Two stage analysis of learning curves on laparoscopic study of surgeons

Sintayehu Aynalem

promotor :
De heer Abel TILAHUN ESHETE

Eindverhandeling voorgedragen tot het bekomen van de graad
Master of science in Applied Statistics

universiteit
hasselt

Abstract

The performance of many repeated tasks change with experience, with improvement being most rapid at first and then tails off over time until a steady state is reached. The term ‘learning curve’ is often used as shorthand to describe this phenomenon. The objectives of this study were to evaluate individual learning curves for surgeons performing laparoscopic activities for two different tasks important for a surgeon (the nuts and the ropes) that are timed and also to evaluate the possibility of using the information from psychological test and gender in predicting surgical performance. Moreover, it was of interest to know if the measures of performance from the learning curve, psychological test and gender can be used for predicting the results during the 'real life test’. Two-stage analysis was implemented to achieve the stated objectives. In the first stage two approaches were used: the simple method which summarises individual measures by initial measurement, the difference between the first and last measurement or just the final measurement, and the complex analysis. For the complex method, the exponential curve was used to derive the measures of performance of a surgeon from which proxies for learning such as initial level, length of learning, final skill level, rate of learning and time taken to reach a plateau were derived. In the second stage, the proxies for learning were regressed against several covariates. When we look at the training ropes, females performed better than males according to most occasions of measures of performance. Surgeons with higher special cognitive ability were found to have lower starting level. It is also noted that surgeons with larger motivation perform better as compared to those with lower motivation.

Key words: *learning curve, laparoscopic surgery, plateau*

Acknowledgements

This thesis would never have been started without the help and stimulation of a number of persons. Many thanks go to Prof Dr Marc Aerts and Prof Dr Herbert Thijs for providing me with this project. I am very grateful to my internal supervisor Drs Abel Tilahun and external supervisor Dr Steffen Fieuws for giving me their unlimited supports and very helpful and important remarks on this paper. This experience has brought me in closer contact with research and data analysis. Thank you for taking time from your extremely busy schedule to help me. Thank you so much. You gave me hope. And I just continued the best path of my life...needless to know I'll be out of this soon! God bless your wonderful deeds super-abundantly in the precious and powerful name of Jesus Christ our Lord and Saviour.

Table of contents

1. Introduction

Laparoscopic abdominal surgery is a modern surgical technique which requires adaptive skills to overcome the difficulties of working with an indirect view of the workspace through the endoscope camera and the remote manipulation of tools. Laparoscopic abdominal surgery involves some advantages such as less pain, short hospital stay and satisfactory exploration of the abdominal space since it was the first time described by Seem (1983). As experience is a key factor of proficiency at laparoscopy, surgical simulators have received attention as learning devices. One approach for continuing development of training simulators is examining surgeons' cognitive processes and skills. The introduction of laparoscopic techniques to abdominal surgery was associated with many complications, which led to the development of skills laboratories to train laparoscopic surgeons.

The structured curriculum can enable trainees to be confident in their skills prior to assisting in and performing the initial laparoscopic procedures, safe in the knowledge that they have achieved preset expert criteria. The ultimate aim is to reduce their time on real patients, leading to acquisition of proficiency at an earlier stage than training on patients alone. This may lead to a reduction in the number of unnecessary complications occurring due to a failure of technical skills, and the time and expense spent acquiring basic laparoscopic skills in the operating room. Identifying these skills can provide a basis of how to approach training. To develop an evidence-based virtual reality laparoscopic training curriculum for novice laparoscopic surgeons to achieve a proficient level of skill prior to participating in live cases.

It is a fundamental human characteristic that a person engaged in a repetitive task will improve his performance over time. Learning curves have long been recognised outside health technology assessment, in psychology, manufacturing and aviation. If data are gathered on this phenomenon, a curve representing a decrease in effort per unit for repetitive operations can be captured by a learning curve. A learning curve is a well known tool which describes the relation between the performance of a task and the number of repetitions of that task. A learning curve model may be used as a prediction tool in various applications of operations planning and control. The learning curve effect states that the more times a task are performed, the less time will be required on subsequent iteration. This is not surprising because we have all seen this and perhaps know it in some intuitive sense. What is interesting is that the rate and shape of the learning curve of improvement fairly changes across repetitive operations. The pattern is a rapid improvement followed by ever lesser improvements with further practice. Occurring concurrently is a decrease in variance in performance as the behaviour reaches an apparent plateau.

A change over time in the performance of a technology because of learning complicates evaluation and impedes rigorous evaluation. Variables that are good proxies for learning need to be identified. Useful parameters for describing skill acquisition in laparoscopic assessment are the initial level, length of learning, final skill level and rate of learning. Methods for estimation of the time taken to reach an asymptote (plateau) should be explored further. Typically, these will be some measures of performance (i.e., behaviour when learning is disabled), but other metrics, including characteristics when learning is disabled, are also legitimate. However, as Provost, Fawcett, and Kohavi (1998) have argued, it is important that these variables make direct contact with the goals of the research.

In this paper, learning curve analysis was proposed and used to answer three main questions relevant to the field of intelligent performance systems. These are stated as follows:

1. How can we describe learning behaviour for two different basic trainings in terms of an existing cognitive model? The aim of this project was to determine whether there were existing methods that had been, or could be, adapted to the purpose of allowing for ‘the learning curve’. Thus we need to find the learning curve for each surgeon and identify the initial difficulty level of each training and how fast can a surgeon learn each practice (i.e., who is the fast learner? what is the learning rate or slope?). We can then provide parameters that indicate as the surgeon behaviour reaches an apparent plateau. Identifying the point at which the curve flattens (the asymptote) would allow subsequent evaluation free of any learning effect.

2. Can we use the information from results of psychological test to predict indicators of learning performances? Moreover, can we use the information from results of psychological test and indicators of learning performance to predict the results obtained during the 'real life test’ which involved stitching on a pig?

3. Is there any method which quantifies learning behaviour that outperforms simple approaches such as just using the first time or last time to summarise acquisition behaviour?

2. Study overview and Datasets

The study was carried out in Gasthuisberg hospital in 2006 involving 27surgical residents. Of the total surgeons 66.7% were male (18) and 33.3%were females (9). All the participants performed 30 repetitions of two tasks on the Invasive Surgical Trainer namely nuts and ropes. It was assumed that these two trainings are important practices in the surgical activities that a surgeon has to accomplish. Each attempt was timed (in seconds) for both trainings. In addition to the two tasks, at the beginning of the study, each surgeon was measured for their motivation (motivatie) and specific cognitive ability (schlauch) and both are continuous variables. These are psychological tests.

After the previous tests the surgeons also performed an operation on a pig. This is the 'real life test' which involved stitching on a pig. The results obtained during the exam in the 'real life test' are quantified by five indices: two subjective assessments and three objective assessments. The subjective assessments are scores on the exam. The three performance indices, during objective assessment, are obtained by placing sensors on the hands of the surgeon. Variables such as time taken, number of movement, and path distance are measured.

The variables in the data set were as follows;

- *time*: the time(in seconds) taken to complete the tasks in the thirty trials by each surgeon
- *motivatie*: measure for motivation (the higher the value the better the motivation)
- *schlauch*: result of a visual-spatial test for each performed task(the higher the better)
- *gender*: the gender of the surgeon (coded 1 for male and 0 for female)
- *score1*:mesure of the first subjective assessment (score up to 10, the higher the better)
- *score2*:mesure of the second subjective assessment(score up to 50,the higher the better)
- *mo*: the number of movements made while the operation is performed on the pig
- *path*: the total distance covered while the operation is performed on the pig
- *otime*: operation time which is the time taken from first incision to last stitch on the pig

3. Methodology

Exploratory data analysis was conducted in order to capture the structure and get insight into the data. Individual profiles were used to explore the data set and see how the profile of an individual evolves over repetition. They also suggest the variability seen within the data, and may provide information whether linear or non-linear models are plausible. In addition to individual, data exploration was also performed using correlation matrix of the variables to investigate the relationship among them. Two stage analyses, involves fitting a model for each subject separately the first stage while the second stage explains the variability in the subject-specific regression coefficients using known covariates.

3.1 Two-Stage Analysis

As discussed in the previous sections, this study involved 27 surgeons performing a simulated task on 30 consecutive trials with one of the objectives being characterizing the performance of the surgeons. To this end, a two stage analysis is thought to be an appropriate modelling tool. Two approaches are to be followed to summarize the thirty repetitions of each trial by a set of summary statistics which serve as measures of performance for each surgeon separately. These are the simple approach and the complex approach.

The first possibility which is termed as simple methods involved summarizing the thirty measurements by either the final measurement or the increment. The second stage explains the variability in the subject-specific regression coefficients using known covariates.

Complex curves can be used to quantify the learning behaviour of surgeons separately. In the two-stage analysis, the first stage involves fitting a non-linear regression model for each subject separately while the second stage explains the variability in the subject-specific regression coefficients using known covariates.

3.1.1 First Stage summary measures

As stated above, there are two approaches that are discussed in this project to quantify learning behaviour of the surgeons which are termed here as: Simple methods and complex methods. These will be in the following sections.

3.1.1.1 Simple methods

As has been noted in the data description, every surgeon had trials recorded for thirty times. One possible way to assess the competence of the surgeons is to condense the measurements from the thirty trials into few summary statistics that are indicators of performance. Three simple summary measures were considered here, namely, the initial measure, the last observed time and the increment i.e. difference between the last and the first observed times. These analyses were carried out to see if there is an easier way to quantify measures of performance of a surgeon using methods that are effective and simpler to use. The basic advantage of these kinds of analyses is that they are computationally simple.

3.1.1.2 Complex methods

The above simple methods entail some draw backs in the form of loss of information and the fact that they may not give clues on the learning process prior to the last repetition of the task. If for example we consider the increment, which is the difference between the first and final performance, we might be able to compare evolutions between surgeons after correcting for baseline differences but miss the process of learning the surgeons have gone through and lose information that have been collected in the intermediate tasks. These draw backs are even more pronounced if we consider the last observed time only, as we can not determine the rate of learning of the surgeons and also lose a great deal of information by ignoring previous measurements. These measures also aid in quantifying most of the necessary properties of learning process such as the initial level, length of learning (the difference between the first and the last), the final level (the level at plateau) and the rate of learning by using three parameters obtained from the model. One way to circumvent these draw backs is to adopt methods that use the whole data and yet are able to give effective summary measures that can be used to evaluate the performance of surgeons. One such method is the learning curve analysis which is the concern of the next section. It is also the project's aim to verify if the learning curve methods (complex methods) indeed outperform the simple methods.

3.1.1.2.1 Learning curves

Learning curves depict the performance of the surgeons with respect to some measure of their ability over time. Time is generally represented by the number of occasions the knowledge element has been used. Learning curves describe the evolution of performance over trials *t*. They can be modelled by the following equation:

$$f(t) = a + bg(t)$$

where *a* is the asymptote of the curve which is also the plateau. The function *g* (*t*) describes the type of curvature present in the learning curve. As such, *g* (*t*) is called the core of the learning curve and is often a function of a third parameter, the learning rate parameter *c* (Paul, 1994). It is not advocated that learning curve analyses involve complex function, rather that they should employ the simplest functions that can answer the questions being posed. The functions used should be parsimonious; that is, they should not use more parameters than are necessary. The most commonly used learning curves are the power curve and the exponential curve. After the fit of the either of the above models (the power or the exponential curve), the proxies for learning describing learning in laparoscopic assessment are then taken as: the rate of learning, the final level, length of learning and the initial level. The following section gives a brief introduction of these two learning curves.

3.1.1.2.1.1 The power curve

When performance is measured by response times, many authors suggested the use of the power curve (Newell and Rosenbloom, 1981, Logan, 1988). If the student is learning the knowledge elements being measured, the learning curve will follow a so-called "power law of practise". Evidence of such a curve indicates that the student is learning the knowledge elements, or, conversely, that the elements represent what the student is learning. Therefore, when comparing two models we might argue that the model fitted with power law is somehow superior. Its core function is given by $\boldsymbol{g(t) = t^{-c}}$.

In order to summarize the measurements from the thirty trials by each surgeon the following model will be used.

$$Y_{ij} = a_i + b_i t_{ij}^{-c_i} + \varepsilon_{ij}; \quad i = 1,\ldots,27 \; and \; j = 1,\ldots,30$$

where Y_{ij} is the number of seconds taken to perform the training in repetition j for surgeon i, t_{ij} is the repetition at j^{th} trial for surgeon i and $a_i, b_i \; and \; c_i$ are unknown surgeon-specific parameters. a_i indicates asymptote or the plateau time for surgeon i ; c_i indicates the rate of learning for surgeon i.

3.1.1.2.1.2 The exponential curve

Heathcote raised some concerns over the recent years in the use of the power curve and suggested the use of the exponential curve, given by the core ***g (t) = e**$^{-c\,t}$*, instead. (Heathcote, Brown and Mewhort, 2000).

If a choice is made to use the exponential curve, summary measures for each surgeon are derived in the same manner from the fit of the following model for each surgeon.

$$Y_{ij} = a_i + b_i \exp(-c_i t_{ij}) + \varepsilon_{ij}; \quad i = 1,\ldots,27 \; and \; j = 1,\ldots,30$$

where Y_{ij} is the number of seconds taken to perform the training in repetition j for surgeon i, t_{ij} is the repetition at j^{th} trial for surgeon i and $a_i, b_i \; and \; c_i$ are unknown surgeon-specific parameters. a_i indicates asymptote or the plateau time for surgeon i ; c_i indicates the rate of learning for surgeon i.

3.1.2 Second Stage Model building

In the second stage, a multiple linear regression model is fitted using the summary statistics obtained in the first stage as responses and variables such as motivatie, schlauch and gender as covariates. In addition, the parameter estimates from the first stage of the two-stage analysis can be used as predictors for models whose responses are from the exam on the 'real life test' such as: score1, score2, number of movements made (mo), path distance taken (path) and time taken in exam (otime).

The general formulation of the second stage, taking performance measures as response, is as follows:

$$S_i = K_i\beta + \varepsilon_{1i}$$

where, S_i is a subject-specific parameter estimate from the first stage. S_i is one of the four measurement for performance i.e. initial level, length of learning, rate of learning and final level. K_i is a row vector of the intercept , motivatie, schlauch and gender. β is a vector of regression parameters for the four predictors including the intercept. ε_{1i} explain the observed variability between the surgeons, with respect to the performance measures.

Another interest also lies on assessing if the parameter estimates from the non-linear regression models can be used as predictors for models whose responses are from the exam on the 'real life test' such as: score1, score2, number of movements made (mo), path distance taken (path) and time taken in exam (otime).

The general formulation of the second stage, taking performance measures as covariates, is as follows:

$$R_i = S_i\gamma + \varepsilon_{2i}$$

where, R_i is one of the variables from the exam on the 'real life test' such as: score1, score2, number of movements made (mo), path distance taken (path) and time taken in exam (otime). S_i is a row vector of parameter estimate from the first stage i.e. intercept, initial level, length of learning, rate of learning and final level. γ is a vector of regression parameters for the predictors including the intercept. ε_{2i} explain the observed variability between the surgeons, with respect to their performance in the 'real life test'.

Subsets of predictor variables to be included in the model were chosen using a multiple linear regression where Cp and adjusted r-square criterion used to select important subset of variables. All possible important two way interactions were included for the predictor variables. The variables in the model were then checked for significance using the backward selection method after the variables are checked for the existence of multicollinearity which is tested using the variance inflation factor (VIF). A multiple linear regression model has some fundamental assumptions which include normality and constancy of variance of the error terms that need to be tested before the model qualifies for prediction or inference. Formal test for normality was done using the Shapiro-Wilk test and informally the plot of residuals was done. The plot of observed versus predicted values is used to assess if the linearity assumption holds. To test the constancy of variance the modified levene’s test was performed. The leverage values were used to detect outliers with respect to the predictor variables. Influence analyses in case of outliers were carried out by calculating DFFITS and DFBETAS.

All statistical analyses are carried out using SAS 9.1 and S-plus 6.2 and most of the time the significant level is taken at 5% level.

4. Results

4.1 Data exploration

It can be seen from the individual profiles in figures 4.1a and 4.1b that most of the identified curves have similar basic shape that decreases to an asymptote in both of the tasks. This demonstrates a decrease in operation time and a kind of horizontal pattern after some iteration of the training in both tasks. The most widely cited shape of learning curve across all fields such as power law and the exponential curve can be appropriate to fit the data.

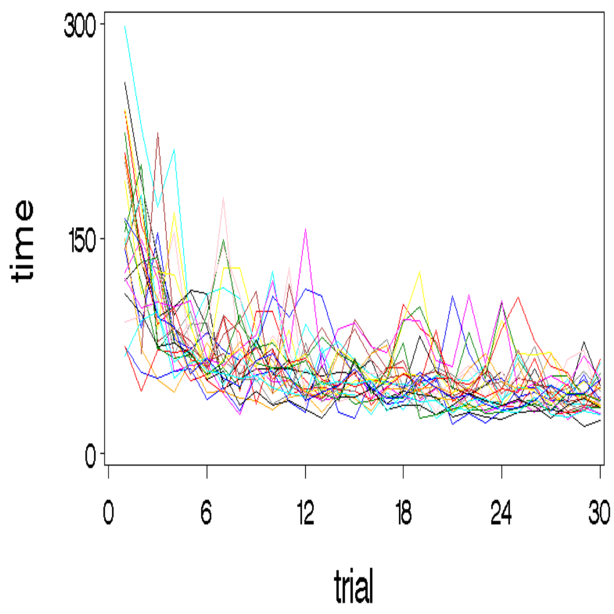


Figure 4.1a: *Individual profile for training nuts*

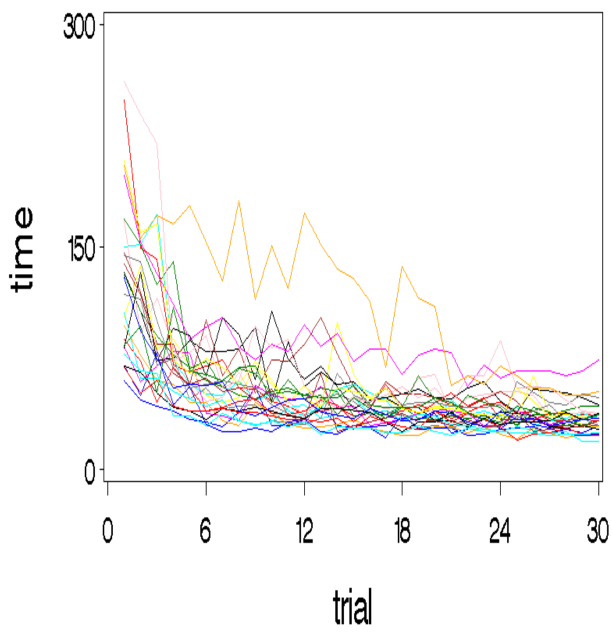


Figure 4.1b: *Individual profile for training ropes*

4.2 Statistical analysis

4.2.1 Two-stage analysis with simple methods in first stage

Three summary measures, the last observed time (final) and the difference between the last and the first observed times (length) and the starting value (initial) were used to quantify performance of each surgeon.

The three measures of performance (initial, length and final) were regressed on the covariates using multiple linear regressions. The various covariates, including every possible interaction term, were included in the model.

Analysis of the relationships between the measures of performance from the simple method taken as predictors and the results obtained during the exam (the 'real life test' which involved stitching on a pig) taken as responses was done.

4.2.1.1 Summary measures as response for ropes

For the training nuts, it has been noted that none of the covariates are statistically significantly related with all the three responses. The scatter plot of the responses with the covariates also doesn't show a pattern of relationships.

In the ropes training, for the responses, initial level and length, gender and schlauch were found to be statistically significant. The corresponding adj R-square values were 0.4620 and 0.5018 respectively for initial level and length. The parameter estimates and p-values are shown in tables 4.2.2 and 4.2.5.

Table 4.2.2: *Parameter estimates of the response initial*

Parameter	Estimate	P-value
intercept	256.81	<0.0001
gender	-56.63	0.0037
schlauch	-6.27	0.0176

0=female,1=male

All assumptions of the models were checked and found to be fulfilled. The plot of the observed versus the predicted value for both models is shown in figure 4.2.1 and 4.2.2.From the plots it seems reasonable to assume the linear line.

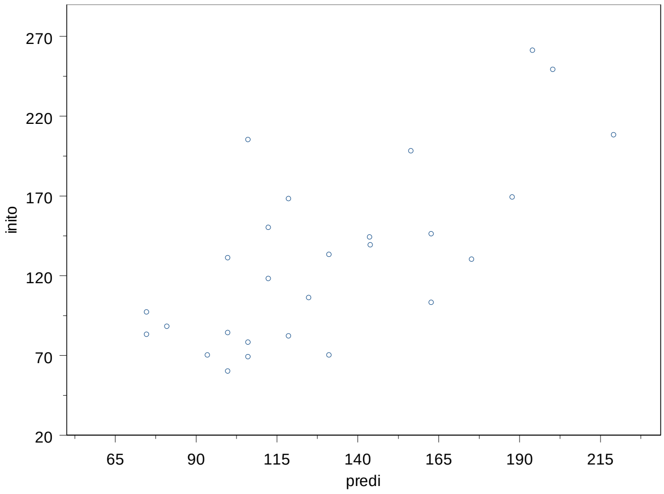


Figure 4.2.1: *The observed versus predicted value plot for initial*

The negative sign for gender in table 4.2.2 indicates that females performed better than males. In addition, the negative sign of the schlauch indicates; the higher special cognitive ability a surgeon has, the lower his/her starting level will become. Therefore, surgeons who have large special cognitive ability have good performance; as good performance mean smaller initial time.

Table 4.2.3: *Parameter estimates of the response length*

Parameter	Estimate	P-value
intercept	218.97	<0.0001
gender	-51.86	0.0028
schlauch	-6.29	0.0080

0=female, 1=male

As can also be seen from the table 4.2.3, the negative sign of the schlauch indicates that the higher special cognitive ability a surgeon has, the lower his length of learning becomes. This may be because of the fact that the surgeons with smaller length have lower starting level since initial level and length have high correlation as shown in table 4.2.1. This in turn means the surgeon ends gaining small change (learning) after his initial time.

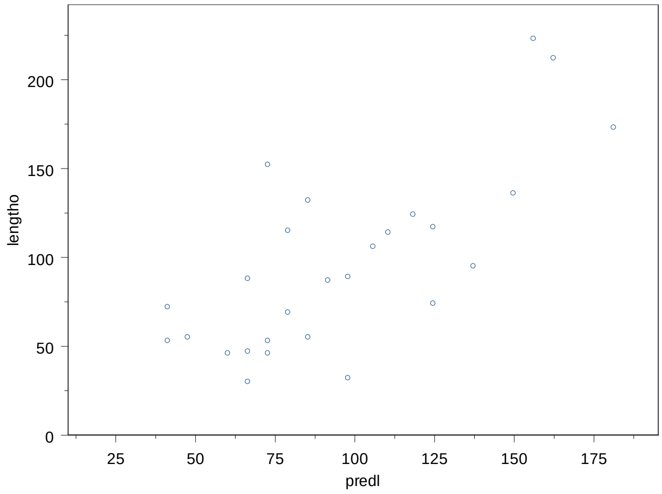


Figure 4.2.2: *The observed versus predicted value plot for length*

In attempting to assess the relationships between the summary statistic final level of the surgeons and variables such as motivatie, schlauch and gender, it was discovered from the scatter plot that there seems no relationship between the response and the covariates. In addition, a multiple linear regression model was fitted for this purpose. This also shows insignificant linear relationship.

4.2.1.2 Summary measures as predictor for ropes

From analysis of the relationships between the measures of performance from the simple method and the results obtained during the exam (the 'real life test' which involved stitching on a pig), it is noted that the variables have no relationships.

For ropes training, the covariates initial observation and length are significantly related with the response otime (operation time taken from first incision to last stitch on the pig) at 10%. The result of the model parameter estimates is displayed in the Table 4.2.4.

Table 4.2.4: *Parameter estimates of the response otime*

Parameter	Estimate	P-value
intercept	141.95	0.0579
initial	4.68	0.0266
length	-5.01	0.0289

The F-value obtained is 2.79with (p= 0.0814) with adjusted R-square 0.1210. This indicates measures of learning behaviour from the simple approach are not strongly related with the variables that indicate the surgeons' skill in the 'real life test'. The plot in figure 4.2.3 shows that plot of observed versus predicted values. There seems to be a quite but not good match between the observed and predicted values. Assuming the linearity assumption is not too bad to call for remedial measures, we can conclude that the positive coefficient for gender indicates surgeons with large baseline differences will have large time to complete the task in ‘real life test’ because other assumptions are found fulfilled.

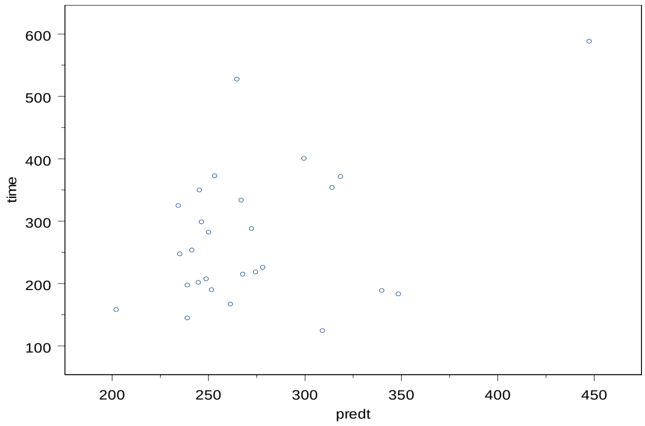


Figure 4.2.3: *The observed versus predicted value plot for otime*

4.2.2 Two-stage analysis with complex methods in first stage

In the “complex methods” section discussed above, the summary measures from the fit of the three-parameter non-linear regression model were used to quantify the performance of each surgeon. These values are expected to be superior to the summary measures obtained from the simpler methods in that they are based on all measurements as opposed to later which summarizes the whole thirty measurements for a surgeon just by taking the first, last or the difference between the last and first.

The power curve model and the exponential model are used for the two training data sets. Both curves are very close in fitting the observed measures in the time to complete the tasks in the two trainings. Within each surgeon there was no apparent difference between the statistical fits of the two models however the exponential model seems to perform slightly better. In addition, the sum of absolute residuals for the exponential is also smaller than the power curve model for both trainings as shown in the appendix figreA5. Thus for the remaining analysis we consider the exponential curve analysis.

All the exponential curve fits in the non-linear regression model were statistically significant at 1% level for each of the surgeons. This was tested using the statistical difference between a full model containing all the three parameters and a reduced model with only intercept.

Table 4.2.5: *Measures of performance from the exponential curve for the training nuts*

Surgeon	Estimates of measures of performance			
	Initial level	Rate of learning(c)	Final level (plateau or a)	Length of learning
1	256.6	0.4029	52.2	204.3
2	242.8	0.6944	65.4	177.4
3	150.3	0.3075	38.4	111.9
4	115.3	0.0354	1E-08	115.3
5	280.1	0.2164	46.2	233.9
6	146.5	0.2905	36.3	110.2
7	192.5	0.3572	49.8	142.7
8	182.6	0.2150	48.4	134.3
9	149.3	1.1191	40.4	108.8
10	160.1	0.3140	48.6	111. 5
11	222.4	0.3287	62.9	159.5
12	129.0	0.1032	18.9	110.1
13	198.6	0.3871	44.9	153.7
14	211.9	0.4864	69.4	142.4
15	69.6	0.2295	38.1	31.5
16	167.1	0.3564	42.5	124.6
17	107.0	0.0191	1E-08	107.0
18	325.2	0.5806	48.8	276.4
19	119.9	0.0630	17.6	102.3
20	203.9	0.5443	53.1	150.9
21	330.2	0.9655	65.9	264.2
22	169.6	0.2348	44.8	124.8
23	114.6	0.2662	33.3	81.3
24	66.5	0.0248	22.9	43.6
25	169.1	0.2821	40.7	128.4
26	168.2	0.3613	43.5	124.7
27	85.1	0.0784	22.1	62.9

For each of the surgeons in the training nuts, the three parameter estimates for the exponential model are shown in Table 4.2.5. The fitted curves for the training nuts suggested that the starting values for the surgeons ranges from 66.5 seconds (surgeon 24) to 325.2 seconds (surgeon 18) and the rates of learning range from 0.019(surgeon 17) to 1.12(surgeon 9). This means surgeon 9 is the fastest learner for this training. A plateau time for the operating time was reached rapidly, at around 69.4 seconds for surgeon 14. There were minor complications in this training, two of the surgeons (surgeon 4 and 17) happened to have a plateau time very close to zero. It is not realistic to consider a value zero for the time it takes to complete a specific job. But this was because the two surgeons perform this training with performance that kept on declining until the end of the repetition of the training. This indicates these surgeons never reached their plateau level during the thirty repetitions. This was also seen in the plot of observed and predicted time to complete the task for these two specific surgeons in appendix FigureA6.

If we focus now on the ropes training, we can learn that all the exponential curve fits in the non-linear regression model were statistically significant at 1% level for each of the surgeons. This was tasted using the difference between full and reduced model as explained above for the nuts.

Table 4.2.6: *Measures of performance from the exponential curve for the training rope*

Surgeon	Estimates of measures of performance			
	Initial level	Rate of learning(c)	Final level (plateau or a)	Length of learning
1	114.6	0.1372	47.2	67.4
2	241.3	0.4511	39.9	201.3
3	127.5	0.2538	38.4	89.0
4	52.4	0.1519	26.8	25.6
5	172.7	0.2948	38.2	134.5
6	78.3	0.1257	31.6	46.7
7	126.3	0.3925	39.0	87.2
8	161.1	0.4599	45.3	115.8
9	94.3	0.2802	26.1	68.1
10	113.7	0.0601	17.2	96.5
11	116.6	0.2040	41.8	74.9
12	107.3	0.0482	1E-08	107.3
13	86.4	0.3975	40.7	45.7
14	176.1	0.2419	40.4	135.8
15	73.1	0.2925	28.5	44.6
16	97.6	0.3873	32.1	65.5
17	192.7	0.3508	73.5	119.2
18	153.7	0.3602	41.8	111.9
19	283.2	0.3743	53.5	229.7
20	197.1	0.0416	1E-08	197.1
21	135.6	0.2859	32.9	102.6
22	210.6	0.3353	43.1	167.4
23	72.1	0.3924	34.9	37.1
24	63.9	0.1829	28.3	35.5
25	90.5	0.0933	30.1	60.4
26	125.0	0.3691	36.3	88.8
27	80.3	0.3322	25.9	54.4

From Table 4.2.6, we can see that for the training ropes the starting values of the surgeons ranges from 52.36 seconds (surgeon 4) to 283.2 seconds (surgeon 19) and the rates of learning range from 0.0420(surgeon 20) to 0.4598(surgeon 8). This means surgeon 8 is the fastest learner for this training. Surgeon 17 was very rapid to reach plateau. Here also there are two surgeons (surgeon 12 and 20) whose time to reach plateau was very close to zero. This is also clearly seen in the plots of the observed and predicted time to complete the task for these two specific surgeons in appendix FigureB6.

Before proceeding to the second stage modelling, the Pearson correlation was used to describe the degree of correlation between the complex model summary measures from the first stage and set of variables from the psychological test and gender. In addition, the relation between the summary measures from the complex methods of the first stage and the five variables from the exams on stitching in the pig are explored using the Pearson correlation for both trainings. These are given in appendix A TablesA1 and A2 for the training nuts and appendix B Tables B1 and B2 for the training ropes.

4.2.2.1 Summary measures as response for training nuts

Let's now consider the nuts training. In the second stage of the modelling process, the effects of the various covariates (motivatie, schlauch and gender), including their possible interaction terms, were fitted on the responses (initial, length, rate and final performance) using multiple linear regression. It is learned that all the three predictors are found to be insignificantly related with initial level and length of learning.

For the case of rate of learning, the F-value obtained is 3.01 with (p= 0.03) which shows the covariates in the model are important in explaining the variation observed in the response variable (rate). The adjusted R-square for this model is found to be 0.2792. However, the diagnostic checking for this model show that the Shapiro-Wilk test was significant (p= 0.0239) showing that the residuals are not normally distributed. The plot of residuals is also highly positively skewed. In assessing the constancy of variance the modified levene's test shows that the variance was not homoscedastic (p= 0.0390). The plot of observed versus predicted values is assessed .This plot shows there is a remarkable deviation from linearity.

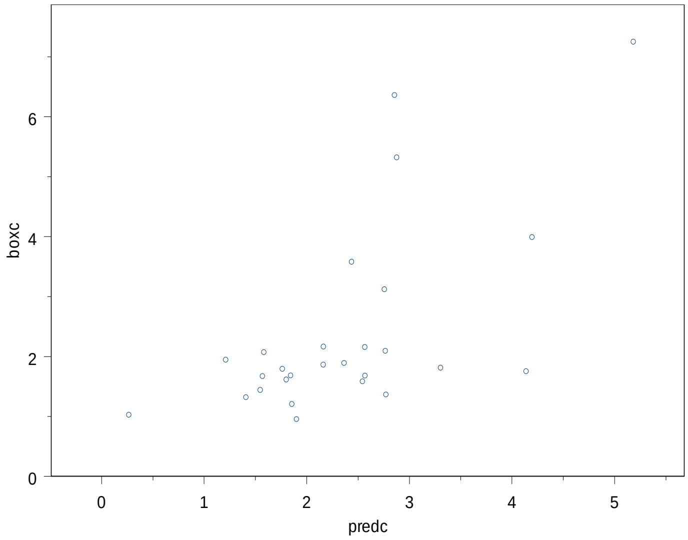


Figure 4.2.4: *The observed versus predicted value plot for* $\frac{1}{sqrt(rate)}$

To resolve these assumptional problems the Box-Cox transformation was used. This suggests the inverse square root transformation i.e. $\frac{1}{sqrt(rate)}$. After this transformation the model is checked again for the diagnostics taking the transformed rate as a response. It was learned that the formally tested assumptions were found to be satisfied. The Shapiro-Wilk test for normality was insignificant (p=0.0559). Besides this, the plot of residuals shows fairly better non-skewed distribution as shown in the appendix figureA1. The modified levene's test also shows a constant variance (p=0.1795).The plot of observed versus predicted values shows there is no remarkable deviation from linearity which shows the linearity assumption seems satisfied as shown in the figure 4.2.4. There were Outlier observations. However, since, they are not influential there will be no danger from this observations. The parameter estimates after transformation is given in Table 4.2.7.

Table 4.2.7: *Parameter estimates of response* $\frac{1}{sqrt(rate)}$

Parameter	**Estimate**	**P-value**
intercept	3.01624	<.0001
$schlauch_o$	0.15165	0.1178
$motivatie_o$	-0.29722	0.0014
gender	-0.42331	0.4850
$motivatie_o$*$schlauch_o$	-0.04375	0.0191
$motivatie_o$*gender	0.29756	0.0095

0=female, 1=male   subscript o indicates deviation from the mean is taken

The model is spelled out as follows:

$$\frac{1}{sqrt(rate)} = 3.01624 + 0.15165 * \mathbf{schlauch_o} - 0.29722 * \mathbf{motivatie_o} - 0.42331 * \mathbf{gender}$$

$$-0.04375 * \mathbf{motivatie_o} * \mathbf{schlauch_o} + 0.29756 * \mathbf{motivatie_o * gender}$$

The transformed model is significant with p= 0.0418 and adjusted R-square= 0.2602. Also, the p-values for the covariates Motivatie$_o$, motivatie$_o$*schlauch$_o$, and motivatie$_o$*gender is less than 5%.Since the response is taken as an inverse square root, parameter estimates should be interpreted carefully and also care must be taken in interpreting these results with regard to the significant interaction terms and the fact that covariates motivatie and schlauch are centred. Centering the continuous variables is needed because they would yield an estimate or p-value which has a relevant meaning. As a result, the small p-value (p=0.0080) of motivatie and the interactions and the negative sign for the coefficient of motivatie indicate that, a female with an average schlauch that has lower than average motivation will have a lower response ($\frac{1}{sqrt(rate)}$). After retransforming back to rate, consequently, females with higher than average motivational score and average schlauch will have a higher rate of learning. Meanwhile, this relationship is not adequate enough to make predictions, since the adjusted R-square is very small.

For the case of final performance, the results of the model are shown in Table 4.2.8. The adjusted R-square for the model is found to be 0.2747.

Table 4.2.8: *Parameter estimates of the response final performance*

Parameter	**Estimate**	**P-value**
intercept	46.20328	<.0001
motivatie$_o$	2.3110	0.0046
gender$_o$	-8.62477	0.1775
motivatie$_o$*gender	-2.93860	0.0064

0=female, 1=male subscript o indicates deviation from the mean is taken

The F-value obtained is 4.28 with (p= 0.0153) which shows the covariates in the model are important in explaining the variation observed in the response variable, final level.

The fitted model was checked for validity of the model. The test for normality using Shapiro-Wilk test shows non-normal distribution for the residuals (p=0.0116). But, the plot of residuals shows a peaked but not extremely skewed distribution as seen in the appendix figureA2. It is assumed that this deviation is not too dangerous to call for remedial measures. The modified levene's test shows there is no problem with constancy of variance of the residuals (p=0.1457).The plots of the observed versus the predicted values for response final level is given in figure 4.2.5. The plot shows that the linearity assumption seems to hold. Outlier and influential tests identified that no individual value is influential.

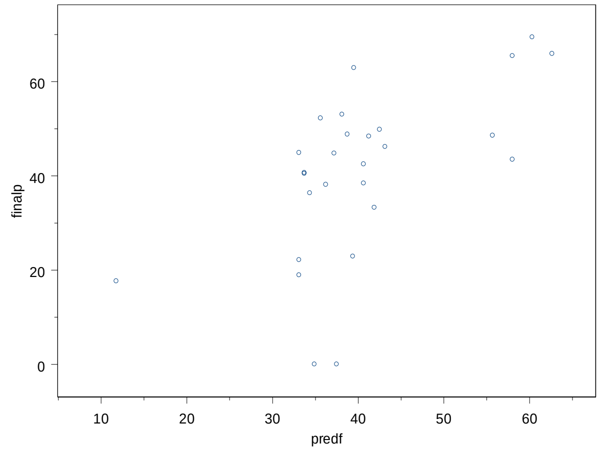


Figure 4.2.5: *The observed versus predicted value plot for*

4.2.2.2 Summary measures as predictor for training nuts

As mentioned before, there was also interest to study the relationship of the measures of performance from complex method and the results obtained during the 'real life test' which involved stitching on a pig and see if prediction is possible from these models. This need arises to know who will be performing better in practice just by looking at his learning behaviour performance measures in the training.

From the models for the five variables that are from the subjective and objective examination assessment of the surgeons, only one of the models was significant. This result is displayed in Table 4.2.9.

Table 4.2.9: *Parameter estimates of the response path distances*

Parameter	Estimate	P-value
intercept	65.6253	0.0024
rate	-18.0399	0.5688
gender	5.1362	0.7569
rate2	0.1431	<.0001
rate2*gender	-0.1392	<.0001

Rate2=rate^-2 and 0=female, 1=male

The adj R-square of the model is 81% of the variation in the number of path distances the surgeon made while stitching on the pig. The F-value obtained is 27.94 with (p<.0001) which shows the covariates in the model are important in explaining the variation observed in the response variable (path). The need for considering the quadratic rate arises from the exploration of the scatter plot of path with rate as shown in the appendix A figureA4. The fitted model was checked for validity of assumptions. The test for normality using Shapiro-Wilk test shows non-normal distribution for the residuals (p=0.0150). The plot of residuals shows a slight deviation from normal as seen in appendix figureA3. The modified levene’s test shows there is no problem with constancy of the residuals (p=0.0759) at 5%level. Surgeon 17 was found to be an outlier and also an influential in this fit. A model was refitted without surgeon 17 and the results found were substantially different from the previous. In the later case all the predictors have insignificant linear relationship with the response. The plot of predicted and observed values is given for both models in figure 4.2.6 and figure 4.2.7.The plot in figure 4.2.6 show that this surgeon is indeed influential. From the plot of figure 4.2.7 we can see that there is no pattern indicating linearity.

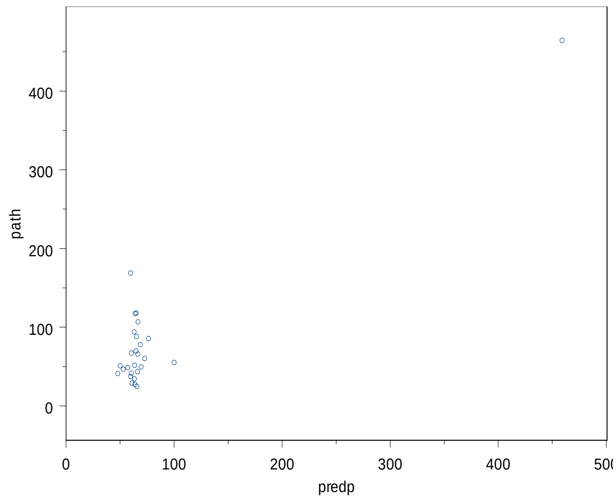


Figure 4.2.6: *The observed versus predicted value*

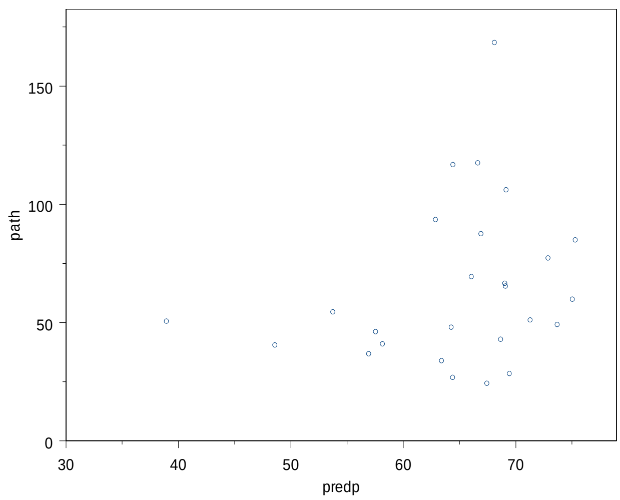


Figure 4.2.7: *The observed versus predicted value plot without surgeon 17*

4.2.2.3 Summary measures as response for training ropes

If we resort our attention to the ropes training, the effects of the various covariates (motivatie, schlauch and gender), including their possible interaction terms, on the responses (initial, rate, length and final performance) were assessed.

First let's consider initial level as a response. The model fit is shown in the Table 4.2.10. The adjusted R-square for this model is found to be below 0.4531.

Table 4.2.10: *Parameter estimates of the response initial*

Parameter	**Estimate**	**P-value**
intercept	265.72120	<0.0001
gender	-53.26737	0.0074
schlauch	-7.04975	0.0107

0=female, 1=male

All assumptions of the model were checked. The test for normality using Shapiro-Wilk test shows normal distribution for the residuals (p= 0.2201). The plot of residuals also shows approximately normal distribution as seen in appendix figureB1. The modified levene's test shows there is no problem with constancy of the residuals (p= 0.0508). The plot of the observed versus the predicted values for response initial is given in the figureB2. The plot shows that the linearity assumption looks reasonable. In addition, the model has no outliers. Conclusions drawn for the simple methods carry over here as well.

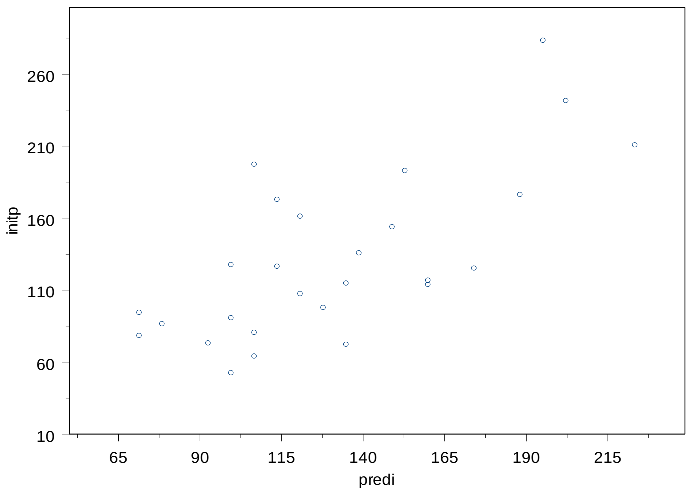


Figure 4.2.8: *The observed versus predicted value for initial*

When length of learning is used as response, the model fit is shown in the Table 4.2.11. The adjusted R-square for this model is found to be below 0.3782.

Table 4.2.11: *Parameter estimates of the response length of learning*

Parameter	Estimate	P-value
intercept	216.92	<.0001
gender	-43.23	0.0254
schlauch	-6.51	0.0168

0=female,1=male

The F-value obtained is 8.91with (p=0.0013) which shows the covariates in the model are important in explaining the variation observed in the response variable (length). All assumptions of the model were checked. The test for normality using Shapiro-Wilk test shows normal distribution for the residuals (p= 0.0705). The plot of residuals also shows approximately normal distribution as displayed in appendix figureB2. The modified levene's test shows there is no problem with constancy of the residuals (p= 0.2103). The plot of the observed versus the predicted values for response length is given in the figure 4.2.9 .The plot shows that the linearity assumption seems to reasonable. In addition, the model has no outliers. Conclusions drawn for the simple methods carry over here for the model with response length of learning.

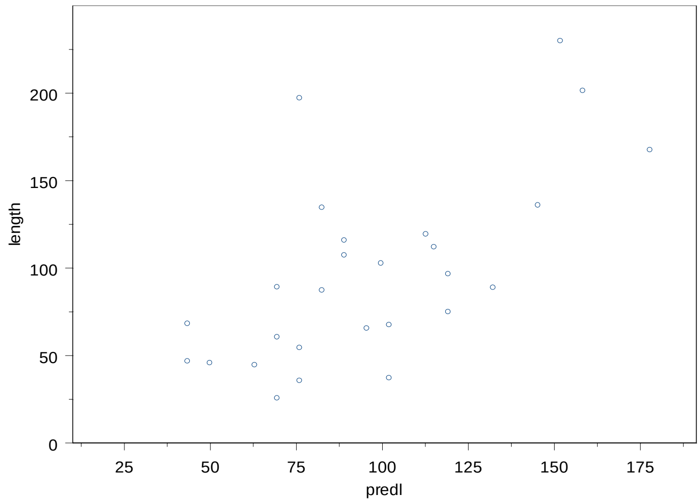


Figure 4.2.9: *The observed versus predicted value for length*

For the case of final level, it is observed that the model is significant with F-value 5.64 and (p= 0.0098) while adjusted R-square of 0.2629.

Table 4.2.12: *Parameter estimates of the response final performance*

Parameter	Estimate	P-value
intercept	79.45	<.0001
gender	-11.39	0.0347
motivatie	-1.01	0.0183

0=female, 1=male

The fitted model was checked for validity of assumptions. The test for normality using Shapiro-Wilk test shows non-normal distribution for the residuals (p= 0.0210). But, the plot of residuals shows a non-skewed distribution as seen in appendix figureB3. It is supposed that the assumption of normality is not violated because this deviation is only because of the high peakedness of the distribution. The modified levene's test shows there is no problem with constancy of the variance of the residuals (p= 0.2765). The plot of the observed versus the predicted values for response final level is given in the figure 4.2.10. The plot shows that the linearity assumption seems to hold. There were no outliers and influential cases.

The negative sign and small p-value for gender indicates that the final time level for female is smaller than that of males. This means the performance of females is better than males when evaluated using the final performance. At 5% level of significance there is evidence that surgeon with higher motivation score will have lesser final time.

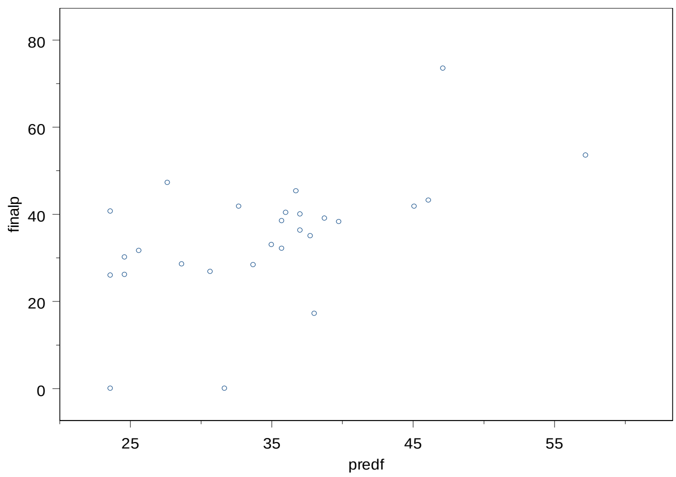


Figure 4.2.10: *The observed versus predicted value for final*

4.2.2.4 Summary measures as predictor for training ropes

From analysis of the relationships between the measures of performance from the complex method and the results obtained during the exam (the 'real life test' which involved stitching on a pig), only the response otime (operation time taken from first incision to last stitch on the pig) is found to be significantly related with the measures. The F-value obtained is 9.89 with (p=0.0007) with adjusted R-square 0.4061.The results of the model parameter estimates is displayed in Table 4.2.13.

Table 4.2.13: *Parameter estimates of the response otime*

Parameter	Estimate	P-value
intercept	211.93	0.0001
rate	-593.46	0.0020
final	6.43	0.0002

All the model assumptions were found to be satisfied .The test for normality using Shapiro-Wilk suggests normality (p=0.7841) and is aided by the plot of residuals as seen in appendix figureB7. The test for constancy of variance was satisfied as suggested by the modified levene's test (p=0.2765). The plot of the residuals versus the predicted values for response otime is given in the appendix figureB4. The plot shows that the linearity assumption seems to hold since there is no trend or pattern in the plot. There were no outliers and influential cases identified.

The negative sign and small p-value (p=0.002) for rate indicates that the higher the learning rate of the surgeon, the smaller operation time he/she take to finish the stitching on the pig. The positive sign and p-value (p=0.0002) for final indicates that the smaller the final time in learning a surgeon has, the small time he/she requires to complete the job in practice.

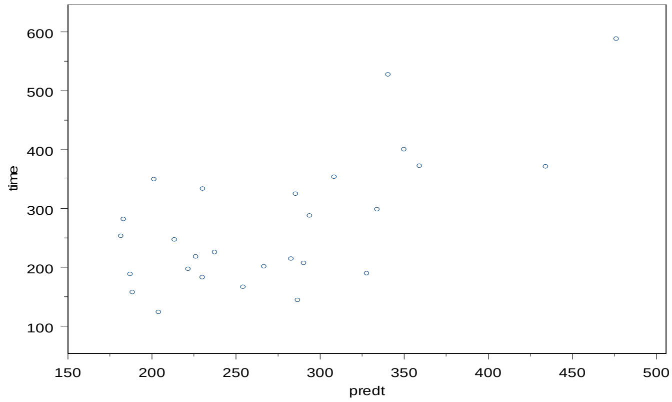


Figure 4.2.10: *The observed versus predicted value for otime*

4.2.3 Summary of the comparison between the simple and complex approaches

Table 4.2.14: *Adj R-square for Simple versus complex approaches*

	Response	Covariates	Adj R-square
Simple approach for ropes	otime	intercept initial length	0.12
Complex approach for ropes	otime	intercept rate final	0.41

As mentioned in the methodology part, the variables used to quantify the learning in the simple approach are initial level, length and final level while for complex (learning curve) approach we used initial level, length, final level and the rate of learning. From Table 4.2.14, it is shown that the variable otime was significantly related with initial level and length in the case of the simple approach but the adj R-square was 0.12 which is very small as compared to the adj R-square for the complex approach (adj R-square=0.41).This indicates that the covariates used to quantify the learning behaviour using the simple approach have low relationships as that of the covariates from the complex approach.

5. Discussion and Conclusions

This study focused on describing statistical techniques that directly assessed the learning curve effect of laparoscopic surgery. The other aims were to investigate if it is possible to use the information from results of psychological test to predict indicators of learning performances and also how to use the information from results of psychological test and indicators of learning performance to predict the results obtained during the exam in the 'real life test’ which involved stitching on a pig.

As a first stage of investigation, the information from each surgeon can be condensed into few statistics that are indicators for performance, called summary statistic. These were done in two different methods. The motivation in considering complex method is because of the fact that obtained values for the measure of performance give better estimates as compared to the quantities obtained from the simpler methods. This is because the performance indicators from the complex method have smaller measurement errors. Variables that are good proxies for learning need to be identified. In this study the identified proxies for learning were: initial level, length of learning, rate of learning and final skill level. The power and exponential curves were contending curves to represent the learning curve effect. Based on the plot and residual analysis, the exponential curve is selected. Surgeon 9 is the fastest learner for the training nuts and a plateau time was reached rapidly by surgeon 14. For the training ropes, surgeon 8 is the fastest learner and surgeon 17 was very rapid to reach plateau.

In the second stage of the two-stage analysis, subsets of predictor variables were included in the model based on mallows Cp and the coefficient of multiple determination (adjusted R-squared). In the analysis of variance, the F test was used to test if there was any relationship between the predictors and response (significance of parameters).

For the training nuts, it was observed from the analysis that the variables schlauch, motivatie, gender, interaction of motivatie and schlauch as well as interaction of motivatie and gender have a significant relationship with the rate of learning. Considering the signs of the estimates and the fact that the covariates are centered, it was concluded that females with higher than average motivational score and average schlauch will have a higher rate of learning. The covariates motivatie, gender and interaction between motivatie and gender have relationship with the final time level of the surgeon, which is the one of the measures of performance. In studying the relationship of the measures of performance and psychological tests versus the results obtained during the exam (the 'real life test' which involved stitching on a pig), only the model with response path or distances covered while stitching on the pig (path) was significant. The model explained 81.13% of the variation in the response but surgeon 17 was found to be highly influential. A model without this surgeon suggests that there is no significant relationship between all of the covariates and the response path.

When the training ropes is considered, there was evidence, at 5% level, that the covariates gender and schlauch have a significant relationship with both initial level and length of learning. Females' performance is better than males' performance. The higher special cognitive ability a surgeon has, the lower his/her starting level will become. Therefore, surgeons who have large special cognitive ability have good performance; as good performance mean smaller initial time. Covariates gender and motivatie are found to be significantly related with the final skill level. In using the final as a measure of performance, the same conclusion was reached as that of initial time with respect to gender i.e. Females' performance is better than males' performance. There is a decrease in final level for surgeons having large motivational score. This means surgeons with larger motivation will perform better. Therefore, one may possibly tell with certain degree of uncertainty whether a particular surgeon will learn the important task faster by looking his motivational score. However care must be taken in using this conclusion as it is based on a model which might not be good to be used for predictions because of the small adjusted R-square. Only the variable time (operation time, the time taken from first incision to last stitch on the pig) is found to be significantly related with measures of performance and psychological tests. There is a decrease in time to complete the job in practice for surgeons whose rates are larger. Moreover, the smaller the final time on the last repetition of the exercise of a surgeon, the small time he/she will take to complete the job in practice.

In considering the two trainings generally, it was observed that there was evidence of relationships between the learning behaviour of a surgeon and his surgerical activities in practice. In addition, the covariates used as a measure of motivation and cognitive ability of the surgeon can be used to identify which surgeon will learn or perform better in the training. Meanwhile, the models fitted for both trainings were not adequate enough for prediction since adjusted R-squares were small. It has to be recalled from the previous sections that two methods have been followed in the first stage namely the simple and the complex methods. The results from these two methods were compared. So, the variables from the complex approach which are used to quantify the learning behaviour of surgeons' have a better relationship with the prediction of performance for the 'real life test' on the pig than the variables from the simple approach. Besides these, arguments stated earlier in the methodology for the simple approaches such as loss of information and the fact that the methods might not capture most of the behaviour of the learning process, like rate which is the important quantifier for learning as was seen in this study. Therefore, it is advisable to use the learning curve (complex) approach in the first stage of the two-stage analysis to capture estimates for the learning behaviour of surgeons.

6. Recommendations

Issues surrounding learning curve effects that require to be addressed are that a particular study asking for prior experience or training of the surgeon in the training under investigation. The results of the questionnaire could then be used to investigate if prior experience or training influences the learning curve. In addition, alternative and better proxies for learning should be investigated. These possibly include surgical near-misses and quality measures from the field of surgery.

7. References

1. Health Technology Assessment NHS R&D HTA Programme: *Statistical assessment of the learning curves of health technologies*: CR Ramsay, AM Grant, SA Wallace, PH Garthwaite, AF Monk, IT Russell

2. Neter,J.,KutnerM.h., Nachtsheim,C.J., Wasserma,W.(1996).*Applied linear statistical models*.MCGraw-Hill,New York.

3. Kenneth R Koedinger2, Brent Martin1, Antonija Mitrovic1 and Santosh Mathan2: *On Using Learning Curves to Evaluate ITS*

4. Rajesh Aggarwal, MRCS, Teodor P. Grantcharov, PhD, Jens R. Eriksen, MD, Dorthe Blirup, Viggo B. Kristiansen, MD, Peter Funch-Jensen, DSc, and Ara Darzi, MD :***An Evidence-Based Virtual Reality Training Program for Novice Laparoscopic Surgeons***

5. Ritter, F. E., & Schooler, L. J. (2002). *The learning curve in International encyclopaedia of the social and behavioural sciences.* 8602-8605. Amsterdam: Pergamon.http://www.iesbs.com/

6. Verbeke G., Molenberghs G. (2000).*Linear mixed models for longitudinal data*. Springer-Verlag, New york.

7. SAS online documentation (http://support.sas.com/91doc/docMainpage.jsp). [Viewed on 07/05/2007]

8. http://www.ispub.com/ostia/index.php?xmlFilePath=journals/ijs/vol9n2/military.xml

9. http://www.bristol-inquiry.org.uk/Documents/r_o_l_c_r_s1.pdf

10. http://www.xitact.com/news/articles/TheLearningCurve.pdf

11. http://www.ncchta.org/fullmono/mon512.pdf

8. Appendices

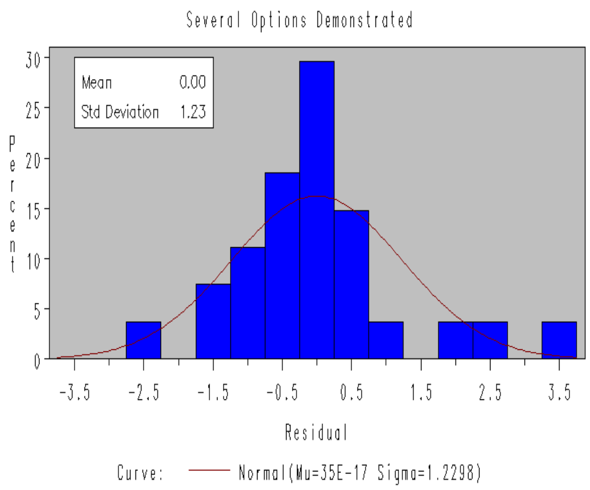
Appendix A: Training Nuts

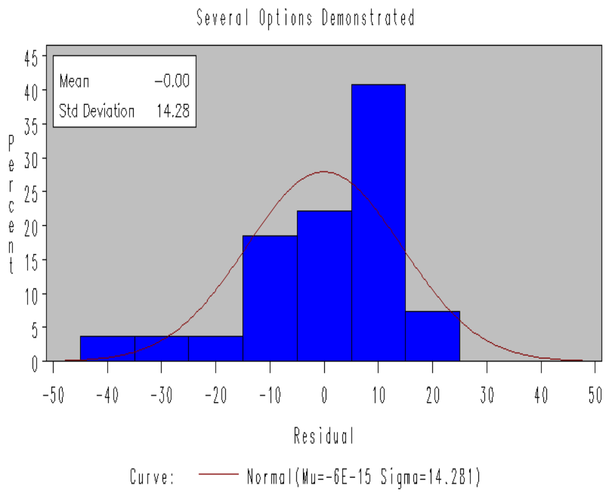
TableA1: *Pearsonean Correlation (p-value) for the performance estimates from complex methods and the psychological test variables*

	Initial	Length	Final	Rate	Motivatie	Schlauch
Initial	1.00	0.98(<.0001)	0.71 (<.0001)	0.59 (0.0010)	0.09 (0.6647)	-0.21 (0.2943)
Length		1.00	0.55 (0.0032)	0.51 (0.0064)	0.03 (0.8650)	-0.19 (0.3213)
Final			1.00	0.66 (0.0002)	0.23 (0.2526)	-0.17 (0.3973)
Rate				1.00	0.39 (0.0402)	0.16 (0.4207)
Motivatie					1.00	0.37 (0.06)
Schlauch						1.00

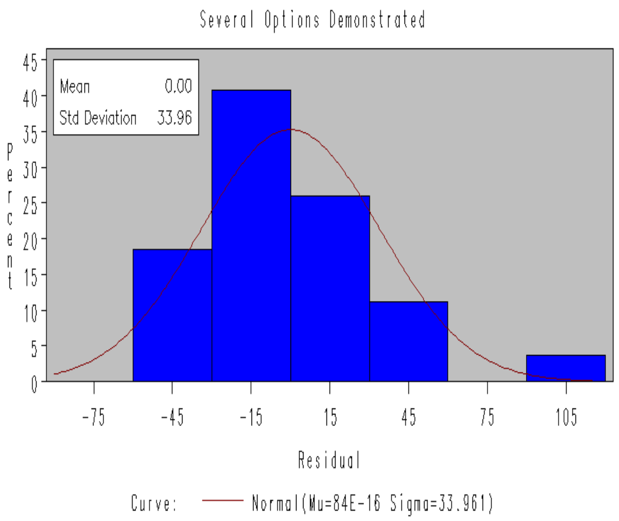
TableA2: *Pearsonean Correlation (p-value) for the performance estimates from complex methods and the exam score variables*

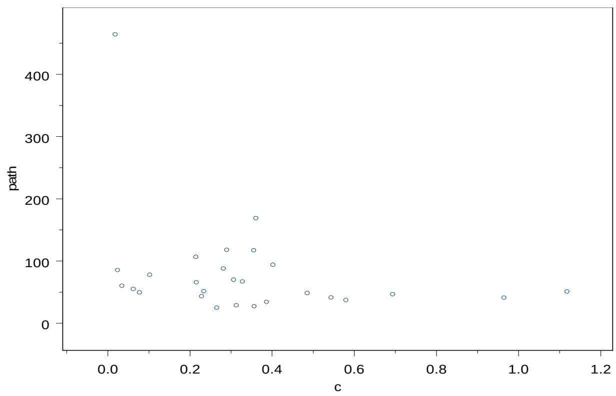
	Initial	Length	Final	Rate	Score1	Score2	Mo	Path	Time
Initial	1.00	0.98 (<.0001)	0.71 (<.0001)	0.59 (0.0010)	-0.26 (0.1805)	-0.30 (0.1275)	-0.33 (0.0930)	-0.23 (0.2577)	-0.19 (0.3244)
Length		1.00	0.55 (0.0032)	0.51 (0.0064)	-0.26 (0.1914)	-0.31 (0.1162)	-0.24 (0.2279)	-0.13 (0.5222)	-0.13 (0.5336)
Final			1.00	0.66 (0.0002)	-0.19 (0.3435)	-0.16 (0.4162)	-0.50 (0.0077)	-0.46 (0.0161)	-0.36 (0.0659)
Rate				1.00	0.05 (0.7993)	0.04 (0.8391)	-0.38 (0.0483)	-0.30 (0.1208)	-0.29 (0.1365)
Score1					1.00	0.97 (<.0001)	0.01 (0.9578)	0.04 (0.8279)	-0.25 (0.2145)
Score2						1.00	-0.02 (0.9286)	-0.02 (0.9169)	-0.30 (0.1253)
Mo							1.00	0.84 (<.0001)	0.67 (0.0001)
Path								1.00	0.66 (0.0002)
Time									1.00





FigureA1: *The plot of residuals for response inverse sqrt of rate* **FigureA2:** *The plot of residuals for response final*

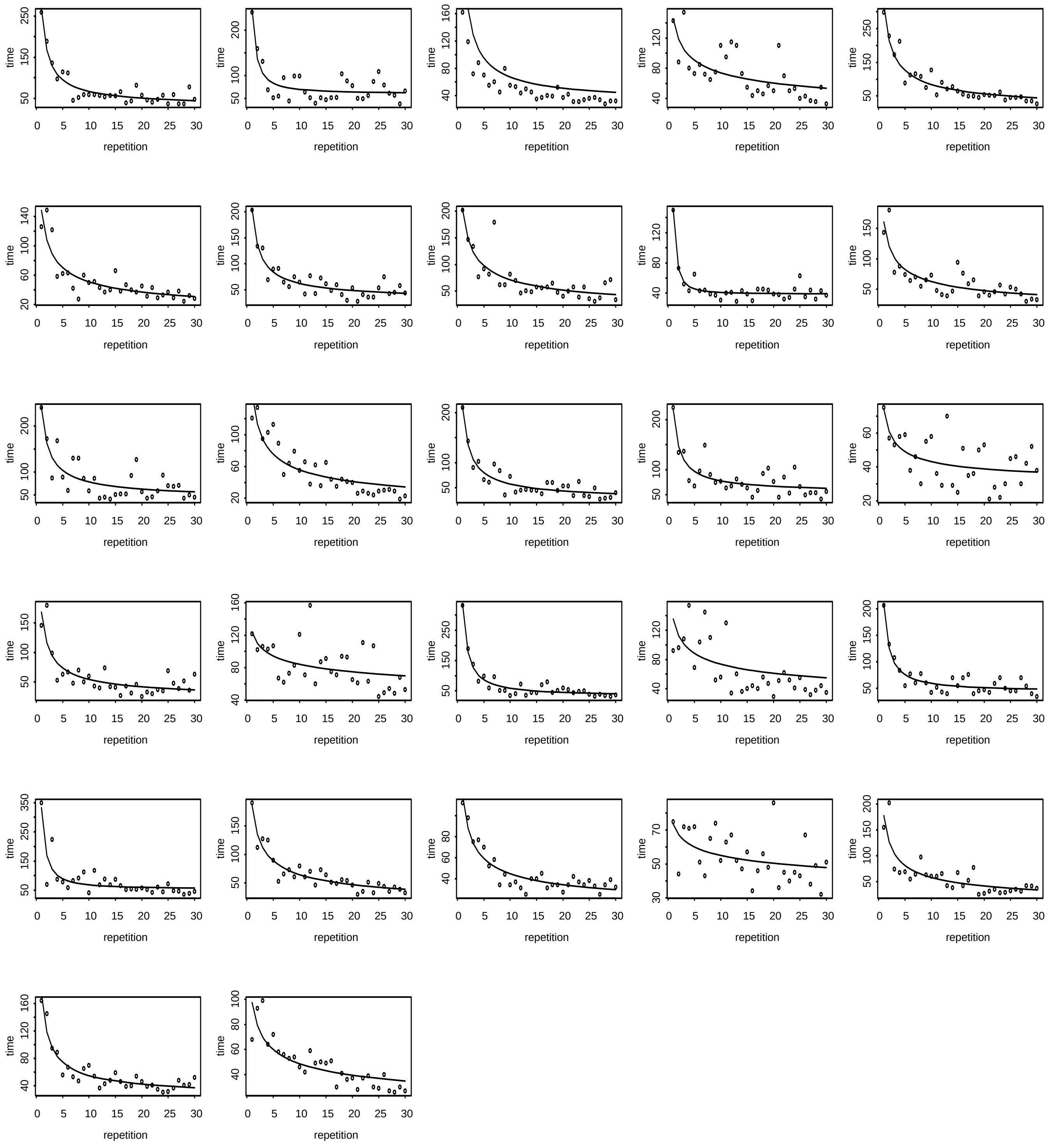




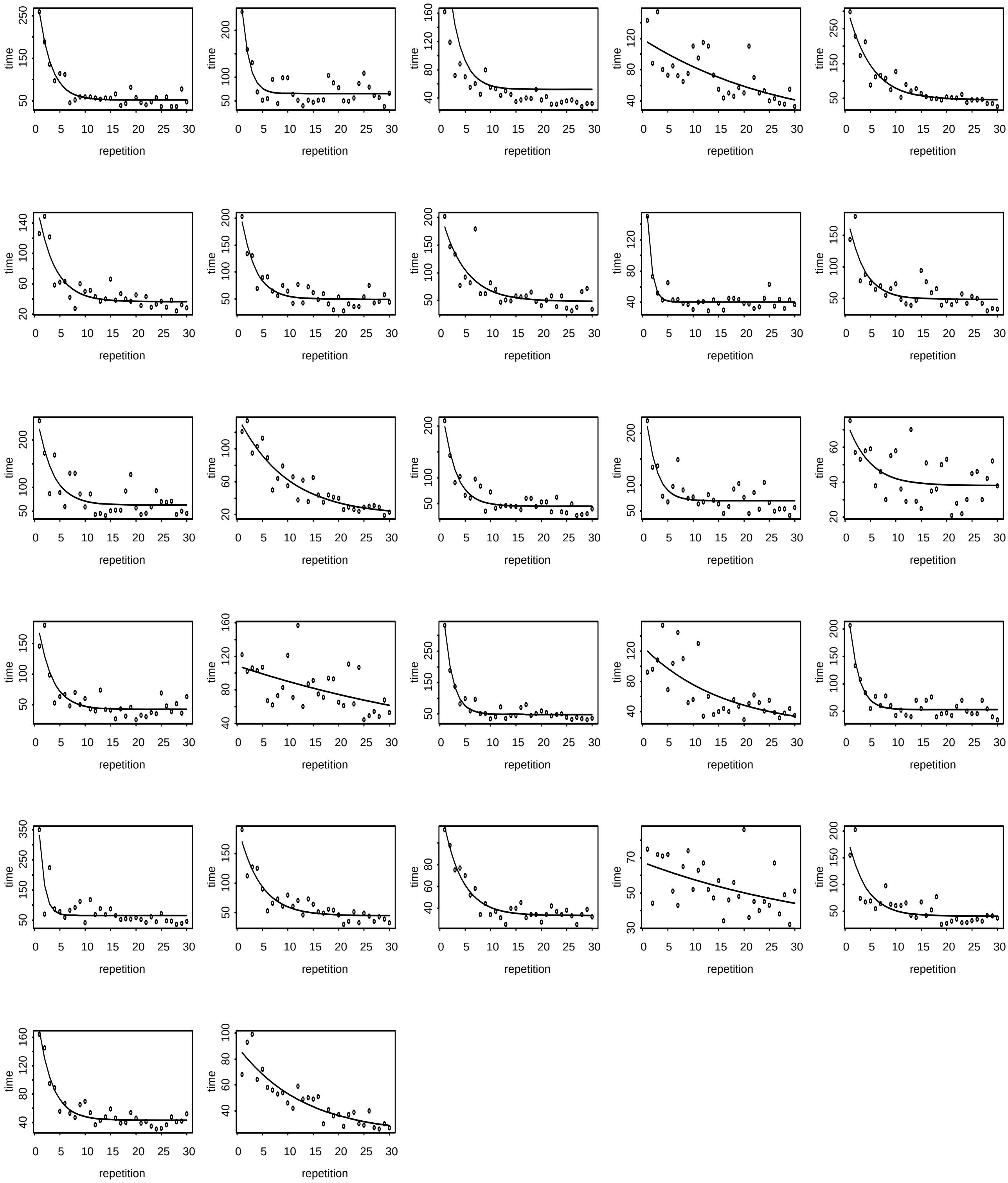
FigureA3: *The plot of residuals for response path* **FigureA4:** *The scatter plot of path versus rate*

TableA5: *The sum of absolute residuals for the power and exponential model*

Training	Sum of absolute residual for power	Sum of absolute residual for exponential
Nuts	10132.75	10033.84
Ropes	6521.09	5895.69



FigureA5: *Observed (dots) and fitted (line) plots for the training nuts using power curve approach*



FigureA6: *Observed (dots) and fitted (line) plots for the training nuts using exponential curve*

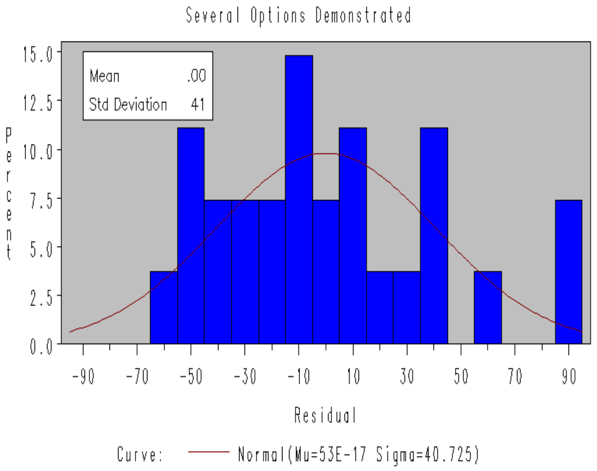
Appendix B: Training Ropes

TableB1: *Pearsonean Correlation (p-value) for the performance estimates from complex methods and the psychological test variables*

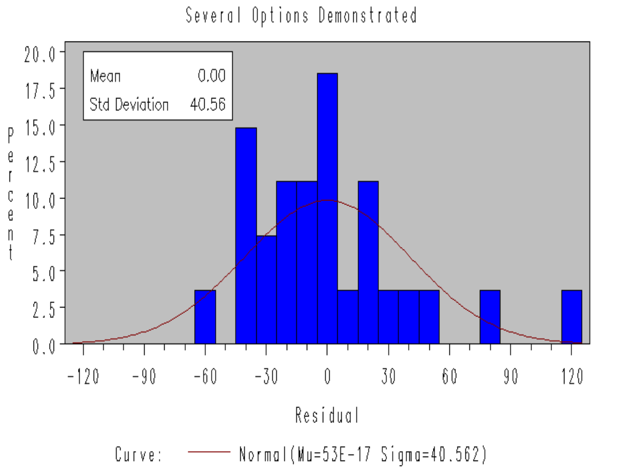
	Initial	Length	Final	Rate	Motivatie	Schlauch
Initial	1.00	0.97(<.0001)	0.38 (0.0517)	0.31 (0.1206)	-0.40 (0.0379)	-0.56 (0.0023)
Length		1.00	0.13 (0.5062)	0.16 (0.4239)	-0.32 (0.1089)	-0.54 (0.0038)
Final			1.00	0.62 (0.0006)	-0.42 (0.0286)	-0.23 (0.2488)
Rate				1.00	-0.39 (0.0665)	-0.31 (0.1149)
Motivatie					1.00	0.38 (0.0530)
Schlauch						1.00

TableB2: *Pearsonean Correlation (p-value) for the performance estimates from complex methods and the exam score variables*

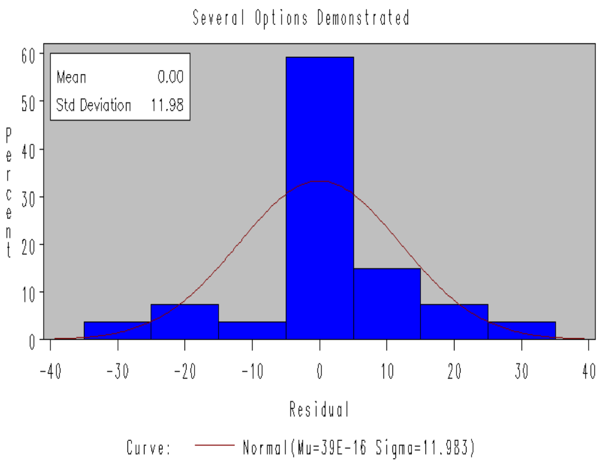
	Initial	Length	Final	Rate	Score1	Score2	Mo	Path	Time
Initial	1.00	0.97 (<.0001)	0.38 (0.0517)	0.31 (0.1206)	-0.09 (0.6590)	-0.16 (0.4280)	-0.03 (0.8974)	0.14 (0.4895)	0.07 (0.7283)
Length		1.00	0.13 (0.5062)	0.16 (0.4239)	-0.02 (0.9366)	-0.06 (0.7604)	-0.16 (0.4345)	0.01 (0.9690)	-0.04 (0.8476)
Final			1.00	0.62 (0.0006)	-0.29 (0.1397)	-0.40 (0.0382)	0.48 (0.0122)	0.52 (0.0055)	0.42 (0.0294)
Rate				1.00	-0.13 (0.5239)	-0.19 (0.3205)	0.04 (0.8327)	0.09 (0.6278)	-0.16 (0.4391)
Score1					1.00	0.97 (<.0001)	0.01 (0.9578)	0.04 (0.8279)	-0.25 (0.2145)
Score2						1.00	-0.02 (0.9286)	-0.02 (0.9169)	-0.30 (0.1253)
Mo							1.00	0.84 (<.0001)	0.67 (0.0001)
Path								1.00	0.66 (0.0002)
Time									1.00



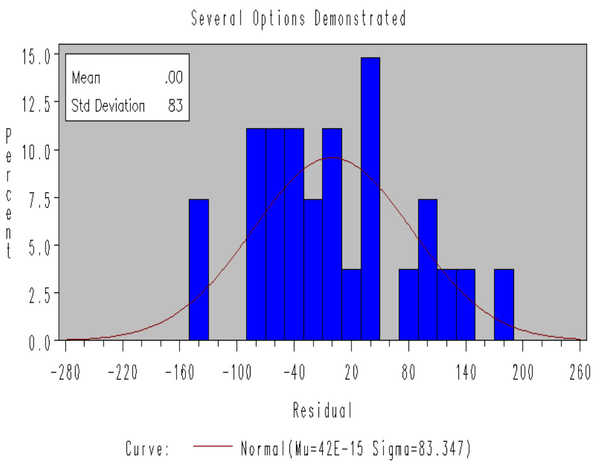
FigureB1: *The plot of residuals for response initial level*



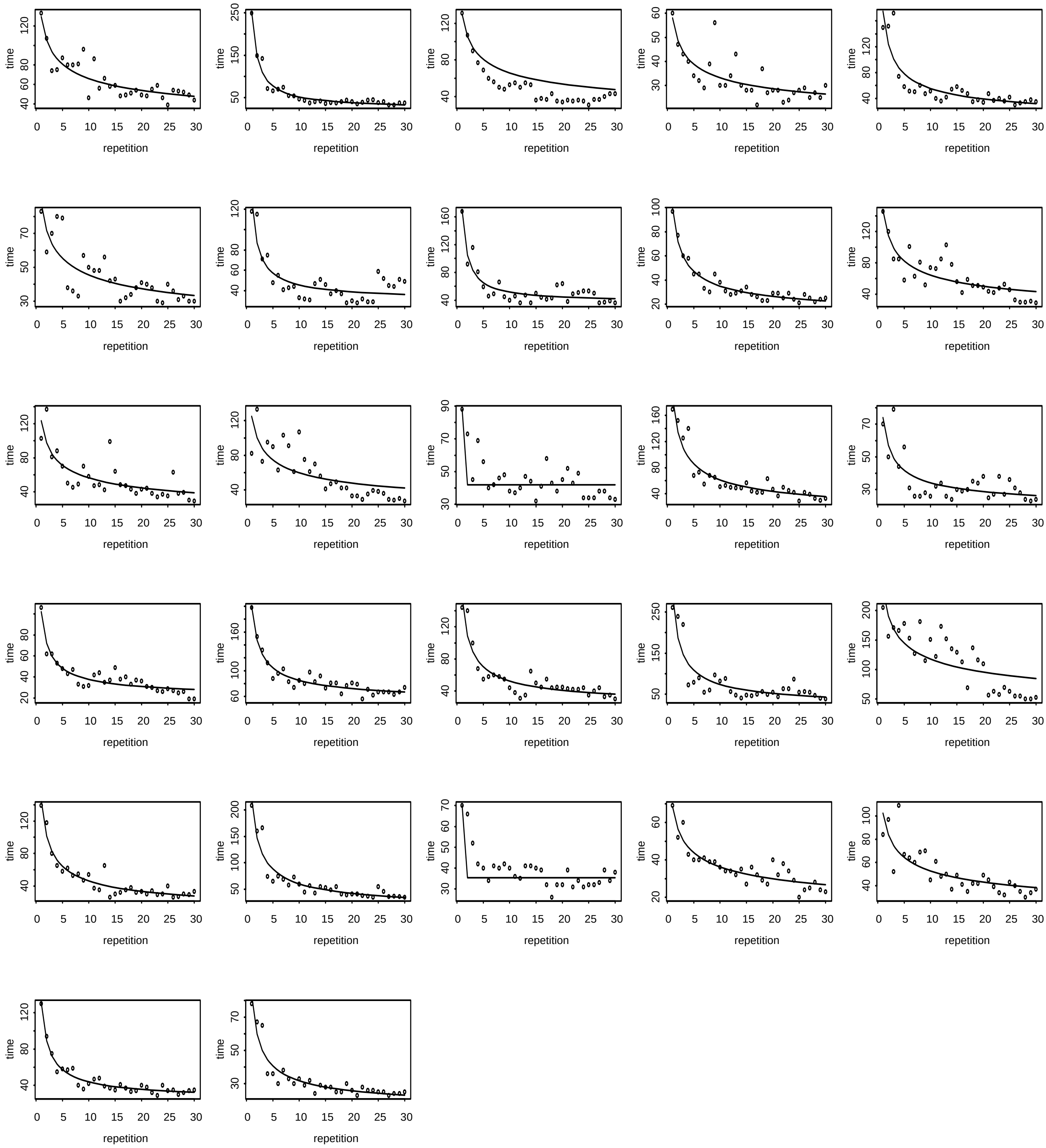
FigureB2: *The plot of residuals for response length*



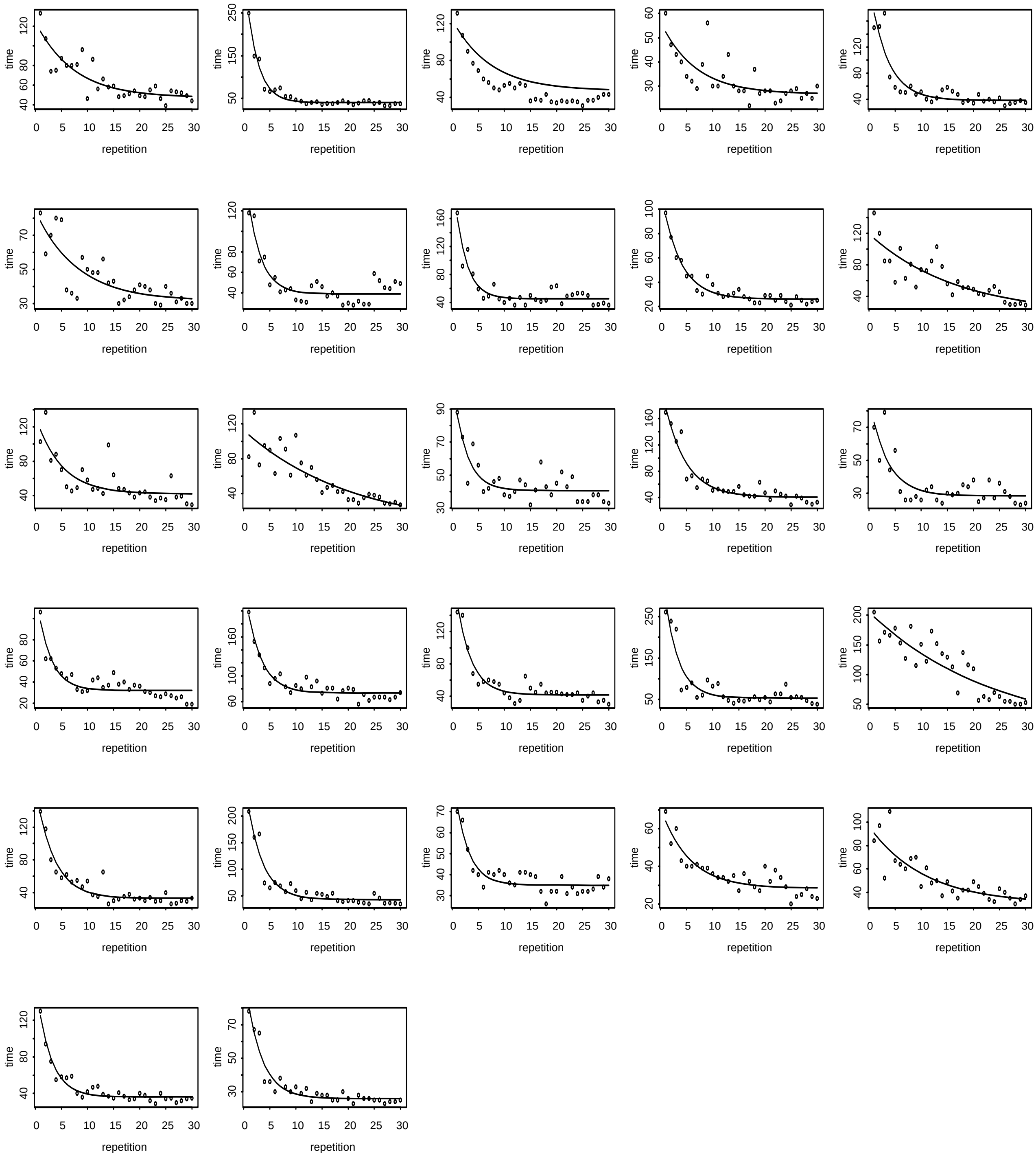
FigureB3: *The plot of residuals for response final level*



FigureB4: *The plot of residuals for response otime*



FigureB5: *Observed (dots) and fitted (line) plots for the training ropes using power curve approach*



FigureB6: *Observed (dots) and fitted (line) plots for the training ropes using exponential curve approach*

Auteursrechterlijke overeenkomst

Opdat de Universiteit Hasselt uw eindverhandeling wereldwijd kan reproduceren, vertalen en distribueren is uw akkoord voor deze overeenkomst noodzakelijk. Gelieve de tijd te nemen om deze overeenkomst door te nemen, de gevraagde informatie in te vullen (en de overeenkomst te ondertekenen en af te geven).

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Two-stage analysis of learning curves on laparoscopic study of surgeons
Richting: **Master of science in Applied Statistics** Jaar: **2007**
in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Ik ga akkoord,

Sintayehu Aynalem

Datum: **08.06.2007**