]
	ω	0.32 [-0.84; 0.97]

Estimates with 95% credibility limits

3.3. Model criticism

After checking the prediction intervals (25%, 50% and 75% quartiles) to ensure the certainty of the posterior prediction, it seems that most observations fall within the predicted range. However, it is worth mentioning that there are certain outliers observed, particularly with regards to those falling below our predicted range, that could reflect the effects of difficulty in task variation. Residual plots show large variation, which is expected since the learning curve of each participant had a different shape.

The prediction intervals and residual plots can be found in appendix C, for further examination.

4. Discussion

The aim of the current paper was to investigate the extent to which MSL is involved in the general motor learning process of dexterity tasks. For the investigation, an innovative paradigm was employed, named *varied-fixed learning (VFL)* paradigm, and modelled using a novel approach to the analysis of learning curves. Findings suggested that the proportion of MSL to general motor learning is present but not very pronounced, while estimates were highly uncertain. The presence of MSL, even though small, implies that variability in laparoscopic training tasks is essential, since promotion of MSL through repetitive tasks might lead to better performance that is however not representative of transferable skills. Variability will ensure that skills mastered in training are transferable to real settings. However, the large uncertainty in the proportion of MSL calls for further, more exhaustive investigation to attain more certain results and derive to conclusive arguments with regards to the extent of variation needed. After an initial scan through our findings, our discussion focuses on the limitations of the current study, in an attempt to investigate the possible reasons behind the uncertain findings. Based on these limitations, we proceed to the provision of guidelines for adjustments that need to take place so that future research can move towards an optimised way of exploring MSL and its role when mastering a new complex procedure. Finally, we discuss the feasibility and implications of the proposed paradigm and approach to analysis in the field of applied research.

4.1. Exploration of current findings

The component of MSL was rather low, since its addition to holistic skill acquisition offered minimal extra advance in performance. However, since this research constituted only a pilot exploratory investigation regarding the feasibility of this novel way to analysis and, as mentioned, findings were uncertain, it would be naive to make decisive conclusions on the exact proportion of MSL from this research alone. Before the examination of experimental weaknesses that might have caused uncertainty in the results, we first focus on some observations of the current findings that are rather noteworthy and could inform future research.

4.1.1. Difference in MSL proportion across tasks

An interesting remark is the difference in MSL proportion between the two tasks. More specifically, when comparing MSL across tasks, one can see that its proportion, while small in both, was almost twice as large on the mirror-drawing compared to the clips-and-string task. This implies dependency of MSL on the nature of the task, either in terms of complexity, or cognitive versus motor effort required for task execution. For example, the mirror-drawing task required much more perceptual effort, due to the fulcrum effect created from the reflection of the participant’s movements on the mirror. However, other than the advanced need for effortful hand-eye coordination due to the perceived inversion of movement, motor requirements on that task were not as demanding, since participants knew how to operate a pencil. On the other hand, the clips-and-string task required much more motor effort, since participants had to learn how to handle the tweezers and carefully pass the string through small clips. Cognitive effort was not as demanding as in the mirror-drawing task, since no recreation of the environment was needed.

Trying to further inspect the dependence of MSL on task nature, a more elaborate glance on majority of paradigms used in previous MSL studies provide further insights. For example, types of apparatus used in such studies, such as the *discrete sequence production* task (DSP; Abrahamse et al., 2013), the *m x n* task (Hikosaka, Rand, Miyachi, & Miyashita, 1995) or the *serial reaction time* task (SRT; Nissen & Bullemer, 1978), include small repetitive sequences of movements, that are timed closely to each other and do not diverge in movement required for execution. Moreover, authors like Sakai, Hikosaka, and Nakamura (2004), have stressed that not only the serial order of a sequence, but also the timing of movement, referred to as rhythm, is determinant of a sequential performance, and this rhythm reflects the underlying chunking process.

Bridging these remarks to our findings, the mirror drawing task favoured continuity much more that the clips-and-string task, since in the latter, movement was occasionally interrupted by confounding obstacles. For example, it was a frequent observation throughout the experiment that the string tangled to clips, and participants had to stop to untangle the string and then continue carrying out the task. This might have impeded the potential for motor chunking of sequences, since conflicting stimuli interrupted the sequence at varying points, preventing the sequence to acquire a patterned serial order or timing.

This suggests that the context-dependency of MSL largely lies on the continuous nature with regards to the flow of movements required for their execution, that should also be in close temporal proximity to promote sequential chunking. Such an observation is important in informing future simulator tasks, in terms of how continuous and complex they should be.

4.1.2. Difference in MSL proportion across participants

Another insight is provided from differences in MSL proportions across participants. Examination of random effects suggested that there is dispersion of this proportion across individuals, meaning that the population mean of MSL is not as generalizable to everyone, with many individuals diverging from the mean. Specifically, while MSL is promoted via the repetition of the same movements, the extent to which individuals acquire learning of specific motor sequences it is not only task-dependent, but also depends on individual differences.

This assertion warns researchers to proceed with caution in future simulator tasks design, since there is a chance of biased performance depending on the variability of individuals' competence in acquiring MSL. For example, if a laparoscopic task were to be highly repetitive, then a biased increase in performance would be present in individuals with high MSL, that would, however, not reflect their true potential for general motor learning. In other words, individuals with high proportion of MSL would show better performance in repetitive simulator tasks than under realistic, diverse conditions. The chance of such an optimistic bias in learning during training needs to be minimized.

Our estimations of random effects were highly uncertain, but the important insights they provide deem necessary a more exhaustive investigation of random effects in future replications of the study, to get more certain estimates. The root of the current uncertainty, and necessary steps for acquisition of more certain results are explored in paragraph *4.2. Study limitations and amendments for future replications* of the discussion section. Given that future replications yield certain random effects that still reveal a great variation in participant-level MSL proportion, researchers could dive into exploring the possible, more theoretically-based explanations to such a variation.

4.1.3. Task correlation

Concerning findings on the psychometric use of dexterity tasks, weak links for internal consistency were found in participants' maximum performance, which were also highly uncertain, suggesting poor reliability between low-fidelity dexterity tasks. This low correlation between the two tasks could be attributed to the earlier discussed differences in task-nature and cognitive-motor demands between the two tasks. More research is required for more certain findings, though their potential for psychometric testing is increasingly doubted, since past research has also found highly uncertain weak links (Arendt, 2017; Kaschub, 2016).

4.2. Study limitations and amendments for future replications

Having examined our findings, we now inspect in more detail the experimental paradigm used and the flaws of our pilot study, which possibly account for the high uncertainty in findings. Weaknesses discussed form the basis of our guidelines for better study replications in future research.

The thorough investigation of individual learning curves reported in the results section, showed that the rate of learning from the varied part of our experimental paradigm was almost seven times as fast as that of the fixed part on the mirror-drawing task. This implies that MSL could possibly bring further advance in performance than those noted in our study, but this was not detected within the limited number of trials since maximum performance was possibly not reached. Also, some dissemination on the time taken to complete each trial was noted, especially in the clips-and-string task. One can conclude, then, that the parameters of each learning curve are rather uncertain. For example, inability to reach maximum performance (parameter asymptote) could have led to uncertainty in the amplitude and rate parameters as well, since those are highly interconnected within our model. Large dissemination could also lead to less accurate rate estimation, which would in turn affect the amplitude parameter's certainty. It is thus important that data collection enables a recording of time-on-task that is as accurate as possible. Below are some key alterations that are believed to bring about improvements to the results' robustness.

4.2.1. Increased number of trials

To estimate the MSL proportion with sufficient certainty, it is important that researchers in future replications record a longer series of fixed trials in both experimental conditions. Increased number of trials will ensure a stable asymptote to one's learning curve, that in turn affects the certainty of amplitude and rate parameters, and will therefore yield greater certainty in the results. Considering that the rate of the second part of the mirror-drawing task was almost seven times slower than that of the first part, one would need to include approximately seven times as much trials in the second part of the task. However, such an excessive number of trials would most likely generate participant fatigue, depriving them of further advance in performance. Given that the rate estimates are uncertain, and thus the difference between the two rates could be smaller, we suggest a number of trials that would lay to approximately 20 trials for the first and 30 trials for the second part, to prohibit fatigue and ensure some stability in the asymptote. In case researchers wish to include more trials, introduction of spaced-practice is a possibility, which has also been correlated to advanced motor skill acquisition (Kwon, Kwon, & Lee, 2015; Schreuder, Wolswijk, Zweemer, Schijven, & Verheijen 2011). Researchers could also consider including a second set of fixed trials in each task. In this way, one can also derive correlations between the rate patterns of learning within the learning curve of each participant, and get more robust results with regards to this learning parameter.

4.2.2. Configuration difficulty as a random effect

Researchers should also pay great attention to the configurations of the dexterity tasks employed. In the mirror-drawing task, noise in data was minimal, reflecting a similar difficulty level across varied trials. However, there was high noise in the data generated from the clips-and-string task, which might suggest that some configurations of the apparatus were more challenging, while other were easier to execute. This is also supported by certain data outliers in some learning curves with regards to the Time on Task (ToT), where some trials of highly increased performance lied in-between normal-performance trials. An observation attributed to the fact that certain configurations favoured performance. Therefore, experimenters in future replications should have a defined set of configurations with varied difficulty. Having them identified in the data set would allow configurations to be considered a *non-human population* that could also be extracted as a random effect from the random component of the posterior analysis, and provide one with more robust results, by factoring out variance due to difficulty in configuration.

4.2.3. Continuous flow of movement, and smaller movement sequences in tasks

As discussed earlier, the clips-and-string task did not favour continuity, since obstacles frequency interrupted movement. Knowing that movement rhythm, formed by the repetitive serial order and tied timing of moves (Sakai, Hikosaka, & Nakamura, 2004), promotes MSL, future replications could create an alternative to this task. Promotion of MSL would be more certain in a task that enables the uninterrupted execution of motor sequences. For example, instead of passing a string through clips, nails with heads could be used, to allow participants to pass the string *by* nails instead of *through* them. Also, a smaller, less scrambled path would ensure the non-tangling on the string.

4.2.4. Smaller movement sequences

Related to the above, tasks used should also ensure that movement sequences are small enough to allow chunking. Since chunking is directly related to working memory (Seidler, Bo, & Anguera, 2012), which is known for its limited capacity ranging somewhere between five to seven items (Miller, 1956; Cowan, 2001), movement sequences should not be very demanding. The study of Pinzon et al (2016), whose apparatus closely resemblance our tasks in term of the flow of movement required, found that continuous sequences larger than seven turns in reactions (thus with more than seven items to be processed in working memory), were not as successful in following replications. A future apparatus could form a smaller path, or a path whose rhythm is clearly divided in parts.

4.3. Model feasibility and implications

Due to the exploratory nature of the study, findings might not have attained a more concrete answer on the exact proportion of MSL to the general motor learning process, they did however pave the way towards a novel approach to the investigation of learning processes, such as MSL, under a more naturalistic experimental paradigm. Unlike the research that has been conducted so far on MSL, which has been focusing on standardized key-pressing tasks that involve minimal exposure to novel skills (e.g. Abrahamse et al., 2013; Hikosaka et al., 1995; Nissen & Bullemer, 1978), our *varied-fixed learning* (VFL) paradigm offers researchers the opportunity to investigate MSL throughout the observation of an individual's entire learning process of mastering a complex motor procedure. As seen in the resemblance spectrum, this holistic investigation of learning offers the opportunity to conduct research that moves beyond the laboratory, into applied settings, and generates more representative outcomes than other, simplified test suites.

While the study conducted here focused on a within-subject analysis using a repeated measures design, our paradigm and model of analysis enables comparisons across-subjects as well, that can be used for research in other applied disciplines. For example, our model could be used to investigate age differences in MSL proportion. The increasing number of elderly people have turned recent research towards the investigations of elderly's ability to acquire new motor skills that are necessary for their independent living. Studies have already found differences in chunking between younger and older adults (e.g. Barnhoorn, van Asseldonk, & Verwey, 2017), yet those are still conducted using simplified laboratory conditions like the DSP task. Having identified the advantages of tasks that are closer to real life activities, researchers could use the VFP paradigm to compare aging differences in MSL proportion, while executing a dexterity task like tying one's shoelaces. The comparison could be performed in the age group of young adults (20 to 30 years old) where cognitive abilities peak, and older adults, ranging from 60 to 70 years of age, since 65 is considered the mean age for cognitive abilities decline, especially in situations where cognition interacts with motor control (see Li & Lindenberger, 2002 for a review).

We also encourage researchers to follow our proposed analysis of learning curves, since it is an exponential model that does not depend on averages across individuals. Other authors have already discussed the limitations of learning curves using power functions (Heathcote, et al., 2000), mainly concerning the great variability that exists across the learning parameters of each individual, favouring against conclusions based on averages. Our exponential model enables one to investigate learning curves at a population level but also at an individual one, offering greater insights into everyone's learning process, and enabling researchers to make suggestions for improved designs that also consider individual differences. For example, individual learning curves could inform personalised training procedures, which have been suggested as a time-efficient and cost-effective approach to training (Ahmed, Abid, & Bhatti, 2016), by assessing individual trainee needs and enabling an individual-tailored training process.

The ARARY model should also be preferred over other statistical models for investigation of learning curves, since it can be adjusted to include more parameters, and be used in other research around learning, besides MSL. For example, it can be expanded to a three-part ARARARY learning curve, to assess the different behaviours of learning as defined by Rasmussen (1983). A proposition for future research could be to assess participants during performance on a complex simulator task, and investigate individual differences that would offer more insights on the need for proper surgeon assessment. Such a task would need to include three different phases. The first one should require referral to instructions for the participant to carry out the task, including enough trials to automate the action and stop referring to instructions. This part would capture rule-based behaviour. The task could later go on to be repeated, capturing the skill-based behaviour, until an unexpected situation would be presented. The participant would then need to use abstract knowledge to complete the task, exercising knowledge-based behaviour. The learning parameters amplitude, rate, and asymptote would be captured in every part of the simulator task, for assessment of individual differences in learning curves using our model. Large individual difference could imply the necessity for good surgeon assessment or more adaptive training.

As can be understood, the analysis of learning proposed in the present paper offers a promising alternative to research in applied settings. The proposed amendments should be viewed as guidelines, so that future replications could move towards more certain results, that could be generalised to laparoscopic and other fields.

5. Conclusions

Findings and observations of the current research suggest that MSL plays a small, yet important role in learning, since improvements in performance are noted after the promotion of specific movement sequence automation. MSL seems to be rather contextualised, meaning that certain aspects regarding movement order and rhythm need to be present for a task to promote sequence learning. Therefore, experts should proceed with caution in developing sufficiently varied simulator tasks, that can ensure a holistic skill acquisition. Some remarks of the current research also suggest the possibility for individual differences with regards to one's ability in acquiring MSL, which however needs further investigation. It would be hasty to make more explicit suggestions for laparoscopic task variation or individual differences based solely on this research, yet the guidelines suggested in the discussion could generate more precise estimates.

The present research has made some important breakthroughs in research around learning. The introduction of a different approach to investigation of learning poses a development that could revolutionise research in applied domains. Specifically, implementation of the VFL paradigm enabled research around motor learning to be conducted in more naturalistic settings, while individuals practice complex motor procedures, instead of limiting investigations to simplified laboratory tasks. Further, our research supported the feasibility of investigating learning curves using the ARARY model, which shows great potential for future assessment procedures regarding learning parameters involved in simulator training. It also allows the consideration of individual variations in the parameters comprising learning, thus enabling more thorough investigation of individual differences. Finally, the introduced concise conceptual framework of how MSL can be involved in general motor learning depending on task nature, and how it differs from holistic skill acquisition, offers sound groundwork for future developments in learning research.

References

Abrahamse, E. L., Ruitenberg, M. F. L., de Kleine, E., & Verwey, W. B. (2013). Control of automated behaviour: Insights from the discrete sequence production task. *Frontiers in Human Neuroscience, 7,* 1–16. doi:10.3389/fnhum.2013.00082

Adams, J. A. (1976). Issues for a Closed-Loop Theory of Motor Learning. *Motor Control*, 87-107. doi:10.1016/b978-0-12-665950-4.50009-2

Ahmed, A., Abid, M. A., & Bhatti, N. (2016). Balancing standardized testing with personalized training in surgery. *Advances in Medical Education and Practice, 8,* 25-29. doi:10.2147/amep.s122312

Arendt (2017). *Towards reliable and valid prediction of MIS-performance with basic laparoscopic tasks in the LapSim and low-fi dexterity tasks.* (Unpublished master's thesis). University of Twente, Enschede, Netherlands.

Barnhoorn, J. S., van Asseldonk, E. H. F., & Verwey, W. B. (2017). Differences in chunking behaviour between young and older adults diminish with extended practice. *Psychological research = Psychologische Forschung,* 1-11. doi: 10.1007/s00426-017-0963-6

Cook, D. A., Hamstra, S. J., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., . . . Hatala, R. (2012). Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Medical Teacher, 35*(1). doi:10.3109/0142159x.2012.714886

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87-114. doi:10.1017/s0140525x01003922

Fitts, P. M., & Posner, M. I. (1967). Learning and skilled performance. *In Human performance.* Belmont, California: Brooks-Cole.

Gallagher, A. G., Mcclure, N., Mcguigan, J., Ritchie, K., & Sheehy, N. P. (1998). An Ergonomic Analysis of the Fulcrum Effect in the Acquisition of Endoscopic Skills. *Endoscopy, 30*(7), 617-620. doi:10.1055/s-2007-1001366

Gallagher, A. G., & Smith, C. D. (2003). Human-Factors Lessons Learned from the Minimally Invasive Surgery Revolution. *Surgical Innovation, 10*(3), 127-139. doi:10.1177/107155170301000306

Grantcharov, T. P., & Funch-Jensen, P. (2009). Can everyone achieve proficiency with the laparoscopic technique? Learning curve patterns in technical skills acquisition.

American Journal of Surgery, 197(4), 447–9. doi:10.1016/j.amjsurg.2008.01.024

Gray, W. D., & Lindstedt, J. K. (2017). Plateaus, dips, and leaps: Where to look for inventions and discoveries during skilled performance. *Cognitive Science, 41*(7), 1838-1870. Doi: 10.1111/cogs.12412

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: Theory and Practice, 19*(2), 203–17. doi:10.1007/s10459-013-9455-7

Heathcote, A., Brown, S., & Mewhort, D. J. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*(2), 185-207. doi:10.3758/bf03212979

Hikosaka, O., Rand, M. K., Miyachi, S., & Miyashita, K. (1995). Learning of sequential movements in the monkey: process of learning and retention of memory. *Journal of Neurophysiology, 74*(4), 1652-1661. doi:10.1152/jn.1995.74.4.1652

Huijser, S. (2015). *Influence of cognitive abilities on laparoscopic simulator trainer performance* (Unpublished master's thesis). University of Twente, Enschede, Netherlands.

Kaschub, V. L. (2016). *Learning complex motor procedures* (Unpublished bachelor’s thesis). University of Twente, Enschede, Netherlands.

Kwon, Y. H., Kwon, J. W., & Lee, M. H. (2015). Effectiveness of motor sequential learning according to practice schedules in healthy adults; distributed practice versus massed practice. *Journal of Physical Therapy Science,27*(3), 769-772. doi:10.1589/jpts.27.769

Li, K. Z., & Lindenberger, U. (2002). Relations between aging sensory/sensorimotor and cognitive functions. *Neuroscience & Biobehavioral Reviews, 26*(7), 777-783. doi:10.1016/s0149-7634(02)00073-8

Maxwell, J. P., Masters, R. S., & Williams, A. M. (2012). Implicit motor learning, reinvestment and movement disruption: what you don’t know won’t hurt you? In N. J. Hodges (Ed.), Skill acquisition in sport: research, theory and practice. London: Routledge.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 101*(2), 343-352. doi:10.1037//0033-295x.101.2.343

Moulton, C. A., Regehr, G., Lingard, L., Merritt, C., & MacRae, H. (2010). Slowing down to stay out of trouble in the operating room: remaining attentive in automaticity. *Academic Medicine, 8*(10), 1571-1577. doi:10.1097/ACM.0b013e3181f073dd

Mylopoulos, M., & Regehr, G. (2011). Putting the expert together again. *Medical Education, 45*(9), 920-926. doi:10.1111/j.1365-2923.2011.04032.x

Mylopoulos, M., & Woods, N. N. (2017). When I say … adaptive expertise. *Medical Education, 51*(7), 685-686. doi:10.1111/medu.13247

Nácul, M. P., Cavazzola, L. T., & Melo, M. C. (2015). Current status of residency training in laparoscopic surgery in Brazil: A critical review. *ABCD. Arquivos Brasileiros De Cirurgia Digestiva (São Paulo), 28*(1), 81-85. doi:10.1590/s0102-67202015000100020

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1-32. doi:10.1016/0010-0285(87)90002-8

Pinzon, D., Vega R., Sanchez, Y. P., & Zheng, B. (2016). Skill learning from kinesthetic feedback. *American Journal of Surgery, 214*(4), 721-725. doi: 10.1016/j.amjsurg.2016.10.018

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs and symbols, and other distinctions in human performance models. *IEEE Systems, Man, and Cybernetics Society 13*(3), 257 – 266. doi: 10.1109/TSMC.1983.6313160

Ruitenberg, M. F. L., de Kleine, E., van der Lubbe, R. H. J., Verwey, W. B., & Abrahamse, E. L.
(2012). *Context-dependent motor skill and the role of practice. 76*(6), 812-820. doi: 10.1007/s00426-011-0388-6

Sakai, K., Hikosaka, O., & Nakamura, K. (2004). Emergence of rhythm during motor learning. *Trends in Cognitive Sciences, 8*(12), 547-553. doi:10.1016/j.tics.2004.10.005

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*(4), 225-260. doi:10.1037/h0076770

Schreuder, H., Wolswijk, R., Zweemer, R., Schijven, M., & Verheijen, R. (2011). Training and learning robotic surgery, time for a more structured approach: a systematic review. BJOG: An International *Journal of Obstetrics & Gynaecology*, *119*(2), 137-149. doi:10.1111/j.1471-0528.2011.03139.x

Seidler, R. D., Bo, J., & Anguera, J. A. (2012). Neurocognitive Contributions to Motor Skill Learning: The Role of Working Memory. *Journal of Motor Behavior, 44*(6), 445-453. doi:10.1080/00222895.2012.672348

Verwey, W. B. (1996). Buffer loading and chunking in sequential keypressing. *Journal of Experimental Psychology: Human Perception and Performance, 22*(3), 544–562. doi: 10.1037/0096-1523.22.3.544

Willingham, D. B. (2001). Becoming aware of motor skill. *Trends in Cognitive Sciences, 5*(5), 181-182. doi:10.1016/s1364-6613(00)01652-1

Appendix A

Consent Form

This study is part of the Master thesis of Lida Z. David, student at the University of Twente. The study consists of two dexterity tasks (= fine motor activities that require good finger and hand to coordination), of 26 brief trials each.

The goal of the study is to explore the extent to which motor sequence learning is involved in learning a task. Simply put, we are trying to see if, after you learn how to perform a task, repeating the same sequences of movement will increase your performance further. This study is a piece of the puzzle that will help make training for laparoscopic surgery better. (Laparoscopic surgery is a type of surgery done through one or more small incisions, using small surgical instruments to operate. A tiny camera is inserted from one of the incisions and projects the area of interest to a flat screen).

The data recorded will be anonymous, so nobody will know your performance. You have the right to withdraw at any moment during the study. By signing below you agree that:

You are 18+
You understand the privacy of your data
You understand you have the right to withdraw

Before signing, please state the details asked below (these will be used only for demographic reasons and will not be crossed with your data scores):

Name: …………………………………….
Age: ………….
Handedness (left-handed/right handed/ambidextrous): ……………………………
Occupation: ………………………………….

Signature:

Appendix B

Instructions for each task

1. Mirror-drawing Instructions

The task you are about to perform is called “Mirror-drawing”. It consists of 2 parts, the first consisting of 16 trials, and the second 10. The goal is to to trace the frame presented to you in the box (point at frame), by looking at the mirror’s reflection instead of the paper. You have to use the pen provided (point at pencil), and draw in between the two black lines of the frame. You have to be as quick and as cautious as possible. Your performance will be timed. As soon as you are done, you have to state “done!”, so I can stop the timer and note your time. For the first part, I will change the positioning of the paper. I will remind you how many trials are left on the 10th and 14th trial. In the second part you’ll have to do the exact same thing for another 10 trials, only this time I will keep the same frame in front of you. You will start from this point (show point) and trace the outline continuously for all ten trials. Don’t worry for shouting done, I will watch you and keep count. I will again remind you how many trials are left on the 6th. If you have any questions, ask me now or in-between the trials, during the moments you are not being timed. Do you have any questions? (answer questions). Please tell what you have to do, so we make sure you understand.

Are you ready to begin?

Go.

2. Clips-and-string Instructions

This task is called “clips-and-string”. It consists of two parts, 16 trials each. The goal is to pass this string (point at string) through the clips (point at clips) of either the red or blue path, using these two tweezers (point at tweezers). I will tell you each time which path you have to follow. You have make your way around the boxes, and have to be as quick and as cautious as possible. Your performance will be timed. As soon as you are done, you have to state “done!”, so I can stop the timer and note your time. Then, I will change the position of the cardboard the same procedure will be repeated, until you complete the task 16 times. I will remind you how many trials are left on the 10th and 14th trial. This will be part 1. For part 2, you will have to do exactly the same for another 10 trials, only this time the the cardboard will remain still, and you will have to follow the same path for all 10 trails.. I will again remind you how many trials are left on the 6th trial. If you have any questions, ask me now or in-between the trials, during the moments you are not being timed. Do you have any questions? (answer questions). Please tell what you have to do, so we make sure you understand.

Are you ready to begin?

Go.

Appendix C

Syntax used in R

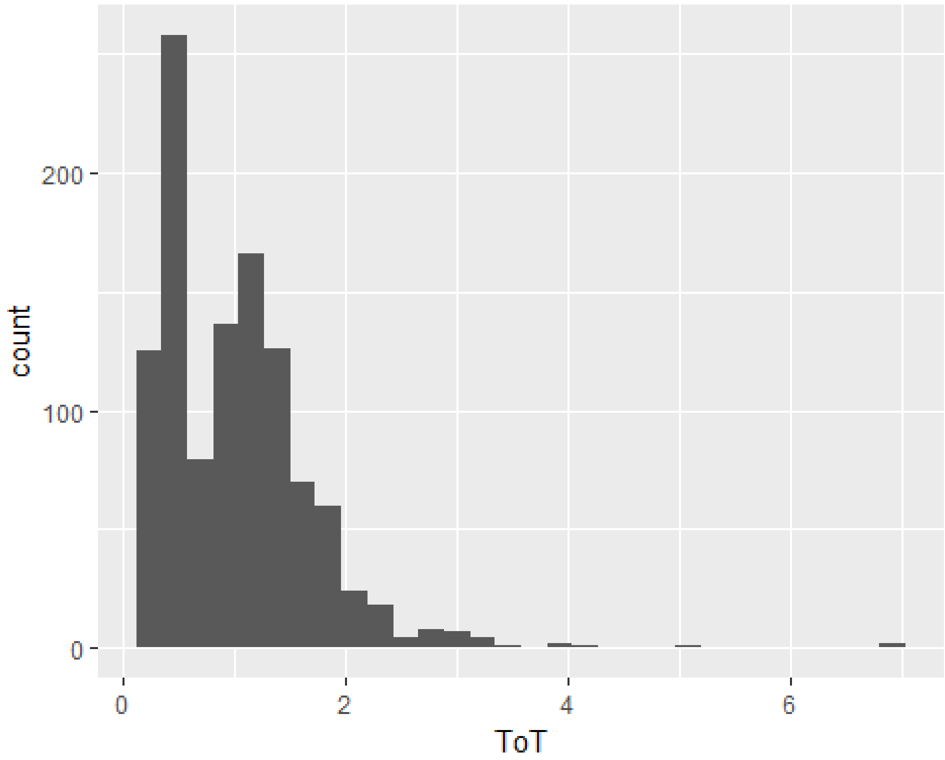
Data analysis protocol Lida David
Martin Schmettow
1 March 2018

Reading in data

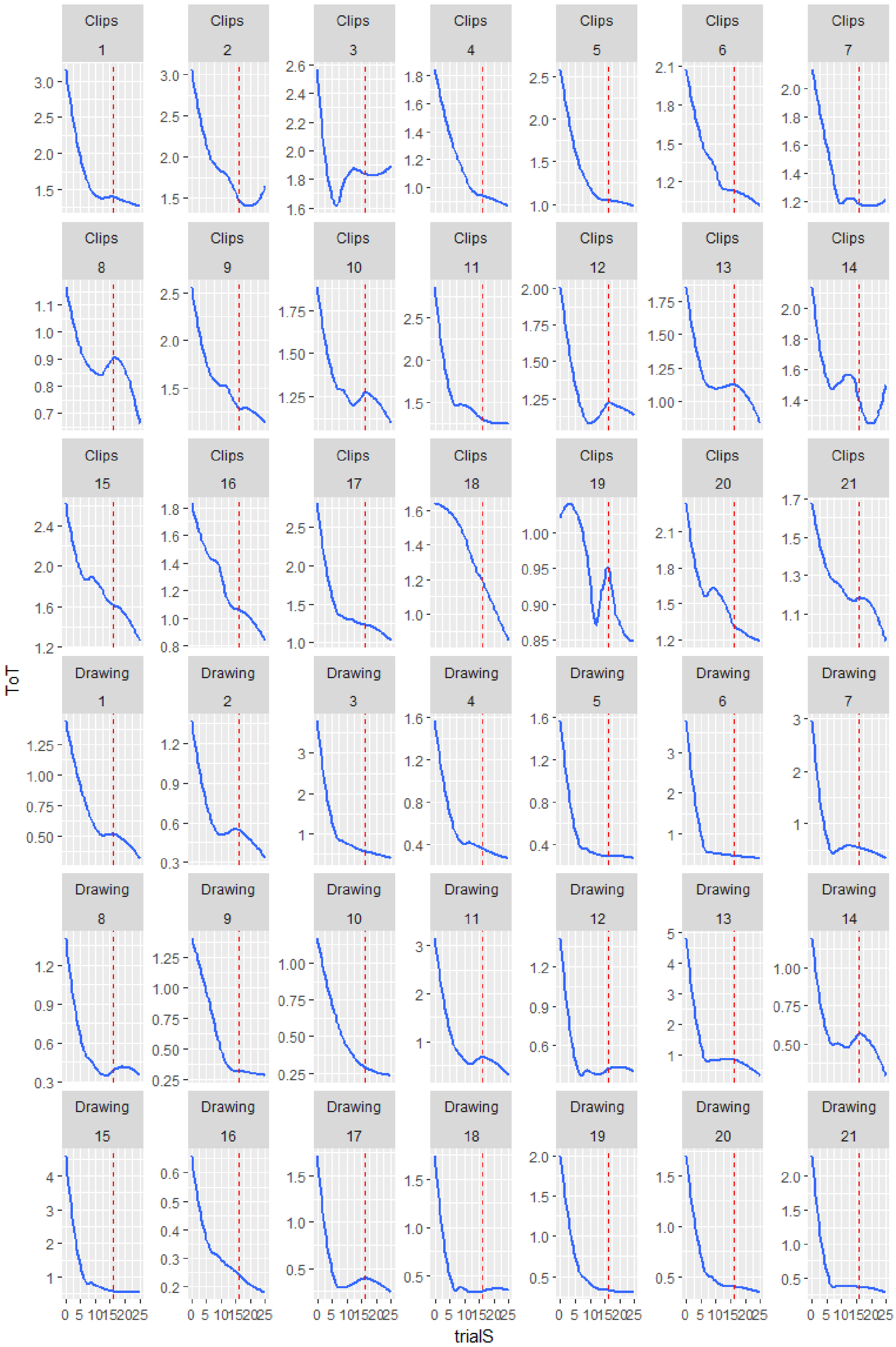
```
LD_18 <-
  readxl::read_excel("Data Thesis.xlsx") %>%
  mutate(trial = as.integer(trial -1)) %>%
  rename(trialS = trial) %>%
  mutate(Condition = ifelse(Condition == 1, "varied", "fixed"),
         trialM = ifelse(Condition == "fixed", trialS - 16, 0),
         Part = as.factor(Part)) %>%
  gather(key = Task, value = ToT, Clips, Drawing) %>%
  mutate(Obs = row_number()) %>%
  select(Obs, Part, Task, Condition, trialS, trialM, ToT)


ToT <-
  str_split_fixed(LD_18$ToT, "\\:", 3) %>%
  as_data_frame() %>%
  mutate(minutes = as.numeric(V2) + as.numeric(V3)/60)

LD_18$ToT <- ToT$minutes

LD_18

save(LD_18, file = "LD_18.Rda")
```

Exploratory data analysis

```
LD_18 %>%
  ggplot(aes(x = ToT)) +
  geom_histogram()

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




```
LD_18 %>%
  ggplot(aes(x = trialS, y = ToT)) +
  geom_smooth(se = F) +
  geom_vline(xintercept = 16, linetype = 2, col = "red") +
  facet_wrap(Task ~ Part, scales = "free_y")

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


Clips 1
Clips 2
Clips 3
Clips 4
Clips 5
Clips 6
Clips 7
Clips 8
Clips 9
Clips 10
Clips 11
Clips 12
Clips 13
Clips 14
Clips 15
Clips 16
Clips 17
Clips 18
Clips 19
Clips 20
Clips 21
Drawing 1
Drawing 2
Drawing 3
Drawing 4
Drawing 5
Drawing 6
Drawing 7
Drawing 8
Drawing 9
Drawing 10
Drawing 11
Drawing 12
Drawing 13
Drawing 14
Drawing 15
Drawing 16
Drawing 17
Drawing 18
Drawing 19
Drawing 20
Drawing 21
ToT
trialS

Regression

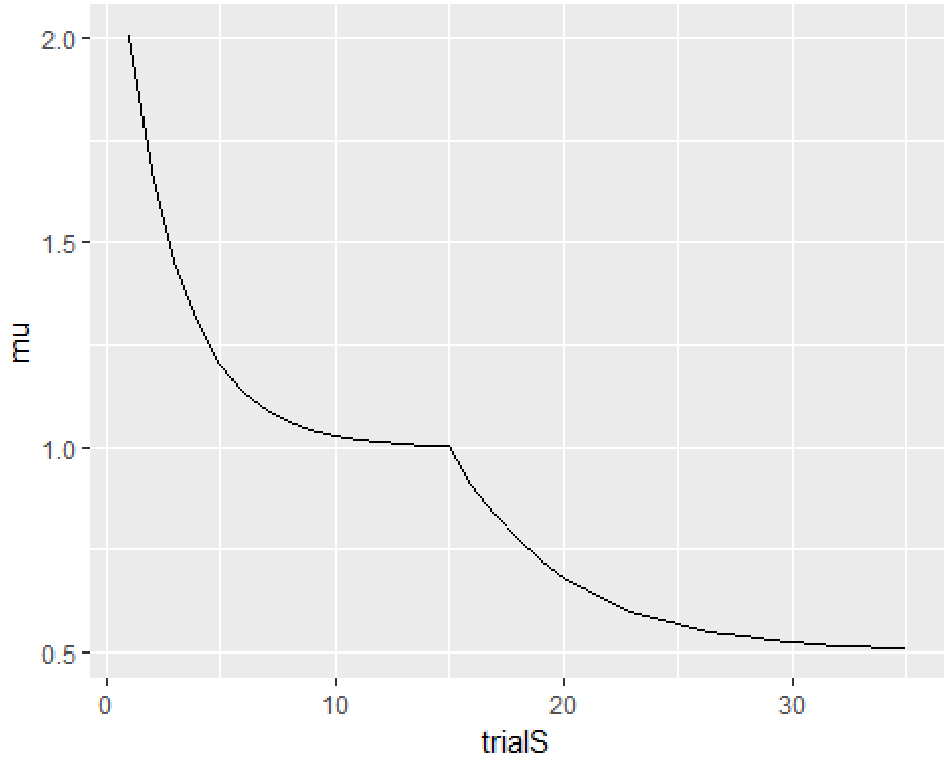
We estimate a two-component learning curve in asym-rate-ampl parametrization. Performance is composed of two components, that have their own learning trajectory. The skill component (SC) contains general coordination and strategy at the task, which is independent of spatial configuration. The motor component (MC) depends on the spatial configuration and is acquired by repeating the exact same sequence.

With the number of trials per component, t_S and t_M, and the learning parameters asymptote (maximum performance) ω, rate ρ_S, ρ_M and amplitude δ_S, δ_M, performance (ToT) takes the following form:

$$\text{performance} = \omega + \delta_S e^{-\rho_S t_S} + \delta_M e^{-\rho_M t_M}$$

The following model specifies a separate learning curves per task-by-participants. Correlations across tasks are included. The amplitude - rate - asymptote parametrization is used, like in the example below:

```
set.seed(42)

n_S = 15
n_M = 20
n = n_S + n_M
asym = 0.5
rateS = 0.4
amplS = 1.5
rateM = 0.2
amplM = 0.5



Part_1 <-
  data_frame(asym = rep(asym, n),
             rateS = rep(rateS, n),
             amplS = rep(amplS, n),
             rateM = rep(rateM, n),
             amplM = rep(amplM, n),
             trialS = 1:n,
             trialM = c(rep(0, n_S),(1:n_M))) %>%
  mutate(mu = asymptote::arary(amplS, rateS, amplM, rateM, asym, trialS, t
rialM))

Part_1  %>%
  ggplot(aes(x = trialS, y = mu)) +
  geom_line()
```



The following formulas are used for regression, the nonlinear function is from package *asymptote*.

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

## ToT ~ asym + amplS * exp(-rateS * trialS) + amplM * exp(-rateM *
##     trialM)
## <environment: namespace:asymptote>

F_ef_arary <- list(
  formula(asym ~ 0 + Task + (0 + Task|corr1|Part)),
  formula(rateS ~ 0 + Task + (0 + Task|corr2|Part)),
  formula(amplS ~ 0 + Task + (0 + Task|corr3|Part)),
  formula(rateM ~ 0 + Task + (0 + Task|corr4|Part)),
  formula(amplM ~ 0 + Task + (0 + Task|corr5|Part)))



# measurement scale
F_pr_arary <- c(set_prior("lognormal(1, 2)", nlpar = "asym", lb = 0),
                set_prior("lognormal(1.5, 1)", nlpar = "amplS", lb = 0),
                set_prior("beta(1.2, 1.2)", nlpar = "rateS", lb = 0, ub =
1),
                set_prior("lognormal(1.5, 1)", nlpar = "amplM", lb = 0),
                set_prior("beta(1.2, 1.2)", nlpar = "rateM", lb = 0, ub =
1))
```

```
F_pr_arary_2 <- c(set_prior("normal(0.5, 10)", nlpar = "asym", lb = 0),
                  set_prior("normal(1, 10)", nlpar = "amplS", lb = 0),
                  set_prior("normal(0.5, 10)", nlpar = "rateS", lb = 0),
                  set_prior("normal(1.5, 10)", nlpar = "amplM", lb = 0),
                  set_prior("normal(0.5, 10)", nlpar = "rateM", lb = 0))


F_pr_arary_gam <- c(set_prior("lognormal(1, 2)", nlpar = "asym", lb = 0),
                  set_prior("lognormal(1.5, 1)", nlpar = "amplS", lb = 0),
                  set_prior("beta(1.2, 1.2)", nlpar = "rateS", lb = 0, ub =
1),
                  set_prior("lognormal(1.5, 1)", nlpar = "amplM", lb = 0),
                  set_prior("beta(1.2, 1.2)", nlpar = "rateM", lb = 0, ub =
1),
                  set_prior("cauchy(0, 5)", class = "shape"))
```

M_5: ARARY with Gamma random component

Here we build the model and compile it with just a few chains.

```
M_5 <-
  brm(bf(ARARY,
         flist = F_ef_arary, nl = TRUE),
      prior = F_pr_arary_2,
      family = Gamma(link = identity),
      data = LD_18,
      iter = 150, warmup = 120,
      init = "0")
```

Using the compiled model above to run the full chains

```
M_5 <-
    brm(fit = M_5,
      data = LD_18,
      iter = 5000, warmup = 3000,
      init = "0",
      control = list(adapt_delta = .95))

M_5


P_5 <- posterior(M_5)
PP_5 <- post_pred(M_5)
LD_18$M_5 <- predict(M_5)$center
LD_18$M_5_resid <- LD_18$ToT - LD_18$M_5

save(M_5, P_5, PP_5, file = "M_5.Rda")
save(LD_18, file = "LD_18.Rda")
```

Results

```
load("M_5.Rda")

T_fixef <-
  P_5 %>% fixef()

T_fixef
```

Estimates with 95% credibility limits

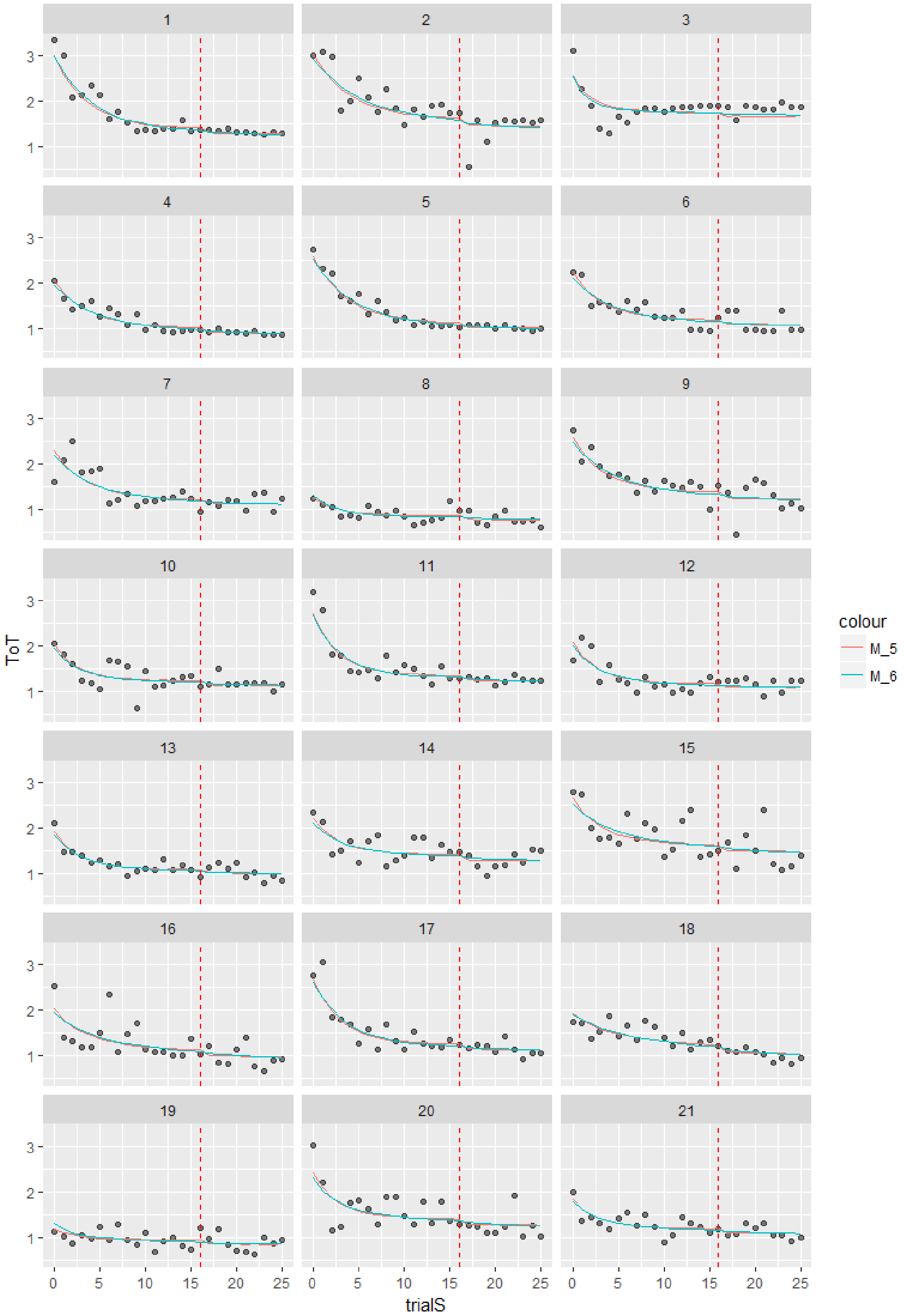
nonlin	fixef	center	lower	upper
asym	TaskClips	1.1577182	1.0291061	1.2638820
asym	TaskDrawing	0.2612140	0.1180004	0.3269353
rateS	TaskClips	0.3220994	0.2141153	0.5059187
rateS	TaskDrawing	0.5377853	0.3907563	0.7115918
amplS	TaskClips	1.0286249	0.7859518	1.3076140
amplS	TaskDrawing	2.0981479	1.5331471	2.6884107
rateM	TaskClips	10.1972400	0.4392050	26.3571275
rateM	TaskDrawing	0.1020007	0.0407116	0.1833435
amplM	TaskClips	0.1054591	0.0395542	0.1794875
amplM	TaskDrawing	0.1866054	0.0900881	0.3412478

Individual learning curves

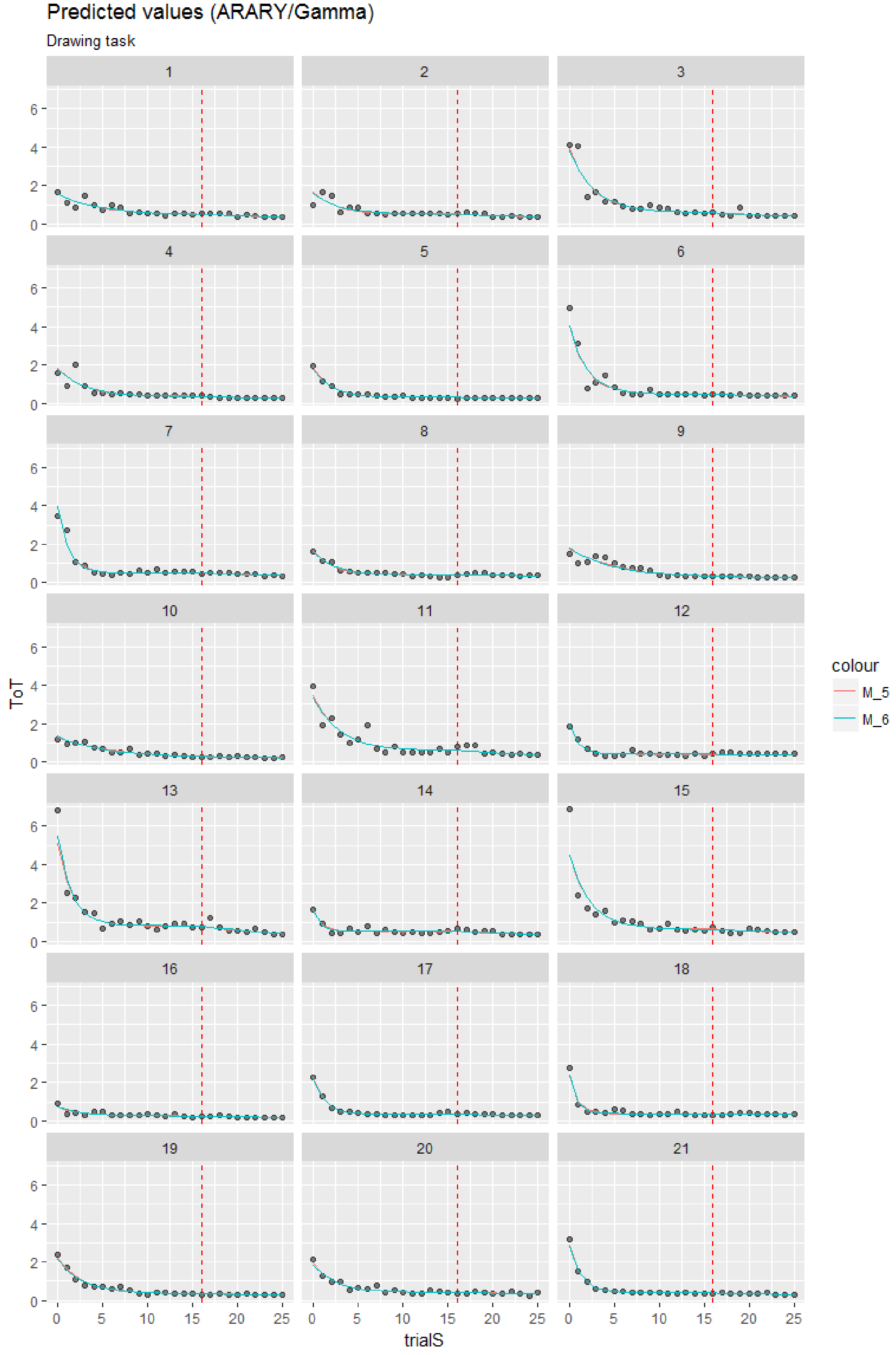
```
LD_18 %>%
  filter(Task == "Clips") %>%
  ggplot(aes(x = trialS, y = ToT)) +
  geom_point(alpha = .5) +
  geom_line(aes(y = M_5, color = "M_5")) +
  geom_line(aes(y = M_6, color = "M_6")) +
  geom_vline(xintercept = 16, linetype = 2, col = "red") +
  facet_wrap(~ Part, nrow = 7) +
  ggtitle(list(title = "Predicted values (ARARY/Gamma)",
               subtitle = "Clips task"))
```

Predicted values (ARARY/Gamma)

Clips task



```
LD_18 %>%
  filter(Task == "Drawing") %>%
  ggplot(aes(x = trialS, y = ToT)) +
  geom_point(alpha = .5) +
  geom_line(aes(y = M_5, color = "M_5")) +
  geom_line(aes(y = M_6, color = "M_6")) +
  geom_vline(xintercept = 16, linetype = 2, col = "red") +
  facet_wrap(~ Part, nrow = 7) +
  ggtitle(list(title = "Predicted values (ARARY/Gamma)",
               subtitle = "Drawing task"))
```


Predicted values (ARARY/Gamma)
Drawing task
ToT
trialS
colour
M_5
M_6

Correlation of learning parameters across tasks

```
## Extracting correlations from the posterior

P_cor <-
  P_5 %>%
  filter(type == "cor")

P_cor_names <-
  P_cor %>%
  distinct(parameter) %>%
  separate(parameter, c("cor", "re_factor", "nonlin", "Task", "nonlin_2",
"task_2"), remove = F) %>%
  select(parameter, nonlin)

T_cor <-
  P_cor %>%
  select(-nonlin) %>%
  left_join(P_cor_names, by = "parameter") %>%
  select(nonlin, value) %>%
  group_by(nonlin) %>%
  summarize(center = median(value),
            lower = quantile(value, .025),
            upper = quantile(value, .975))

T_cor %>% kable(digits = 4)
```

nonlin	center	lower	upper
amplM	-0.1113	-0.9249	0.8896
amplS	-0.0669	-0.5544	0.4857
asym	0.4362	-0.3350	0.9278
rateM	-0.0122	-0.9475	0.9527
rateS	-0.0585	-0.8218	0.7646

Individual-level analysis

The research question (amount of motor learning) has previously been answered on the population level. Due to the multi-level structure of the model, it can also be examined on individual level. However, the participant level random effects are differences towards the population mean. We compute absolute *scores* per participant on the posterior and summarize them.

```
P_5_fixef <-
  posterior(M_5, type = "fixef") #%>%
  #filter(nonlin %in% c("amplS", "amplM")) %>%
  # posterior()

P_5_ranef <-
  posterior(M_5, type = "ranef")# %>%
```

```
  #filter(nonlin == "asym") %>%
  #posterior()

P_5_scores <-
  left_join(P_5_ranef, P_5_fixef,
            by = c("model", "chain", "iter", "fixef", "nonlin"),
            suffix = c("", "_fixef"))

P_5_scores$value = P_5_scores$value + P_5_scores$value_fixef

P_5_scores %>%
  group_by(fixef, nonlin, re_entity) %>%
  summarize(center = median(value),
            lower = quantile(value, .025),
            upper = quantile(value, .975)) %>%
  select(Task = fixef, nonlin, Part = re_entity, center, lower,upper) %>%
  ungroup() %>%
  sample_n(10) %>%
  kable()
```

Task	nonlin	Part	center	lower	upper
TaskDrawing	rateS	12	0.9490302	0.6945260	1.2807423
TaskDrawing	rateS	15	0.4238480	0.3215025	0.5753750
TaskClips	asym	6	1.0849301	0.9587774	1.1946560
TaskDrawing	rateM	12	0.0616261	-0.0264168	0.1988777
TaskDrawing	amplS	15	3.8373221	2.9942492	4.9751266
TaskDrawing	amplM	10	0.0128850	-0.1452042	0.1857661
TaskDrawing	asym	12	0.2942147	0.0964226	0.4350132
TaskClips	amplS	15	1.0893530	0.6094646	1.6894599
TaskDrawing	rateS	5	0.5335334	0.3712991	0.7589465
TaskDrawing	amplS	4	1.4910178	1.0897490	2.0628405

Model criticism

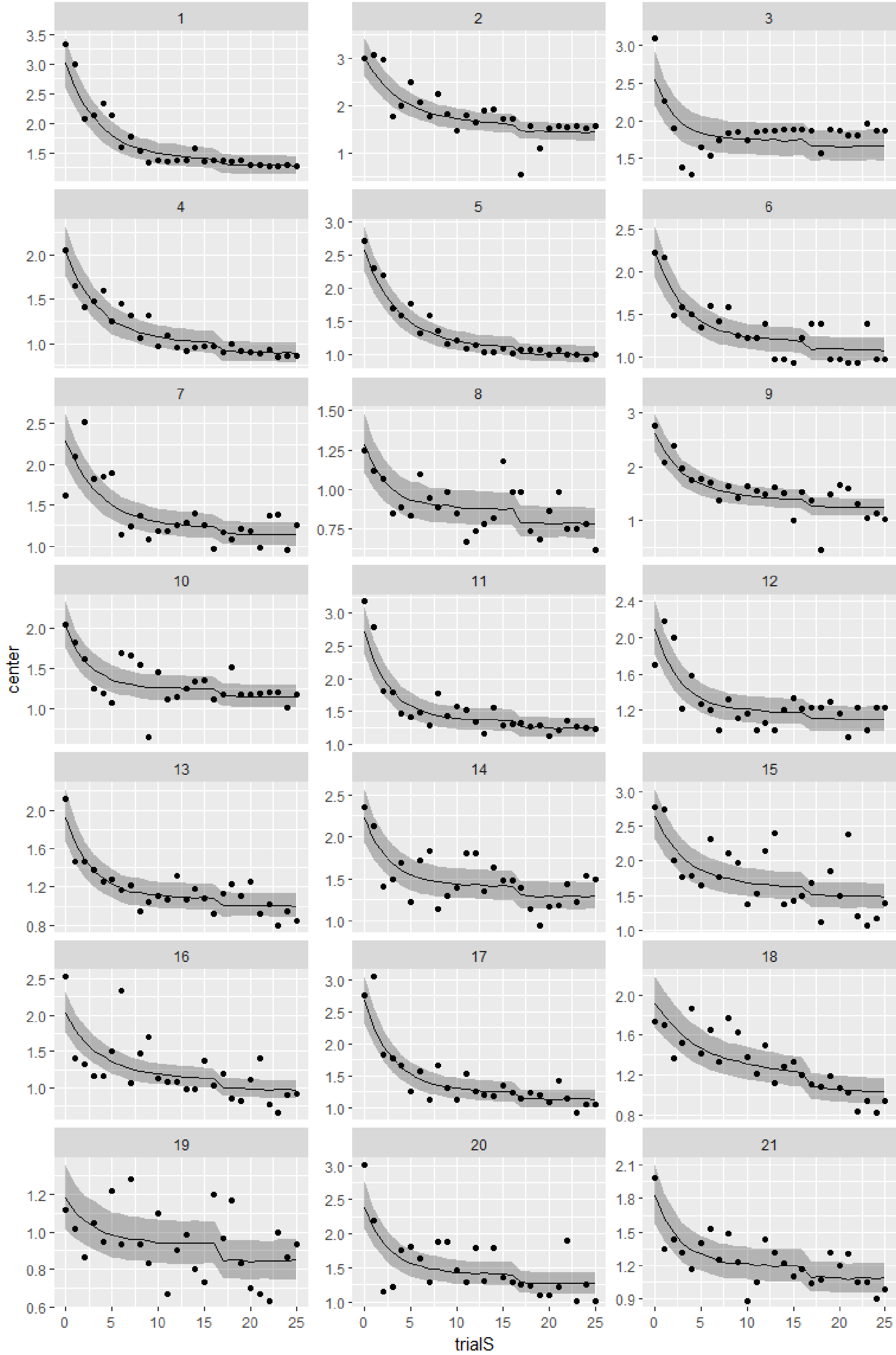
Predicted values analysis

We check whether the prediction intervals (25%, 50% and 75% quantiles) cover the observations.

```
T_PP_5  <-
  PP_5 %>%
  group_by(Obs) %>%
  summarize(lower = quantile(value, .25),
            center = quantile(value, .50),
            upper = quantile(value, .75)) %>%
  right_join(LD_18, by = "Obs")
```

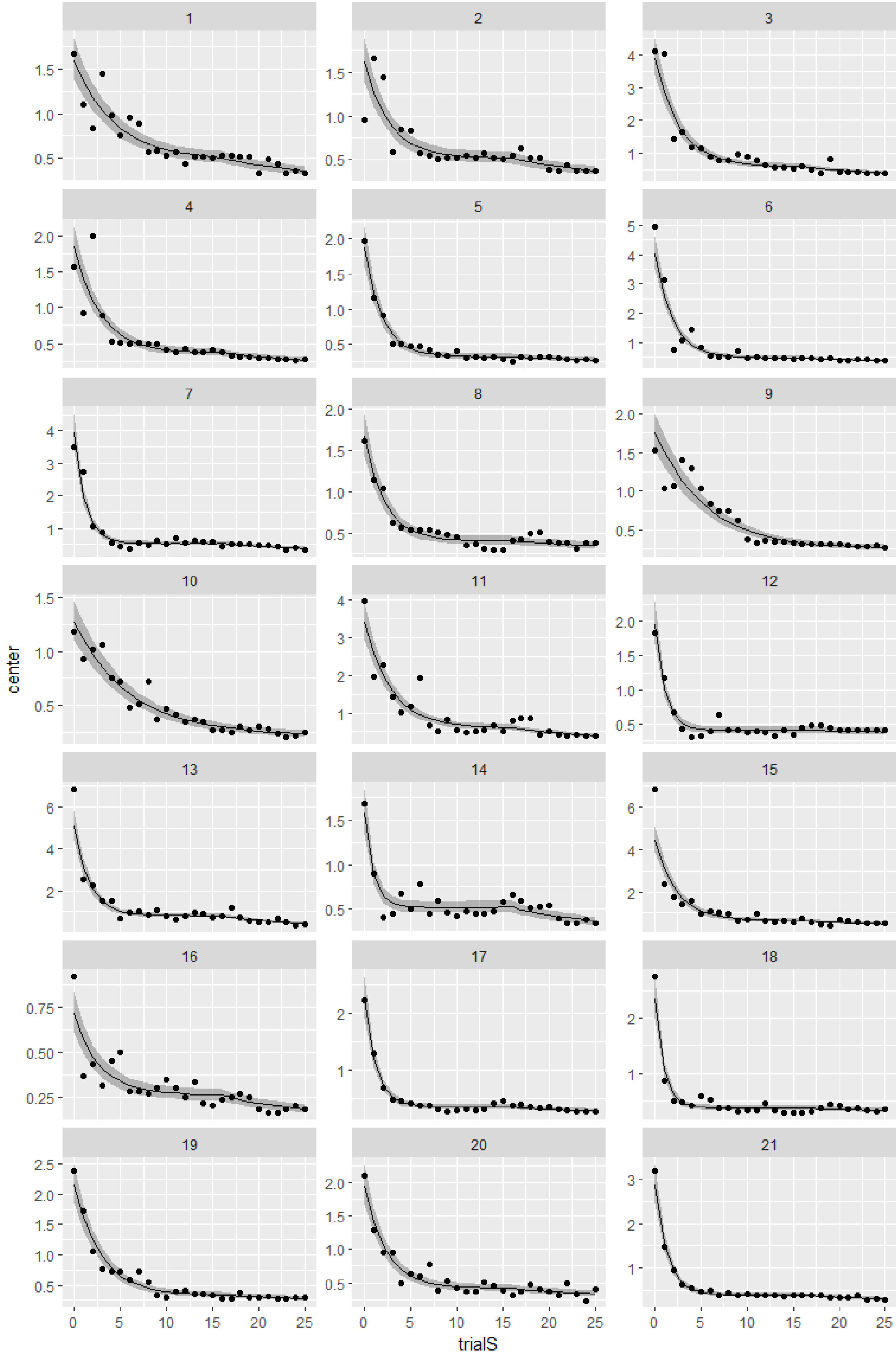
Task: Clips

```
T_PP_5 %>%
  filter(Task == "Clips") %>%
  ggplot(aes(x = trialS)) +
  geom_line(aes(y = center)) +
  geom_ribbon(aes(ymin = lower, ymax = upper), alpha = .3) +
  geom_point(aes(y = ToT)) +
  facet_wrap(~ Part, nrow = 7, scales = "free_y")
```



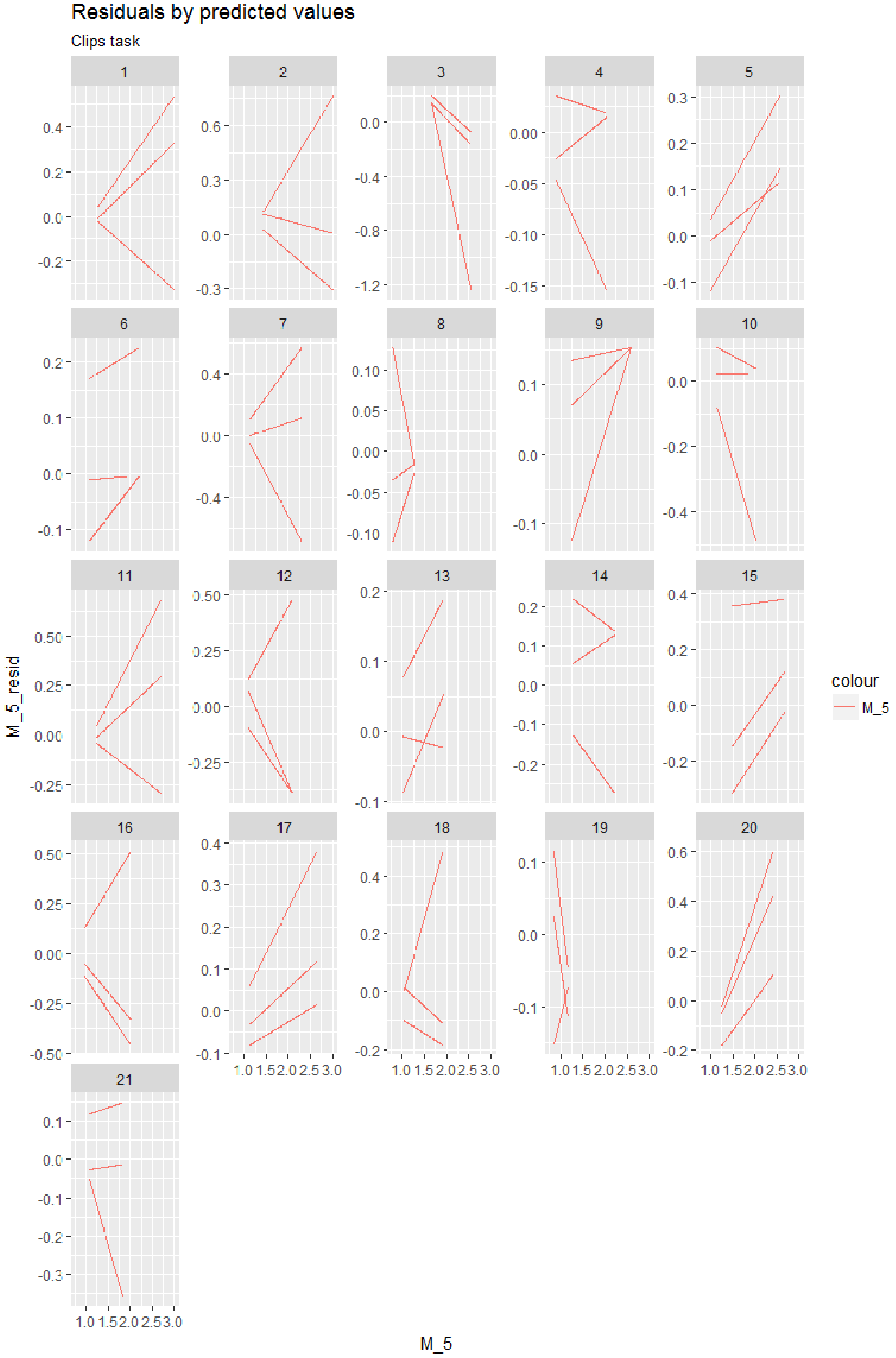
Task: Drawing

```
T_PP_5 %>%
  filter(Task == "Drawing") %>%
  ggplot(aes(x = trialS)) +
  geom_line(aes(y = center)) +
  geom_ribbon(aes(ymin = lower, ymax = upper), alpha = .3) +
  geom_point(aes(y = ToT)) +
  facet_wrap(~ Part, nrow = 7, scales = "free_y")
```

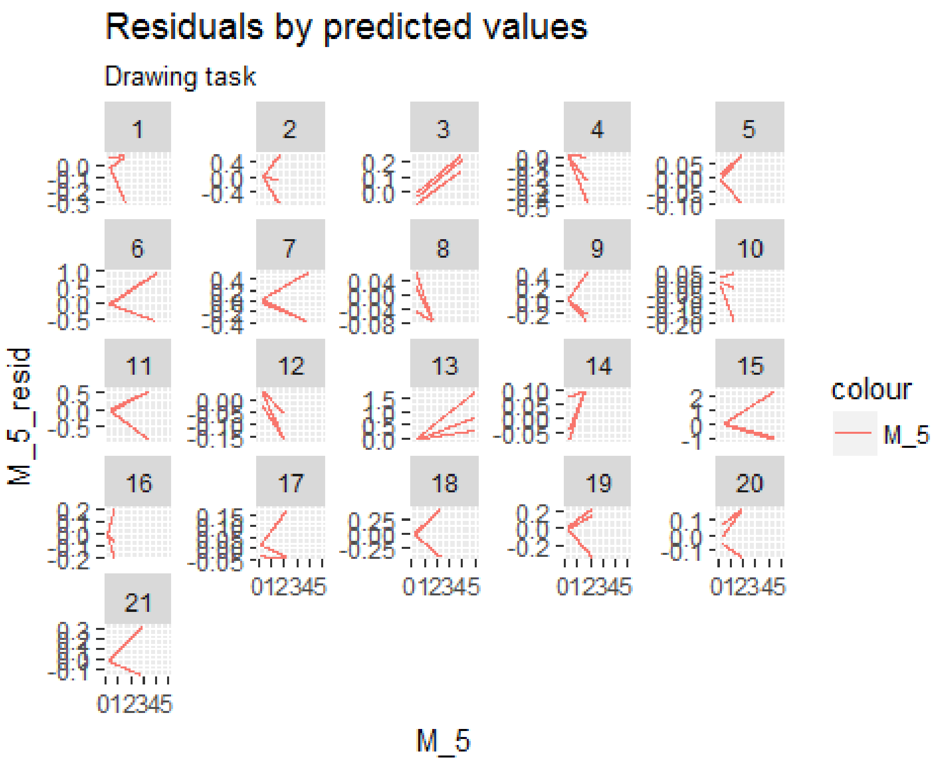

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
center
trialS

Residuals analysis

```
LD_18 %>%
  filter(Task == "Clips") %>%
  ggplot() +
  geom_quantile(aes(x = M_5, y = M_5_resid, col = "M_5"), formula = y ~ x)
+
  facet_wrap(~ Part, scales = "free_y") +
  ggtitle(list(title = "Residuals by predicted values",
               subtitle = "Clips task"))

## Loading required package: SparseM

##
## Attaching package: 'SparseM'

## The following object is masked from 'package:base':
##
##     backsolve
```


Residuals by predicted values
Clips task
M_5_resid
M_5
colour
M_5

```
LD_18 %>%
  filter(Task == "Drawing") %>%
  ggplot() +
  geom_quantile(aes(x = M_5, y = M_5_resid, col = "M_5"), formula = y ~ x)
+
  facet_wrap(~ Part, scales = "free_y") +
  ggtitle(list(title = "Residuals by predicted values",
               subtitle = "Drawing task"))
```



M_6: LARARY/Gamma with linear scaled LC parameters

The following model is an ARARY model, except that parameters have been linearized (L) inside the nonlinear function (*exp*) to have a range $-\infty, \infty$, i.e. all parameters are on a log scale. Estimates on scale measurement are obtained by the mean function *exp*.

```
LARARY <- formula(
  ToT ~ exp(amplS - exp(rateS) * trialS) + exp(amplM - exp(rateM) * trialM
) + exp(asym)
  )

F_pr_larary_gam <- c(set_prior("normal(-1, 5)", nlpar = "asym"),
                     set_prior("normal(0, 10)", nlpar = "amplS"),
                     set_prior("normal(-1, 2)", nlpar = "rateS"),
                     set_prior("normal(0, 10)", nlpar = "amplM"),
                     set_prior("normal(-1, 2)", nlpar = "rateM"))
```

```
M_6 <-
  brm(bf(LARARY,
         flist = F_ef_arary, nl = TRUE),
      prior = F_pr_larary_gam,
      family = Gamma(link = identity),
      data = LD_18,
      iter = 200,
      inits = "0")

M_6 <-
  brm(fit = M_6,
      iter = 6000, warmup = 4000,
      control = list(adapt_delta = .95,
                     max_treedepth = 15),
      inits = "0")


#M_6


P_6 <- posterior(M_6)
PP_6 <- post_pred(M_6)
LD_18$M_6 <- predict(M_6)$center
LD_18$M_6_resid <- LD_18$ToT - LD_18$M_6

save(M_6, P_6, PP_6, file = "M_6.Rda")
save(LD_18, file = "LD_18.Rda")
```

Results

```
#load("M_6.Rda")

P_6_fixef <-  posterior(M_6, type = "fixef")
P_6_ranef <-  posterior(M_6, type = "ranef")
```

The following estimates are on a log scale.

```
fixef(P_6)
```

Estimates with 95% credibility limits

nonlin	fixef	center	lower	upper
asym	TaskClips	0.0903742	-6.0264817	0.2038210
asym	TaskDrawing	-1.6057879	-11.3528430	-1.1510160
rateS	TaskClips	-1.4435217	-1.9117564	-0.8622703
rateS	TaskDrawing	-0.7634813	-1.0602916	-0.4811777
amplS	TaskClips	-0.0523939	-0.3331043	0.1943617
amplS	TaskDrawing	0.5920581	0.2967064	0.8696491
rateM	TaskClips	-1.0785965	-4.8280385	2.8546129

rateM	TaskDrawing	-2.7952877	-3.9099986	-1.7869940
amplM	TaskClips	-3.0740102	-21.7739050	0.1576592
amplM	TaskDrawing	-1.5666336	-2.6576229	-0.7744704

Here are the estimates on scale of measurement:

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

Estimates with 95% credibility limits

nonlin	fixef	center	lower	upper
asym	TaskClips	1.0945838	0.0024140	1.2260787
asym	TaskDrawing	0.2007313	0.0000117	0.3163152
rateS	TaskClips	0.2360948	0.1478205	0.4222025
rateS	TaskDrawing	0.4660412	0.3463548	0.6180551
amplS	TaskClips	0.9489550	0.7166955	1.2145355
amplS	TaskDrawing	1.8077051	1.3454203	2.3860734
rateM	TaskClips	0.3400725	0.0080022	17.3677221
rateM	TaskDrawing	0.0610973	0.0200405	0.1674628
amplM	TaskClips	0.0462354	0.0000000	1.1707671
amplM	TaskDrawing	0.2087467	0.0701147	0.4609479

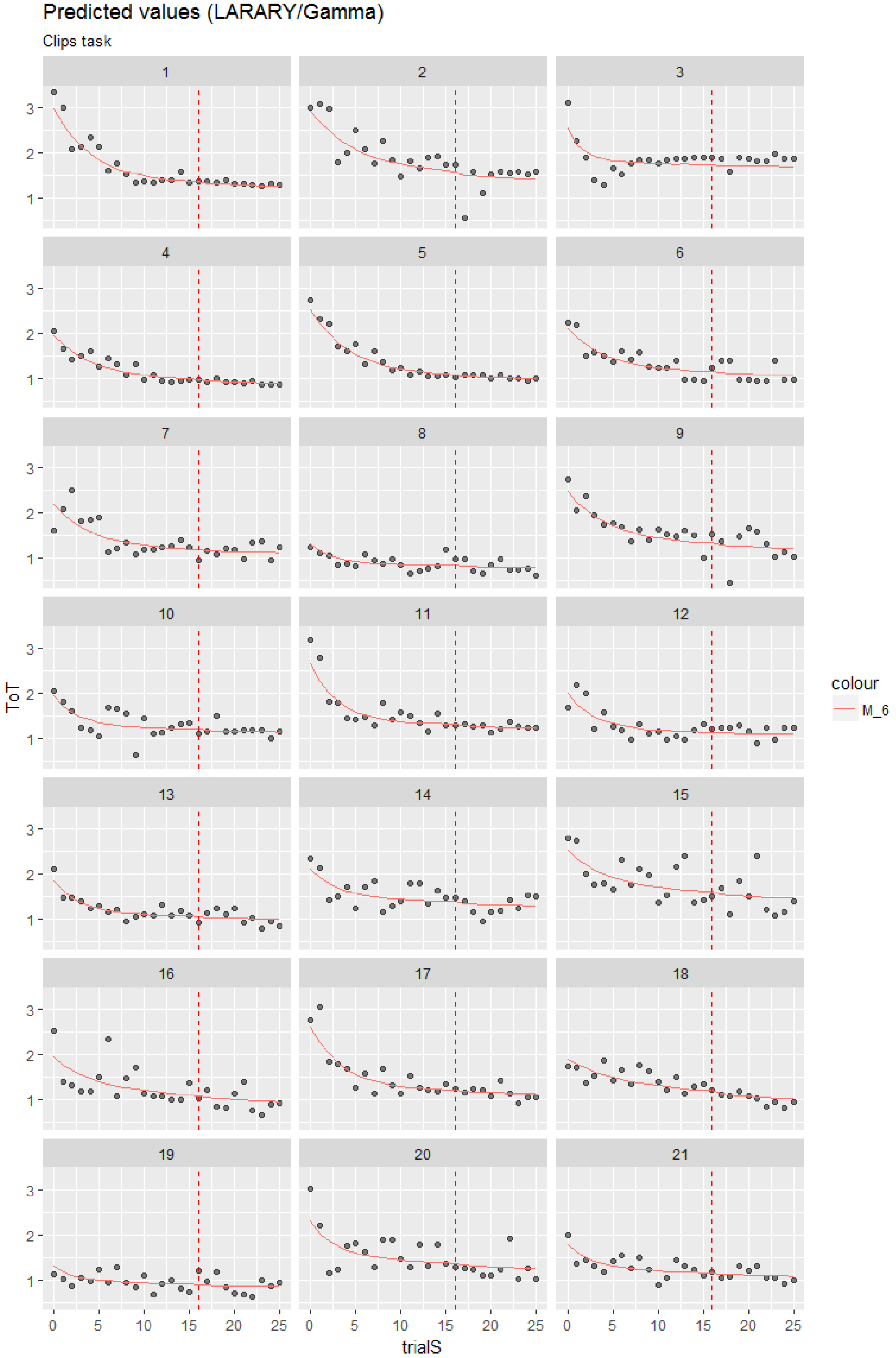
The ratio of amplitudes directly answers the question of what relative impact motor sequence learning has. The proportion of MSL on overall learning is not directly represented in the model, but can be created on the posterior.

```
T_6_prop_fixef <-
  P_6_fixef %>%
  select(chain, iter, fixef, nonlin, value) %>%
  filter(nonlin %in% c("amplS", "amplM")) %>%
  mutate(value = exp(value)) %>%
  spread(nonlin, value) %>%
  mutate(propM = amplM/(amplS + amplM)) %>%
  group_by(fixef) %>%
  summarize(center = median(propM),
            lower = quantile(propM, .025),
            upper = quantile(propM, .975))

kable(T_6_prop_fixef, digits = 4)
```

fixef	center	lower	upper
TaskClips	0.0461	0.0000	0.5529
TaskDrawing	0.1048	0.0361	0.2200

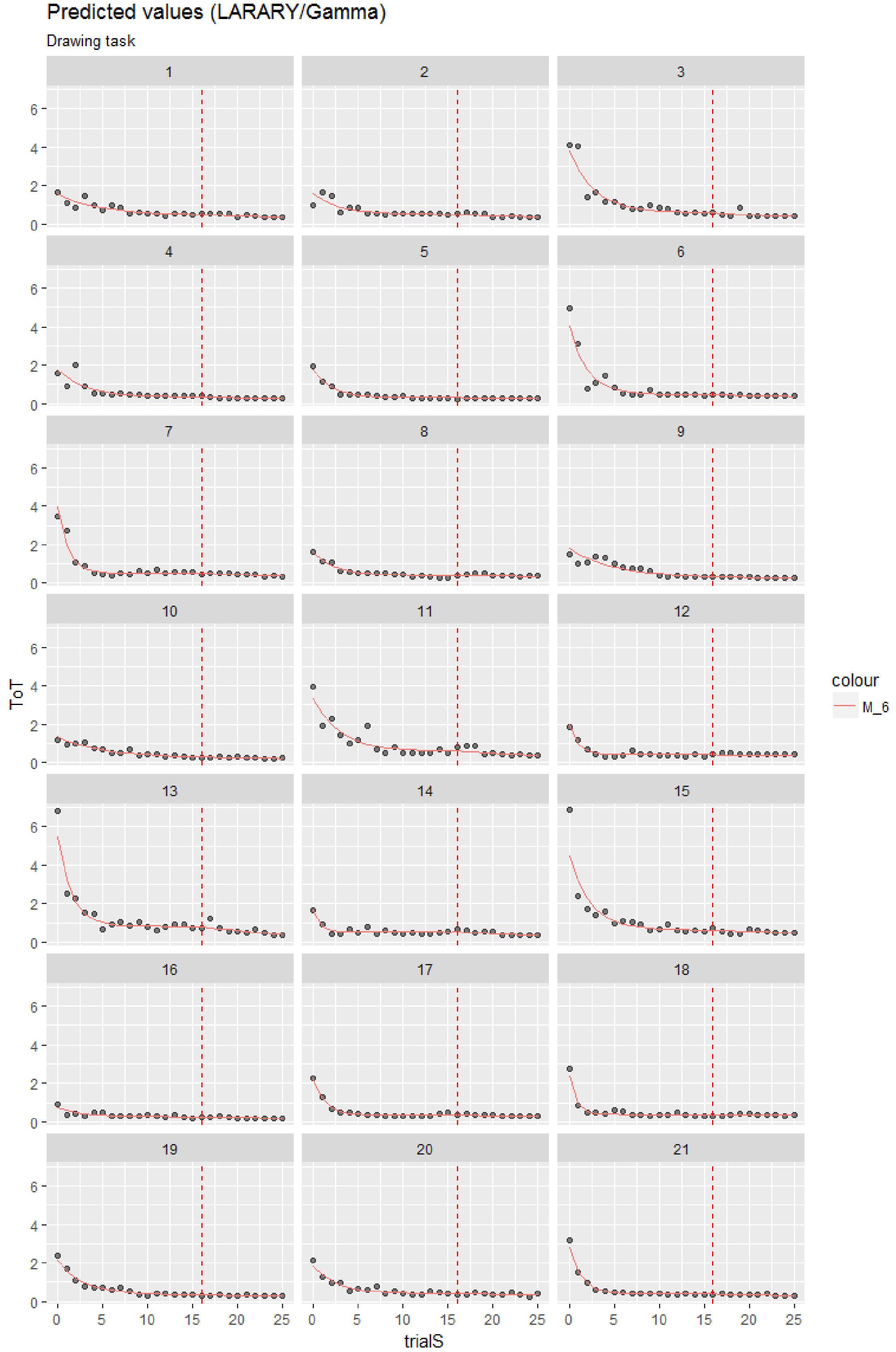
Individual-level anaysis

```
LD_18 %>%
  filter(Task == "Clips") %>%
  ggplot(aes(x = trialS, y = ToT)) +
  geom_point(alpha = .5) +
  geom_line(aes(y = M_6, color = "M_6")) +
  geom_line(aes(y = M_6, color = "M_6")) +
  geom_vline(xintercept = 16, linetype = 2, col = "red") +
  facet_wrap(~ Part, nrow = 7) +
  ggtitle(list(title = "Predicted values (LARARY/Gamma)",
               subtitle = "Clips task"))
```



Predicted values (LARARY/Gamma)
Clips task
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
ToT
trialS
colour
M_6

```
LD_18 %>%
  filter(Task == "Drawing") %>%
  ggplot(aes(x = trialS, y = ToT)) +
  geom_point(alpha = .5) +
  geom_line(aes(y = M_6, color = "M_6")) +
  geom_line(aes(y = M_6, color = "M_6")) +
  geom_vline(xintercept = 16, linetype = 2, col = "red") +
  facet_wrap(~ Part, nrow = 7) +
  ggtitle(list(title = "Predicted values (LARARY/Gamma)",
               subtitle = "Drawing task"))
```


Predicted values (LARARY/Gamma)
Drawing task
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
ToT
trialS
colour
M_6

The research question (amount of motor learning) has previously been answered on the population level. Due to the multi-level structure of the model, it can also be examined on individual level. However, the participant level random effects are differences towards the population mean. We compute absolute *scores* per participant on the posterior and summarize them.

```
P_6_scores <-
  left_join(P_6_ranef, P_6_fixef,
            by = c("model", "chain", "iter", "fixef", "nonlin"),
            suffix = c("", "_fixef"))

P_6_scores$value = exp(P_6_scores$value + P_6_scores$value_fixef)


T_6_scores <-
  P_6_scores %>%
  group_by(fixef, nonlin, re_entity) %>%
  summarize(center = median(value),
            lower = quantile(value, .025),
            upper = quantile(value, .975)) %>%
  select(Task = fixef, nonlin, Part = re_entity, center, lower,upper) %>%
  ungroup()


T_6_scores  %>%
  sample_n(10) %>%
  kable()
```

Task	nonlin	Part	center	lower	upper
TaskClips	rateS	20	0.2406953	0.0823091	1.3439111
TaskDrawing	asym	12	0.2251756	0.0000040	0.4001455
TaskDrawing	rateS	14	1.3330630	0.7133088	2.3221068
TaskClips	asym	19	0.8377900	0.0018241	0.9372280
TaskClips	rateS	2	0.1817135	0.0794867	0.3746833
TaskDrawing	rateS	12	0.9900177	0.6901883	1.3914052
TaskDrawing	rateS	19	0.3299586	0.2519067	0.4350551
TaskClips	amplS	11	1.3617665	0.8826241	2.1773698
TaskDrawing	rateS	5	0.5043384	0.3616376	0.7169677
TaskDrawing	amplM	16	0.1226252	0.0278308	0.2841341

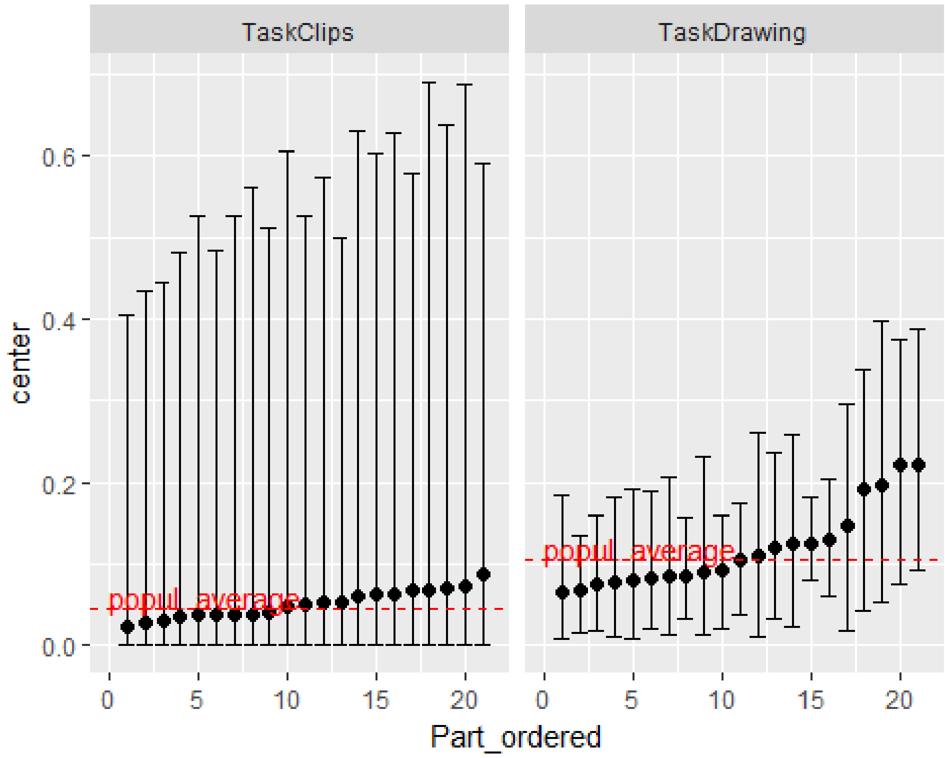
Like above, we compute the relative amount of MSL on participant level:

```
T_6_prop_ranef <-
  P_6_scores %>%
  select(chain, iter, fixef, nonlin, re_entity, value) %>%
  filter(nonlin %in% c("amplS", "amplM")) %>%
  spread(nonlin, value) %>%
```

```
  mutate(propM = amplM/(amplS + amplM)) %>%
  group_by(fixef, re_entity) %>%
  summarize(center = median(propM),
            lower = quantile(propM, .025),
            upper = quantile(propM, .975)) %>%
  ungroup()


T_6_prop_ranef %>%
  sample_n(10) %>%
  kable()
```

fixef	re_entity	center	lower	upper
TaskDrawing	5	0.0863199	0.0146132	0.2074805
TaskClips	14	0.0714128	0.0000000	0.6375967
TaskDrawing	7	0.0863594	0.0330473	0.1559368
TaskDrawing	4	0.1194728	0.0321097	0.2353048
TaskClips	12	0.0393289	0.0000000	0.5608916
TaskClips	8	0.0637771	0.0000000	0.6272537
TaskClips	3	0.0676844	0.0000000	0.6887665
TaskDrawing	19	0.0788545	0.0110844	0.1821584
TaskClips	4	0.0359640	0.0000000	0.4821311
TaskDrawing	15	0.0925961	0.0211989	0.1602233

```
T_6_prop_ranef %>%
  group_by(fixef) %>%
  mutate(Part_ordered = min_rank(center)) %>%
  ggplot(aes(x = Part_ordered, y = center, ymin = lower, ymax = upper)) +
  facet_grid(~fixef) +
  geom_point(size = 2) +
  geom_errorbar() +
  geom_hline(aes(yintercept = center),
             color = "red", linetype   = 2,
             data = T_6_prop_fixef) +
  geom_text(aes(y = center, x = 0), label   = "popul. average",
            color = "red", vjust = 0, hjust = 0,
             data = T_6_prop_fixef)
```



Correlation of learning parameters across tasks

Note that the correlations are computed on the linear predictor level.

```
## Extracting correlations from the posterior

P_cor <-
  P_6 %>%
  filter(type == "cor")

P_cor_names <-
  P_cor %>%
  distinct(parameter) %>%
  separate(parameter, c("cor", "re_factor", "nonlin", "Task", "nonlin_2",
"task_2"), remove = F) %>%
  select(parameter, nonlin)

T_cor <-
  P_cor %>%
  select(-nonlin) %>%
  left_join(P_cor_names, by = "parameter") %>%
  select(nonlin, value) %>%
  group_by(nonlin) %>%
  summarize(center = median(value),
            lower = quantile(value, .025),
            upper = quantile(value, .975))
```

```
T_cor %>% kable(digits = 4)
```

nonlin	center	lower	upper
amplM	0.1305	-0.9264	0.9085
amplS	-0.0342	-0.5311	0.4874
asym	0.3180	-0.8440	0.9466
rateM	-0.0163	-0.9478	0.9469
rateS	-0.0551	-0.7593	0.6940

Model criticism

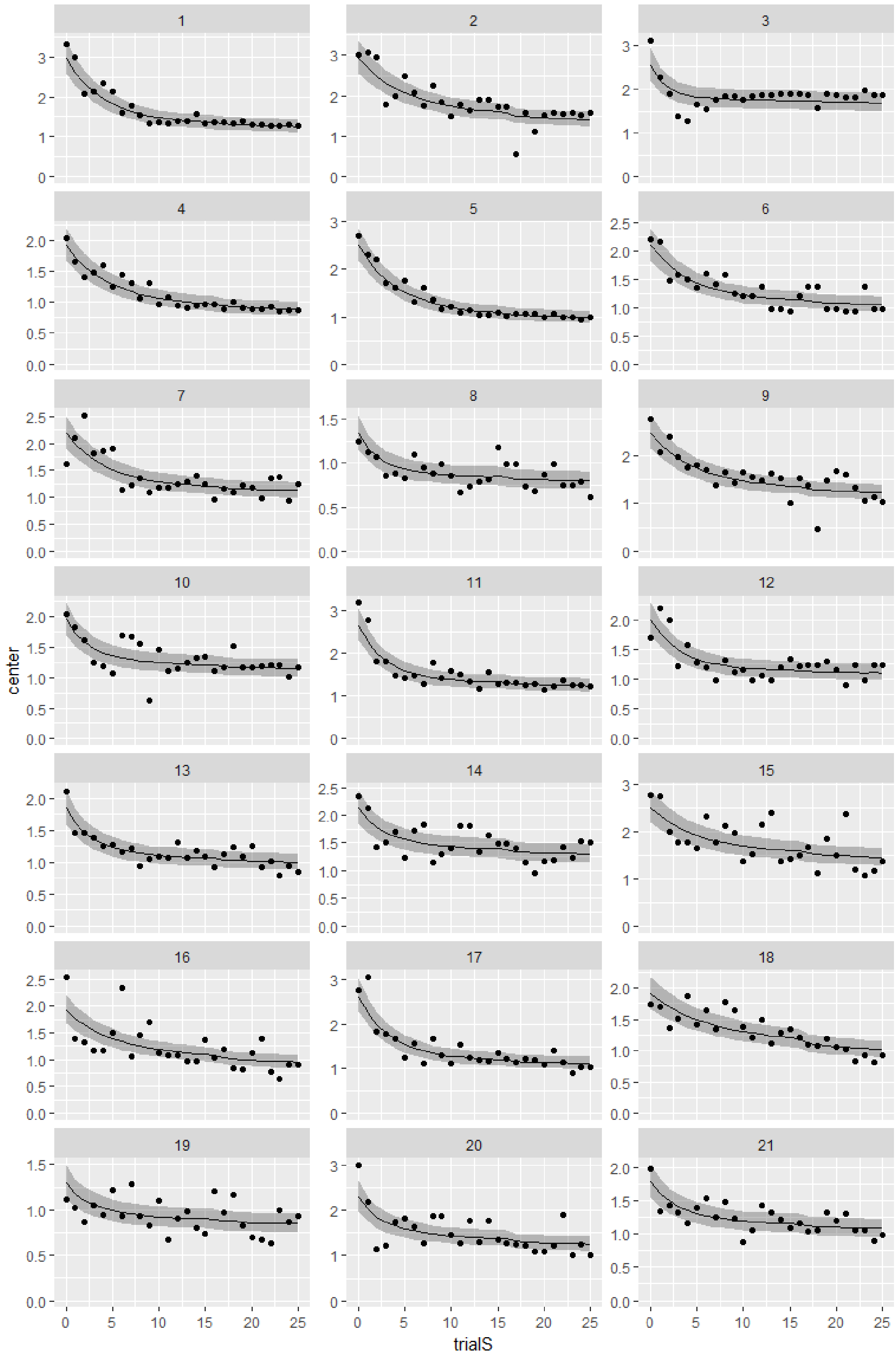
Predicted values analysis

We check whether the prediction intervals (25%, 50% and 75% quantiles) cover the observations.

```
T_PP_6  <-
  PP_6 %>%
  group_by(Obs) %>%
  summarize(lower = quantile(value, .25),
            center = quantile(value, .50),
            upper = quantile(value, .75)) %>%
  right_join(LD_18, by = "Obs")
```

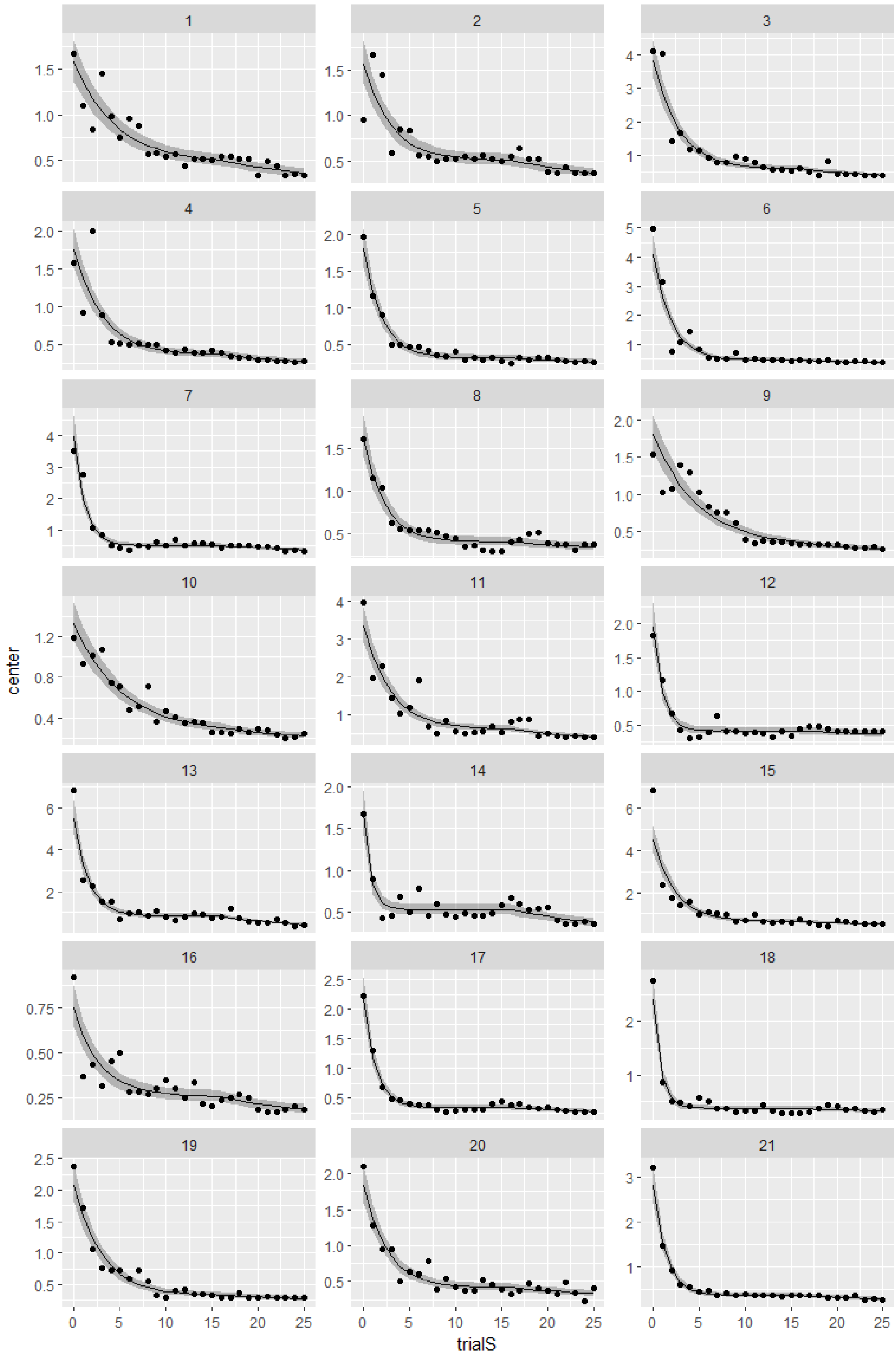
Task: Clips

```
T_PP_6 %>%
  filter(Task == "Clips") %>%
  ggplot(aes(x = trialS)) +
  facet_wrap(~ Part, nrow = 7, scales = "free_y") +
  geom_line(aes(y = center)) +
  geom_ribbon(aes(ymin = lower, ymax = upper), alpha = .3) +
  geom_point(aes(y = ToT)) +
  expand_limits(y = 0)
```



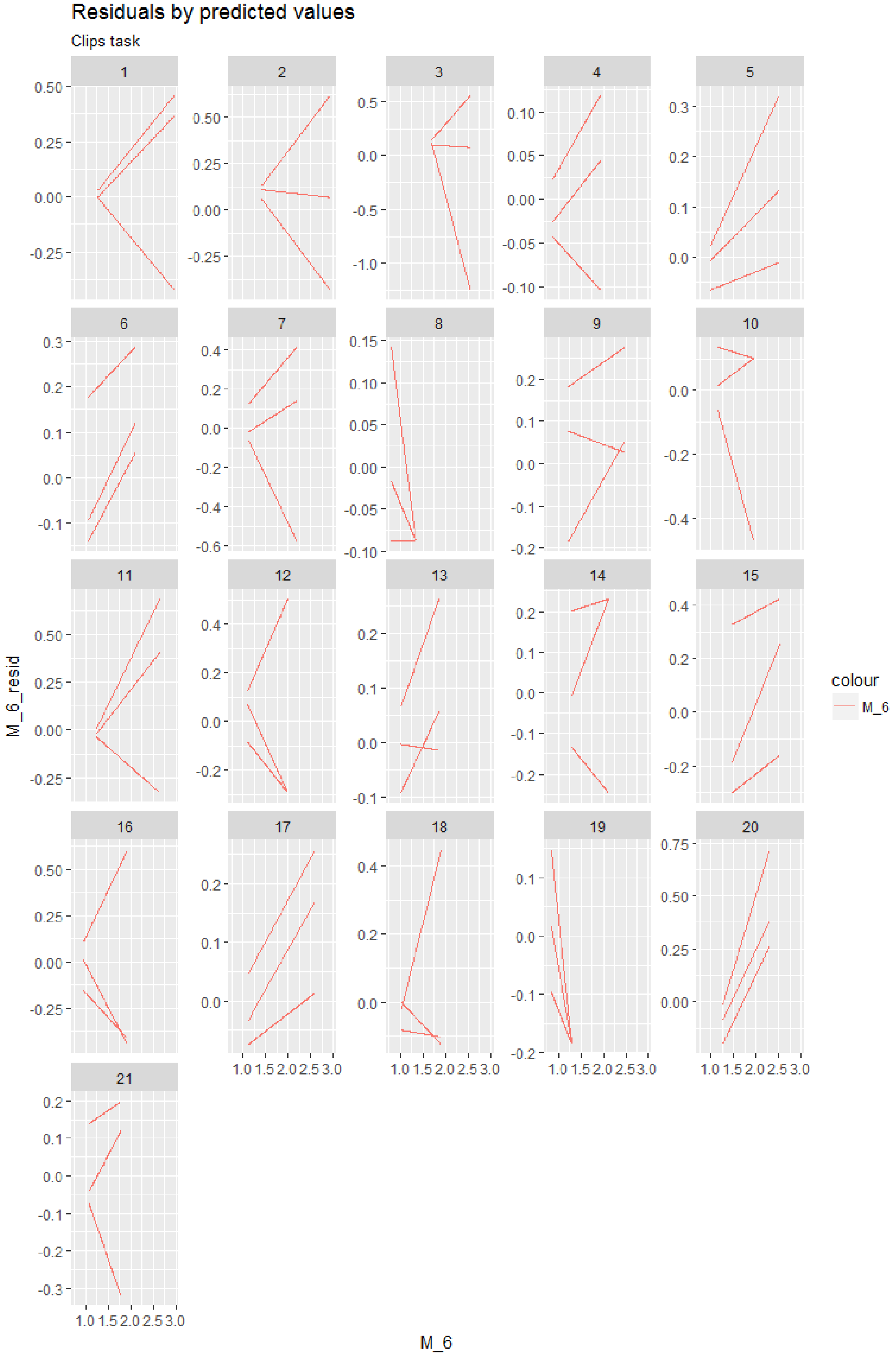
Task: Drawing

```
T_PP_6 %>%
  filter(Task == "Drawing") %>%
  ggplot(aes(x = trialS)) +
  geom_line(aes(y = center)) +
  geom_ribbon(aes(ymin = lower, ymax = upper), alpha = .3) +
  geom_point(aes(y = ToT)) +
  facet_wrap(~ Part, nrow = 7, scales = "free_y")
```

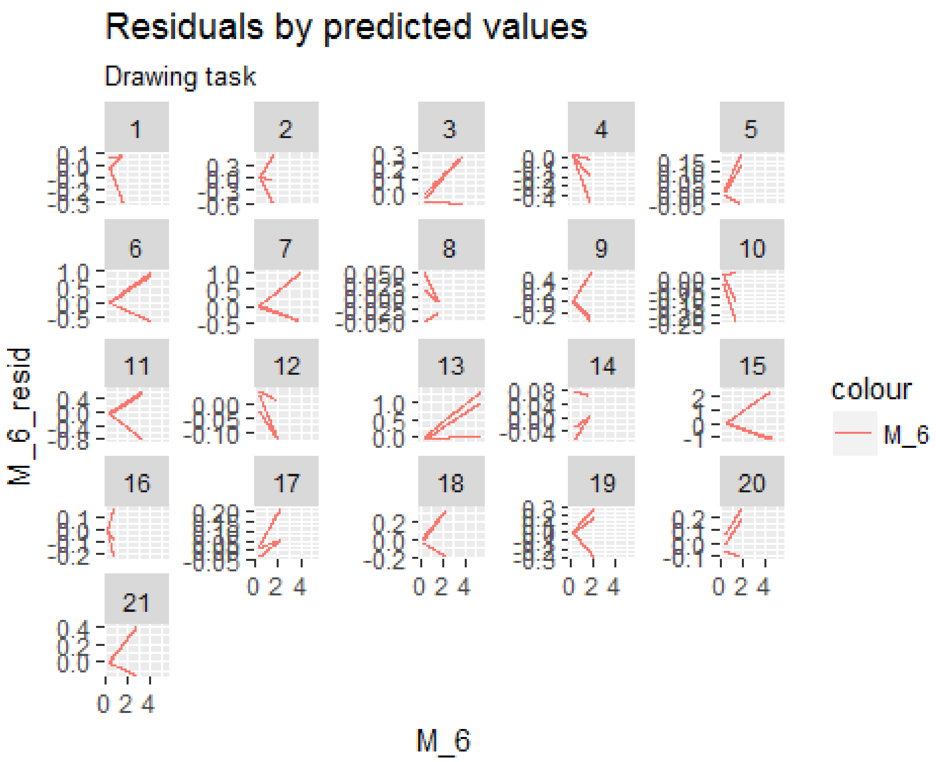

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
center
trialS

Residuals analysis

```
LD_18 %>%
  filter(Task == "Clips") %>%
  ggplot() +
  geom_quantile(aes(x = M_6, y = M_6_resid, col = "M_6"), formula = y ~ x)
+
  facet_wrap(~ Part, scales = "free_y") +
  ggtitle(list(title = "Residuals by predicted values",
               subtitle = "Clips task"))
```


Residuals by predicted values
Clips task
M_6_resid
M_6
colour
M_6

```
LD_18 %>%
  filter(Task == "Drawing") %>%
  ggplot() +
  geom_quantile(aes(x = M_6, y = M_6_resid, col = "M_6"), formula = y ~ x)
+
  facet_wrap(~ Part, scales = "free_y") +
  ggtitle(list(title = "Residuals by predicted values",
               subtitle = "Drawing task"))
```



Model comparison

Do the two models differ in point estimates? Most estimates are close. Only rateM-TaskClips is extremely high at M_5. Such a strong rate is virtually a one-step learning. As this is very unlikely for MSL, the ARARY model seems to fail and we prefer M_6.

```
bind_rows(
  fixef(P_5),
  fixef(P_6, mean.func = exp)
)

##
##
## Table: Estimates with % credibility limits
##
## model    nonlin    fixef              center       lower        upper
## ------   -------   ------------   -----------   ----------   -----------
## M_5      asym      TaskClips        1.1577182    1.0291061    1.2638820
```

```
## M_5      asym     TaskDrawing     0.2612140   0.1180004    0.3269353
## M_5      rateS    TaskClips       0.3220994   0.2141153    0.5059187
## M_5      rateS    TaskDrawing     0.5377853   0.3907563    0.7115918
## M_5      amplS    TaskClips       1.0286249   0.7859518    1.3076140
## M_5      amplS    TaskDrawing     2.0981479   1.5331471    2.6884107
## M_5      rateM    TaskClips      10.1972400   0.4392050   26.3571275
## M_5      rateM    TaskDrawing     0.1020007   0.0407116    0.1833435
## M_5      amplM    TaskClips       0.1054591   0.0395542    0.1794875
## M_5      amplM    TaskDrawing     0.1866054   0.0900881    0.3412478
## M_6      asym     TaskClips       1.0945838   0.0024140    1.2260787
## M_6      asym     TaskDrawing     0.2007313   0.0000117    0.3163152
## M_6      rateS    TaskClips       0.2360948   0.1478205    0.4222025
## M_6      rateS    TaskDrawing     0.4660412   0.3463548    0.6180551
## M_6      amplS    TaskClips       0.9489550   0.7166955    1.2145355
## M_6      amplS    TaskDrawing     1.8077051   1.3454203    2.3860734
## M_6      rateM    TaskClips       0.3400725   0.0080022   17.3677221
## M_6      rateM    TaskDrawing     0.0610973   0.0200405    0.1674628
## M_6      amplM    TaskClips       0.0462354   0.0000000    1.1707671
## M_6      amplM    TaskDrawing     0.2087467   0.0701147    0.4609479
```