Training Self-Regulation Skills: Does The Accuracy of Task-Selection Depend On Who You Are Choosing For and Self-Efficacy?

Jelmer Prinsen

Utrecht University, Faculty of Social Sciences

201600025: Master's Thesis

Dr. V. Hoogerheide, F.M. Lucas & Dr. C. Hulshof

27th of July 2023

Word count: 7371

Abstract

Earlier research has found that self-regulated learning is challenging due to the demand on the learner to monitor and assess their learning progress. The Self-Assessment and Task-Selection (SATS) training proved that self-regulated learning can be trained. However, the person who is choosing and self-efficacy can affect self-regulated learning outcomes. This study, therefore, examined if task-selection accuracy would be dependent on choosing a task for yourself or choosing a task for someone else and subsequently, if that effect would be dependent on self-efficacy. 125 secondary education students participated in this experimental study. Participants followed a Self-Assessment and Task-Selection (SATS) training, afterwards rated their self-efficacy, and solved biology tasks which they chose for themselves and biology tasks they chose for someone else. Results showed that task-selection accuracy did depend on choosing for yourself or someone else with choosing a task for someone else having the highest task-selection accuracy. Subsequently, this effect was not moderated by participants self-efficacy. Results suggest that social aspects can affect self-regulated learning, and therefore learning performance. Replication of this study with choosing a task for someone else and investigation on how social aspects affect self-regulated learning is recommended.

*Keywords: s*elf-regulated learning, task-selection choice, task-selection accuracy, self-efficacy

Introduction

People around the world are constantly learning with the intention of pursuing desired learning outcomes. To harness how people can effectively learn, they have to self-regulate their learning which is important because it can help learners how to judge their own performance (Panadero & Romero, 2014). Developed self-regulated learning abilities are important because they are associated with higher performance, both in elementary and secondary school (Dent & Koenka, 2016; Dignath & Büttner, 2008). Self-regulated learning is the process of making constant assessments and choices about what to study next and how to study it (Bjork et al., 2013; Kostons et al., 2009), and is managed via abilities of monitoring and control (Dent & Koenka, 2016; Kostons et al., 2012; Raaijmakers et al., 2018). Monitoring and control, for example, are used in problem-solving tasks during STEM education where the learner needs to self-assess which task to do in what order. Monitoring refers to self-evaluation of performance during task performance, self-evaluation on a performed task, or self-evaluation on tasks that have yet to be performed (Bjork et al., 2013). Learners can use monitoring to control their learning activities effectively to see if or which task is appropriate to choose next (Bjork et al., 2013; Kostons et al., 2009; Miller, 2003).

However, assessing one's performance and selecting the next appropriate task within the learning process is difficult (Kostons et al., 2009). Generally, learners have a poor memory of their results after finishing a task (Kostons et al., 2010). Accurately monitoring one's own performance is difficult because learners often must split their limited cognitive resources between multiple tasks (Kostons et al., 2010). Poor accuracy of monitoring can, in turn, further lead to learners overestimating or underestimate themselves which can lead to inaccurate control of new task-selection choices. Therefore, training self-regulated learning is important because choosing the next appropriate task can go wrong when a learner struggles with monitoring and control and their learning performance may suffer due to choosing the wrong task (Kostons et al., 2009; Kostons et al., 2010).

Training Self-Regulated Learning Abilities

Previous research from Kostons et al. (2012) has shown that self-regulated learning abilities are trainable, which can improve learning outcomes. Further evidence from Raaijmakers et al. (2018), who used the same training as Kostons et al. (2012), found that learners who had received training in task-selection and self-assessment and afterwards had to work on problem-solving tasks also showed higher learning outcomes. Kostons et al. (2012) and Raaijmakers et al. (2018) used the Self-Assessment and Task-Selection training (SATS) which is based on the social learning theory from Bandura (1977). The social learning theory implies that the incentives and consequences that learners create for themselves give them some control over their actions and behavior (Bandura, 1977). The SATS training lets the learners watch a video with a model giving modeling examples of solving hereditary biology problems and training about self-assessment and task-selection. The model was working on a problem-solving task while at the same time reading the stated problem and afterwards sharing the necessary steps for solving the task by thinking aloud. When the problem-solving task was done, the model rated how much mental effort they invested to solve the problem. The model proceeded to rate how many steps of the task they performed correctly, ranging from 0 to 5 and elaborating that for each correct step one point should be given. Afterwards the model used an algorithm to determine which new task should be selected. In the algorithm, the score of mental effort and the sum of correct steps from the tasks are used to determine which task was suitable to choose next. Learners could choose the next best task out of 75 tasks. The 75 tasks were divided between five levels of complexity with each complexity level having 15 tasks. Each complexity level had three different levels of support. All the complexity levels had 5 tasks and had either high support, low support or no support. Due to multiple complexity levels and support levels the choices learners made could be observed and compared to show differences in task-selection accuracy (Raaijmakers et al., 2018; Van Merriënboer & Kirschner, 2018). After the video learners had to problem-solve tasks from a database in the domain of biology on their own without help.

Findings from Raaijmakers et al. (2018) and Kostons et al. (2012) suggest that learners are trainable in making self-assessments and task-selections after following the SATS training and had higher learning outcomes as a result of that. However, Raaijmakers et al. (2018) also suggested that there can be differences in cognitive demands when choosing a task for yourself or someone else.

Selecting a Task for Yourself or Others

For the SATS training to work optimally and as intended, it is important to uncover what affects the learning outcomes of the SATS training. Findings from Kostons et al. (2010) suggest that learners experience extraneous load on various levels when training self-assessment and task-selection. Extraneous load is the number of resources from the working memory that is not used for learning, therefore potentially disturbing learning potential (Sweller, 1988; Sweller et al., 2019). When performing problem-solving tasks, learners must work on their tasks while at the same time monitoring their own performance. This can place a burden on the cognitive load experienced when doing multiple tasks, imposing split attention (Van Merriënboer & Kirschner, 2018). Doing multiple tasks can therefore result in ineffective processes of performing tasks and imposing extraneous load (Sweller, 1988). This implies that in order to properly execute problem-solving tasks, it is important to keep the extraneous load as low as possible. Prior research done by Raaijmakers et al. (2018) elaborate that participants were cognitively pressured because they had to monitor and assess their own performance and choose a new task for themselves.

However, it is less cognitively demanding when learners are given the information they require to make a decision and can therefore completely dedicate their attention selecting the next best task. For example, when learners are provided the mental effort and performance rating, only their working memory is needed to choose a suitable task. This potentially leads to less extraneous load. In fact, multiple studies had their participants choose a task for someone else (Polman & Vohs, 2016; Raaijmakers et al., 2018) with positive results. However, Bandura (1977) elaborates that a learner's conviction of their efficacy also can play a role in providing the desired results.

Self-Efficacy and Task-Selection Accuracy

Self-efficacy refers to confidence in one's ability to plan and carry out the actions necessary to achieve certain goals (Bandura, 1997; Martin et al., 2021; Putwain et al., 2015). A stronger self-efficacy improves people's capacity to remain cool in the face of adversity, to set higher standards for themselves, and to try new ideas or take more risks (Pajares, 1996; Putwain et al., 2012). Self-efficacy, together with self-assessment and task-selection, can be an effective self-regulatory component (Baars & Wijnia, 2018; Sitzmann & Ely, 2011). Baars and Wijnia (2018) did a study on the relation between motivation profiles and self-regulated learning. One key finding was that learners who experienced higher motivation generally showed higher self-efficacy beliefs and obtained higher learning outcomes as a result. The findings from Baars and Wijnia (2018) are important because their research shows that self-efficacy can indeed affect learning outcomes.

Ridgley et al. (2021) investigated if students could adjust their self-regulated learning approaches when problem-solving tasks. Students demonstrated lower self-efficacy when solving difficult problems, lower performance evaluations, and lower effort (Ridgley et al., 2021). One possible explanation is that learners chose less strategic ways when tasks became too challenging, therefore making the challenge too big and affecting their confidence to get the desired outcome (Komarraju & Nadler, 2013; Ridgley et al., 2021; Putwain et al., 2012). This in turn could mean that when learners, who followed the SATS-training, are working on the task database as used by Raaijmakers et al. (2018) can be affected by a low self-efficacy when working on tasks that are too demanding. Even though the video-modeling examples in the SATS-training from Kostons et al. (2012) and Raaijmakers et al. (2018) did not particularly target learners' self-efficacy, they showed that learning by witnessing models has a beneficial effect on learners' self-efficacy (Bandura, 1997; Schunk & Hanson, 1985). Subsequently, to the best of our knowledge, there is no literature on the importance of self-efficacy in the context of SATS training. However, related research does suggest that self-efficacy can affect the desired learning outcome (Baars & Wijnia, 2018; Komarraju & Nadler, 2013; Ridgley et al., 2021; Putwain et al., 2012), and will therefore be included in this study.

The Present Study

Research has shown that the SATS training can train the self-regulated learning abilities of monitoring and task-selection and subsequently learning performance (Kostons et al., 2010; Kostons et al., 2012; Raaijmakers et al., 2018). However, it is unclear which other factors are affecting the SATS training. This study therefore explored whether selecting a task for yourself or someone else and self-efficacy can affect the accuracy of monitoring and task-selection, and subsequently learning performance by answering the first research question: *Does task-selection* *accuracy depend on choosing a task for yourself or someone else?* Regarding if self-efficacy affected task-selection accuracy, the second research question was: *Does this effect depend on learners' self-efficacy?*

The hypothesis for the first research question was that task-selection accuracy would depend on whether you choose a task for yourself or someone else. It was expected that choosing a task for someone else would be more accurate due to experiencing less extraneous load when making choices without monitoring and assessing your own performance (Sweller, 1988; Sweller et al., 2019). The hypothesis with research question two was that self-efficacy would affect the accuracy of choosing a task for yourself or for someone else due to prior findings that self-efficacy can affect the desired learning outcome (Baars & Wijnia, 2018; Komarraju & Nadler, 2013; Ridgley et al., 2021; Putwain et al., 2012).

Method

Participants

This experimental study used a within-subjects design. Using the G*power test (appendix A) from Faul et al. (2009), this study required 104 participants or more for a statistical power of at least 0.80 (Field, 2018). The study took place at a secondary school within a classroom, including a total of 125 participants. However, one participant did not give informed consent, 8 participants did not have an ID number which made the data untraceable, and one participant had partial data therefore excluding them from the dataset. Consequently, the data of 115 participants were used in this study. The final sample consisted of 57 males (M = 15.51, SD = 0.54) aged between 15 and 17 years old, and 58 females (M = 15.36, SD = 0.52) aged between 15 and 17 years old.

Materials

Qualtrics

Qualtrics (2022) is a program that can collect data from surveys and was used as a platform to answer the informed consent, self-efficacy rating, pretest, SATS training, mental effort, self-assessment of performance and the use of the algorithm. Qualtrics is a program that was provided by the University of Utrecht.

Self-Assessment and Task-Selection Training (SATS) & Task Database

The participants were trained by self-assessment and task-selection training (SATS; cf. Kostons et al., 2012; Raaijmakers et al., 2018), which is a social learning approach (Bandura, 1977) that trains adolescents' self-assessment and task-selection skills. The participants began the training by watching a short three-minute introductory video about biology in which the topic of monohybrid crossing problems (e.g., dominant/recessive, homozygous/heterozygous) was introduced. After the training introduction the participants watched a 20-minute video modeling example where two human models, an adult man and woman, dealt with four tasks with each task having a different complexity and difficulty level, in first person view. During the video the problem statement and all steps done by the models were always displayed on the screen, but the steps were progressively built up as the model was thinking aloud during problem solving. After the problem solving, the model showed how to rate how much mental effort it took to complete the task.

Participants used an algorithm shown in Table 1, in combination with their mental effort and performance assessment scores, to select the next task. For example, when a learner had a mental effort score of 7 and a performance score of 2 then, according to the algorithm (Table 1), the participant should choose a task of a lower complexity level and more support. This meant going one step to the left in the column of Table 2. Going to the left would result in the participant getting a less complex task, going to the right would result in getting a more complex task.

Table 1

Algorithm for Task-Selection Advice With the Use of Performance and Mental Effort

Performance			
4-5	+2	+1	0
2-3	+1	0	-1
0-1	0	-1	-2
	1-3	4-6	7-9 Effort

Table 2

Task Database Containing the 75 Tasks

Niveau 1 - 2 generaties - 1 onbekende - 1 oplossing - deductief			***Niveau 2*** - 2 generaties - 1 onbekende - meerdere oplossingen - deductief			***Niveau 3*** - 2 generaties - 1 onbekende - meerdere oplossingen - inductief			***Niveau 4*** - 3 generaties - 1 onbekende - meerdere oplossingen - inductief en deductief			***Niveau 5*** - 3 generaties - 2 onbekende - meerdere oplossingen - beide kanten op		
Veel hulp	**Enige hulp**	**Alles zelf**	**Veel hulp**	**Enige hulp**	**Alles zelf**	**Veel hulp**	**Enige hulp**	**Alles zelf**	**Veel hulp**	**Enige hulp**	**Alles zelf**	**Veel hulp**	**Enige hulp**	**Alles zelf**
Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur	Oogkleur
Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm	Haarvorm
Kattenvacht	Huntington	Wolman syndroom	Apert syndroom	Bloemkleur	Huntington	Fruitvliegjes	Staart-lengte	Tongrollen	Albinisme	Fruitvliegjes	Hazenlip	Melkallergie	Staart-lengte	Appelboom
Appelboom	Kippen-snavel	Staart-lengte	Fruitvliegjes	Appelboom	Albinisme	Kippen-snavel	P.R.A. ziekte	Depressie	Kattenvacht	Bloemkleur	P.R.A. ziekte	Apert syndroom	Wolman syndroom	Fruitvliegjes
Depressie	Hazenlip	Melkallergie	P.R.A. ziekte	Tongrollen	Mamma-carcinoom	Wolman syndroom	Melkallergie	Apert syndroom	Depressie	Tongrollen	Mamma-carcinoom	Huntington	Bloem-kleur	Kippen-snavel

Self-Efficacy

Participants were asked to rate their self-efficacy (appendix B) with two independent of each other, items which were adapted from Bandura (2006). The first item was "*how confident are you that you can solve these biology problems?"* and was named "SE biology". The second item was "*how confident are you that you can select a suitable next biology problem?"* and was named "SE task-selection". Participants answered on a scale from (1) very, very little, (3) very little, (5) not much/not little, (7) very much, and (9) very, very much (Bandura, 2006), for the translation to Dutch which was used see appendix B. Participants rated their self-efficacy before and after the practice phase of choosing tasks for yourself, and before and after the practice phase of choosing tasks for someone else.

Mental Effort

Mental effort (appendix C) was measured immediately after each task. The rating scale was designed by Paas (1992) and Bratfisch et al. (1972) and ranges from (1) very, very little effort, (3) little effort, (5) neither little nor much effort, (7) much effort, and (9) very, very much effort. The four scorings of mental effort from choosing a task for yourself were used and depicted when participants had to choose a task for someone else four times.

Performance Rating

Participants performance was assessed on a 6-point rating scale ranging from 0 to 5 immediately after each task. The program Qualtrics (2022) automatically scored the performance when participants chose a task for themselves and provided a performance score when participants had to choose a task for someone else. A point was given for every step in the problem-solving procedure participants had completed successfully. For example, a score of 0 meant none of the steps were correct and a score of 5 meant all steps were correct. The same rating scale from Raaijmakers et al. (2018) and Kostons et al. (2012) was used. Again, scorings for performance from choosing a task for yourself were used and displayed when participants had to choose a task for someone else.

Practice Phase

Participants were asked to select a next task from the task database as depicted in Table 2 after completing each of the biology tasks. New tasks were chosen with the algorithm from Table 1 by using the performance rating, which was calculated and provided by the program Qualtrics, and the mental effort from the last task. Afterwards, participants used the task database when choosing a next task for themselves four times, and when choosing a next task for someone else four times. Participants had to first complete the four tasks of choosing a task for yourself before they could continue choosing a task for someone else. This is due to the scores from mental effort and performance rating from choosing a task for yourself being used all four times when participants had to choose a task for someone else. For example, when a participant had scored a 2 on mental effort and a 5 on performance when choosing a task for yourself, that score would be depicted within one of the four tasks of choosing a task for someone else. Because of this, no counterbalancing measures were and could have been taken because else participants would not have been able to choose a task for someone else.

Procedure

Participants first read the information letter (appendix D) and were asked to sign the informed consent letter (appendix D), which was accessible via the URL provided by the attending researcher. After signing the informed consent participants watched an introductory video from the SATS training about Mendel's law of heredity, monohybrid cross problems in the biology domain and how to solve them. After the introductory video participants watched a 20-minute video of a model and immediately after this video, participants were asked to rate their self-efficacy.

Then, an introductory letter regarding what participants had to do was presented, before they moved on to completing the four tasks. After completing each of the four tasks' participants had to choose a suitable next task using the algorithm presented in Table 1.

Next, participants solved four tasks in which they had to interpret the performance and mental effort of a fictitious peer student. Participants first read an introduction explaining what they were supposed to do. Based on information about task complexity, support level, performance and mental effort, participants had to choose which new task the fictitious student should have selected.

Ethics

Participants read the information letter (Appendix D) regarding informed consent (appendix E) and needed to sign the informed consent letter before they could proceed with the experiment. Participants who were 16 years or older could sign their own informed consent papers. Whenever a participant was 15 years or younger, their parents and the participant themselves both had to sign the informed consent in order for their data to be used in this experiment. Participants were able to keep the information letter if requested. All parents of the participants received an information letter two weeks prior to the experiment. The information letter (appendix D) informed parents about this experiment and that they passively were giving consent unless they actively signed the online objection form (appendix E), which they could access with the link which was provided in the information letter. Both participants and parents were also informed that up to 6 weeks after the experiment they still had the chance to retroactively change their informed consent. After six weeks the key file, which included names of the students linked to the participants number, were cleared making it impossible thereafter to look up participants data. The performance data, which was stored separately from the key file, contained the data and participants numbers.

Data was collected and managed confidentially, which meant that participants are anonymous, untraceable, and data was safely stored in YODA (Utrecht University, 2020). Participants were notified that they could leave the experiment at any given time without being questioned, however all data, complete or partial, would be saved. To ensure that the ethical principles are properly followed and adhered to, this research was authorized by the Faculty Ethics Review Board (FERB) of the faculty of Social Sciences at the University of Utrecht. Participants were told that their data will be saved in the data bank of the University of Utrecht's Faculty of Social Sciences for up to at least 10 years.

Data Analysis

The program of SPSS (IBM corp., 2021) was used for the analysis. The first step was to exclude participants who did not give informed consent, did not have an ID number and had partially missing data. The first research question was answered by computing and comparing the mean task-selection accuracy scores from the four tasks of choosing a task for yourself and the four tasks of choosing for someone else. The scores of the means could be positive, neutral or negative. Participants who have gotten a negative value for their task-selection accuracy were recoded to positive ones. For example, if a participant had a task-selection accuracy of –3 the recoded number for this participant would be having a task-selection accuracy of 3. This did affect the outcome because when recoded, data could be properly analyzed and compared. For example, a -1 has the same deviation from 0 as having a score of 1 however being positively coded, data could be compared with each other. A repeated measures ANOVA was used to reveal if there was a significant difference between the means of selecting a task for yourself and selecting a task for someone else.

The second research question was answered by measuring the two items *how confident are you that you can solve these biology problems?* and *how confident are you that you can select a suitable next biology problem?* Both items were measured separately and independently from each other. Both self-efficacy questions were asked before participants worked on selecting a task for yourself or someone else. In this way, participants were only affected by the SATS training experience and were not yet affected by experiences from working on biology problems in the practice phase. This way the self-efficacy questions were unbiased and therefore ruled out the possibility that any prior experience could have affected participants' confidence. During analysis, the means scores of choosing task for yourself and choosing a task for others were used as dependent variables in a repeated measures ANCOVA with the self-efficacy items as two separate covariates.

Results

Descriptives

First for clarification, the task-selection accuracy variables of "*choosing a task for yourself"* were stated as "TS accuracy self" (M = 0.74, SD = 1.16), and "*choosing a task for someone else"* as "TS accuracy other" (M = 0.43, SD = 0.65).

In regard to research question two, the variable "*how confident are you that you can solve these biology problems?"* was named "SE biology" (M = 6.77, SD = 1.68) and variable "*how confident are you that you can select a suitable next biology problem?"* was named "SE task-selection" (M = 7.06, SD = 1.25). Both self-efficacy items used had a range of 1 – 9 and were independent from each other measuring their own item.

Task-Selection Accuracy

To answer research question one: "*does task-selection accuracy depend on choosing a task for yourself or someone else?"* a repeated measures ANOVA was conducted. However, first the assumptions had to be met. The first assumption of independent observations was met. The second assumption of sphericity was irrelevant because only a comparison of two means was conducted (Field, 2018). A Shapiro-Wilk test further revealed that the third assumption of normal distribution was violated for TS accuracy self, W(115) = .64, p = .001, and TS accuracy other, W(115) = .65, p = .001. Due to violation of normality, a non-parametric Friedman test was used to answer research question one.

To determine which were the outliers first a box plot, shown in Figure 1, was used depicting both the variables TS accuracy self and TS accuracy other. Next, to determine the frequency of the outliers the z-scores of every participant were computed for both TS accuracy self and TS accuracy other variables. A full disclosure of outliers per z-score cutoff is shown in appendix G. Cases with a z-score > 3.29 were determined as extreme outliers. In sum the data was big enough that outliers did not gave any major divergent results and were what you would expect to find in a normal database (Field, 2018). Therefore, research question one was answered with all the outliers present. However, to give an indication whether the outliers could have affected the outcome an analysis was also conducted without the outliers and results are found in appendix H.

Figure 1

Visualization of Outliers for Both Conditions of Task-Selection Accuracy

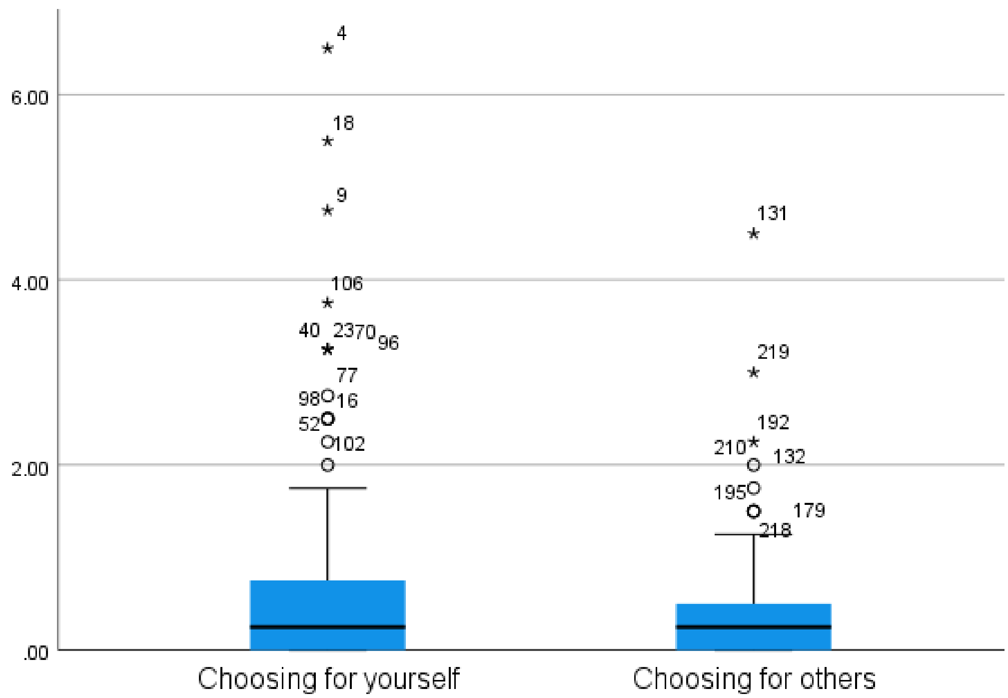


Table 3 shows the outcome of the Friedman test and starts with the inference that there was a significant difference in task-selection accuracy between TS accuracy self and TS accuracy other. TS accuracy other has a mean closer to zero when compared to TS accuracy self, suggesting that participants who chose a task for someone else were more accurate. Furthermore, the standard deviation from TS accuracy self is high in contrast to the standard deviation from TS accuracy other indicating more variation in choices made.

Table 3

Results Non-Parametric Friedman Test

	N	*df*	*M*	*SD*	χ^2	*p*	*Range*
Task-selection accuracy							
TS accuracy self	115	1	0.74	1.16	5.88	.015	0 - ∞
TS accuracy other	115	1	0.43	0.65	5.88	.015	0 - ∞

Self-Efficacy

The two items which were used to answer research question two: "*does this effect depend on learners' self-efficacy?"* are shown in Table 4. SE biology (*M* = 6.77, *SD* = 1.68) and SE task-selection (*M* = 7.06, *SD* = 1.25) scores were, in regards of having a range between 1 – 9, quite high indicating a high confidence. Furthermore, the standard deviations of both items stay relatively close to the mean, suggesting a little bit of variation in confidence. In order to answer research question two, a repeated measures ANCOVA was conducted with TS accuracy self and TS accuracy other as dependent variables and the two self-efficacy items as covariates.

However, prior to conducting a repeated measures ANCOVA, assumptions were checked. The distribution of normality assumption was, according to the Shapiro-Wilk test for both self-efficacy variables, violated for biology problem-solving (W (115) = .91, $p < .001$) and selecting a next task (W (115) = .92, $p < .001$). Nevertheless, Field (2018) states that the assumption of normality becomes less important as sample sizes increase since the sampling distribution will be normal regardless of how our population (or sample) data appear. In this regard, the decision was made to continue as if normality was not violated. The second assumption of independent observations was met. The third assumption of sphericity was irrelevant because, again, we were comparing only two means of dependent variables with two covariates (Field, 2018).

As was the case with handling outliers with research question one, a z-score > 3.29 was used to determine extreme outliers (Field, 2018). Only SE biology had one outlier with a z-score > 3.29 and a full disclosure about the outlier can be found in appendix G. However, according to Field (2018), these numbers are expected when working with a normal database and research question two was therefore answered with the outlier. Again as with research question one, to gain insight whether the outlier would have affected the outcome the repeated measures ANCOVA was also conducted without the outlier and results are found in appendix H.

To ascertain if there was any relationship, negatively or positively, between the variables a correlation analysis has been conducted and results are found in Table 4.

Table 4

Correlations Between Task-Accuracy and Self-Efficacy

	N = 115			
Variable	1	2	3	4
1. TS accuracy self	-			
2. TS accuracy other	.27**	-		
3. SE biology	-.18	-.05	-	
4. SE task-selection	-.04	-.20*	.52**	-

Note. *p < .05. **p < .01. *** = without outlier having a z-score > 3.29.

Table 4 shows a clear significance between the self-efficacy variable SE task-selection and TS accuracy other which suggests a negative linear relationship meaning that for example a participant had low self-efficacy of selecting a new task, the TS accuracy other would be high and vice versa. Furthermore, both self-efficacy variables also have a positive linear relationship meaning that when one of the self-efficacy variables was high than so was the other and vice versa. Lastly, the TS accuracy self and TS accuracy other had a significant positive relationship.

To answer research question two a repeated measures ANCOVA was conducted with task-selection accuracy and SE biology (Table 5), and with task-selection accuracy and SE task-selection (Table 6). Table 5 shows that task-selection accuracy did not depend upon SE biology meaning that it did not have an effect on task-selection accuracy. Furthermore, it is also revealed that SE task-selection did not moderate task-selection accuracy meaning that SE task-selection did not affect task-selection accuracy.

Table 5

Repeated Measures ANCOVA Results for Task-Selection Accuracy and Covariate SE Biology

				N = 115		
Variable	*df*	*SS*	*MS*	*F*	*p*	η^2
TS accuracy	1	3.48	3.48	5.18	.025	.044
TS accuracy X SE biology	1	1.78	1.78	2.65	.106	.023
Error	113	75.77	0.67			

Table 6

Repeated Measures ANCOVA Results for Task-Selection Accuracy and Covariate SE task-selection

	N = 115					
Variable	*df*	*SS*	*MS*	*F*	*p*	η^2
TS accuracy	1	0.04	0.04	0.05	.819	.000
TS accuracy X SE task-selection	1	0.37	0.37	0.55	.461	.005
Error	113	77.18	0.68			

Discussion

The aim of this study was, in regard to and improving the SATS training, to get a better understanding if selecting a task for yourself or for someone else would result in task-selection accuracy differences and therefore indicating if these factors affect the outcome of self-regulated learning. And in addition to this, if task-selection accuracy was dependent on the learners' self-efficacy. Self-regulation of learning is difficult because learners have to be able to control their choices and monitor their performance in their own learning process (Kostons et al., 2012). This study helps to answer which variables can affect self-regulation of learning and subsequent learning performance. Therefore, the importance of this study is evident. The results showed and confirmed, in line with hypothesis one, that task-selection accuracy did depend on whether you choose a task for yourself or someone else. As for hypothesis two, which stated that task-selection accuracy would be dependent upon self-efficacy, the results did not lead to similar conclusions. Results are elaborated in the section below.

As expected, a difference in task-selection accuracy was found between choosing a task for yourself and someone else. An explanation for why learners were more accurate at selecting a task for someone else might be that the difference between task-selection choices could have been affected by the cognitive load experienced when making choices. When learners have to choose a task for themselves, they have to choose the task, work on it and assess which task to choose next. Sweller (1988) elaborates that a learner, while problem solving, simultaneously has to manage multiple goals such as considering the magnitude of the problem, the end goal and a strategy for how to deal with the problem which imposes higher extraneous load on the processing capacity (Sweller, 1988; Van Merriënboer & Kirschner, 2018). However, when learners have to work on only choosing a task for someone else, less processing capacity is needed which can explain why choosing a task for someone else has a higher task-selection accuracy.

Furthermore, it is interesting that selecting a task for someone else had a higher task-selection accuracy because participants, in actuality, chose the next best task from prior scorings of their own when choosing for someone else. Earlier findings did suggest that choosing for someone else (Raaijmakers et al., 2018) omit a high extraneous load when learners only focus on one thing at a time. Nonetheless, it could also be that a social aspect of choosing for yourself versus choosing for others can be affecting the task-selection accuracy. Research done by Destan et al. (2017) investigated how children would judge themselves and other unidentified kids and results suggested that children may have been impacted by such juvenile and self-serving biases whilst evaluating their own performance, but more objective reasoning was used to evaluate the performance of the unidentified other kid. An explanation for only reasoning objectively when judging the other might be a person's desire to maintain their self-esteem or their tendencies to view themselves more favorably than others in a constructive, self-serving way (Leary, 2007). In other words when participants chose a task for someone else it was less threatening to their self-esteem and could therefore judge without being biased.

An alternative explanation might also be that learners were more accustomed and experienced with selecting a new task because they already completed the four tasks of choosing for yourself. In an earlier study, Okita (2014) found that when students learned how to evaluate themselves, they were better equipped to not only pick up the material at hand but also to self-monitor and fix their own errors. Because of this practice the learners received during choosing a task for yourself, they might have chosen more accurately when choosing a task for someone else due to the carryover effect (Afsarinejad & Hedayat, 2002). The carryover effect introduces a transfer of knowledge and multiple studies have used a counterbalance effect to remedy the carryover problem. However, the reason this study did not counterbalance the carryover effect was that the scores from performance and mental effort of the first four tasks of choosing a task for yourself were used when they had to choose a task for someone else. This means that participants could not have started with choosing a task for someone else because the scores were not yet in place to be cognitively assessed. Still, the carryover effect might have affected the test results and future research could investigate if this holds true by implementing a counterbalancing method when replicating this study.

The results regarding hypothesis two were also not as expected due to task-selection accuracy not being dependent upon self-efficacy. This contradicts earlier findings in which self-efficacy did affect self-regulated learning (Baars & Wijnia., 2018). A possible explanation for this finding could be that participants were already quite accurate in selecting the next best task and that self-efficacy, SE biology, did not really contribute to enhance accuracy further. Prior findings from Raaijmakers et al. (2018) and Kostons et al. (2012) showed that due to the SATS training learners made more accurate self-assessments and task-selections. It could be that participants did not really rely on their self-efficacy because it had nothing to do with choosing the next best task. Future research could investigate if measuring self-efficacy before the SATS training might differ in measuring the self-efficacy variables after the SATS training. In this way, the impact the SATS training has on self-efficacy can be investigated.

Next, the result of self-efficacy item SE task-selection also found no main effect and this result was also not in line with the hypothesis for research question two. The mean of SE task-selection was actually high (M = 7.06), implying that the self-efficacy was very high. When interpreting Table 6 it is shown that task-selection accuracy is non-significant while it is significant in Table 3 and Table 5. A potential explanation for this could be that outliers or violation of normal distribution did affect the outcome, however this study did proceeded with all the outliers due to keeping a representation of what you would normally find in a dataset (Field, 2018). A replication of this study could examine if the same results hold true.

Furthermore, a potential reason why no main effect was found even though self-efficacy was high could be that even though participants did not know the question they were going to get, they already had confidence due to the SATS training. They had practice with selecting a new task, in which it did not matter to their confidence which task they had to choose because they knew how the system worked. Another potential reason why selecting SE task-selection did not have a main effect was because participants did not start the actual practice phase yet. Remember, self-efficacy was measured before the participants started working on the tasks. the confidence of participants was not yet affected by their performance (Komarraju & Nadler, 2013; Ridgley et al., 2021; Putwain et al., 2012). Future research could do a replication of this study to measure self-efficacy after the SATS training to rule out if this is indeed the case.

Limitations

The first potential limitation of this study involves the assumption of normal distribution, which was violated for all the variables. The assumption's violation increases the likelihood that the outcomes would be incorrectly positive or negative, which could increase the chance that the study findings are invalid. As an example, Table 6 has task-selection accuracy as non-significant while the findings of research question one found a significant difference between task-selection accuracy of choosing for yourself or someone else. The finding of non-significance of task-selection accuracy at Table 6 could potentially be attributed to outliers or violation of normal distribution. However, according to research by Lumley et al. (2002) an argument can be made that there is no such thing as the ideal distribution of normality. Statistical tests for normality are believed to have poor power in small sample sizes and high power in big sample sizes. Lumley et al. (2002) argue that the ideal distribution of normality fails to exist, making the normality test undesirable since in the former, the distribution generally matters while in the latter, it does not. Micceri (1989), who performed research with 440 individuals, found results that are exactly in accordance with Lumley et al. (2002). Micceri (1989) discovered that with variation, severe tail weight or asymmetries were more often than not the norm.

The second limitation was the heuristic of measuring the task-selection accuracy. The task-selection algorithm, and the SATS training, is backed up by the findings of Raaijmakers et al. (2018) and Kostons et al. (2012). Learners make more accurate self-assessments and accurate task-selections when they receive instructions from the SATS training. However, a nuanced argument can be made that because we are measuring people's choices in a scale that is unending with items that really do not measure just one construct, it is hard to pinpoint what exactly is accuracy because the ideal task-selection accuracy is not just affected by mental effort and performance. The algorithm was designed with a certain systematical heuristic without a certain nuanced sensitivity. For example, when a participant is ready for 5 steps more complexity, but the algorithm has calculated that that number is 2 then the participant is told that his or her task-selection was not accurate if the participant chooses to go for 5 steps of more complexity. Consequently, this also makes it hard to validate and calculate the reliability due to all the complex layers. A possible solution, albeit possibly difficult to create, for future research could be to develop and use an algorithm that adjusts itself accordingly to the choice of the participant afterwards. Using the previous example, when the participant scores high on his/her own chosen task while not recommended from the algorithm, the algorithm adjusts itself to the choice made by the participant.

A third and last limitation was that the experimental conditions were not counterbalanced and/or randomized. In this study, all participants experienced the same sequence of interventions. Not having counterbalanced the sequence of interventions limited the generalizability because the participants were not in an equal distribution across the experiment (Field, 2018). Participants first worked on choosing a task for themselves. It could be that the carryover effect (Afsarinejad & Hedayat, 2002) was present and that participants had practice with choosing a task for yourself and therefore knew how to work on choosing a task for someone else. However, the reason this method was chosen in this study was that participants could assess themselves without them knowing it was their own score. If possible, future research could try to implement a counterbalance method therefore resolving the carryover effect. When doing so the results of the study exclude biases and therefore could be more generalized because the effects of existing knowledge about biology and frequency of interventions are accounted for (Field, 2018).

Implications

The main finding was that task-selection was more accurate when participants chose a task for someone else. This knowledge contributes to making the SATS training more efficient, and therefore we understand more about what affects self-regulated learning. A future implication with regards to the SATS training is that future research does a replication with only choosing a task for someone else to examine which social aspects are at play within and between participants. This is important because it could make the SATS training more efficient when it comes to learning outcome.

Another possible implication for the field could be that teachers use the finding that choosing for someone else has a higher potential of task-selection accuracy to their advantage. For example, in assigning peers to give each other feedback and choosing a task for each other. It would be interesting to use this knowledge in future studies to see what affects the task-selection accuracy when you choose the next best task for someone else, and which social aspects are at play. For example, would task-selection accuracy be affected when there is a strong positive or negative relationship between the person who chooses the next best task and the one who has to do the task? Would task-selection accuracy differ when participants do or do not know or know each other?

Furthermore, the difference of task-selection accuracy between choosing for yourself and choosing for someone else suggests that there are social aspects at play. Task-selection accuracy when choosing for someone else might be more accurate due to experiencing less extraneous load

(Sweller, 1988) and reasoning more objectively due to maintaining self-esteem (Destan et al., 2017; Leary, 2007). Future teachers and educators could, for example, make use of this in their lessons where peers assess each other's work anonymously. When assessing each other's work anonymously, the social aspect of a possible relationship with the peer's work is not a bias towards the assessment.

Future research

Directions for further research would be to replicate this study and add a component about the social aspect of viewing yourself more positively than others in a self-serving manner (Leary, 2007) when choosing a task for yourself or someone else. Comparing outcomes of different measures of, for example, self-esteem could give interesting results whether task-selection accuracy is affected.

Future research could also do a replication study in which participants solely choose for someone else. To further investigate the social aspects which might be at play, it might be interesting to have participants choose a task for someone else who is anonymous and for someone they know.

Lastly, implementing a dynamic algorithm which adjusts itself accordingly to the scoring of the participant could also give interesting results. If it is feasible to develop such an adapting algorithm is the question. However, when possible, the adapting algorithm could account for possible task-selection accuracy differences between advised task and chosen task.

Conclusion

Task-selection accuracy was indeed dependent on choosing for yourself or someone else, where results showed that task-selection accuracy was higher when choosing a task for someone else. These results suggest that there are social aspects at play which might affect the SATS training and therefore learning outcomes.

Up until now, there was little to no knowledge whether self-efficacy could affect task-selection accuracy choices from the SATS-training. The findings from this study however revealed that task-selection accuracy was not dependent on self-efficacy, therefore implying that confidence does not always affect task-selection accuracy.

Future research should further investigate which social aspects are affecting self-regulated learning. Furthermore, conducting a replication study with an adaptive algorithm could potentially give a bigger picture about error-free task-selection accuracy. In short, make sure you choose a task for someone else.

Acknowledgements

I would like to sincerely thank my supervisors, Dr. Vincent and Florence who took the time, helped and guided me through the process of conducting and writing this thesis. Vincent and Florence, you took your own time and resources to listen and give thoughtful and meaningful feedback and for that I am really grateful. Thank you for your patience with me and my process, you are both wonderful people and future students are lucky to learn from you both! However, without the Alberdingk Thijm College I could not have conducted the experiment, for that I thank you Arno. My two sisters Karlijn and Merel, one even living in the high north of Scotland, giving me feedback on my poor English writing I thank you! Further, a special thanks for the existence of Lavazza coffee, my wife who was patient with me and guided me emotionally, and my son for inspiration. Last but not least, my thanks go to Casper for taking the time to approve my thesis plan and giving me extra time to complete my thesis even when he was on his holiday. I am

looking back, and I am happy and proud to have produced this thesis. Kind, humble and appreciative regards from me.

References

Afsarinejad, K., & Hedayat, A. S. (2002). Repeated measurements designs for a model with self and simple mixed carryover effects. *Journal of Statistical Planning and Inference*, *106*(1–2), 449–459. https://doi.org/10.1016/s0378-3758(02)00227-6

Baars, M., & Wijnia, L. (2018). The relation between task-specific motivational profiles and training of self-regulated learning skills. *Learning and Individual Differences*, *64*, 125–137. https://doi.org/10.1016/j.lindif.2018.05.007

Bandura, A. (1977). Social learning theory. Englewood Cliffs: Prentice Hall

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman and company, New York.

Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-efficacy Beliefs of Adolescents*, *5*(1), 307-337

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*(1), 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Bratfisch, O., Borg, G., & Dornic, S. (1972). Perceived item-difficulty in three tests of intellectual performance capacity. *PsycEXTRA Dataset*. https://doi.org/10.1037/e420862004-001

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8

Destan, N., Spiess, M. A., de Bruin, A., van Loon, M., & Roebers, C. M. (2017). 6- and 8-year-olds' performance evaluations: Do they differ between self and unknown others? *Metacognition and Learning*, *12*(3), 315–336. https://doi.org/10.1007/s11409-017-9170-5

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Field, A. P. (2018). *Discovering statistics using Ibm Spss statistics*. SAGE.

IBM Corp. Released 2021. *IBM SPSS Statistics for Windows*, Version 28.0. Armonk, NY: IBM Corp. Retrieved January 27, 2023, from https://www.ibm.com/support/pages/how-cite-ibm-spss-statistics-or-earlier-versions-spss

Komarraju, M., & Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences*, *25*, 67–72. https://doi.org/10.1016/j.lindif.2013.01.005

Kostons, D., van Gog, T., & Paas, F. (2009). How do I do? investigating effects of expertise and performance-process records on self-assessment. *Applied Cognitive Psychology*, *23*(9), 1256–1265. https://doi.org/10.1002/acp.1528

Kostons, D., van Gog, T., & Paas, F. (2010). Self-assessment and task selection in learner-controlled instruction: Differences between effective and ineffective learners. *Computers & Education*, *54*(4), 932–940. https://doi.org/10.1016/j.compedu.2009.09.025

Kostons, D., van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction*, *22*(2), 121–132. https://doi.org/10.1016/j.learninstruc.2011.08.004

Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, *58*(1), 317–344. https://doi.org/10.1146/annurev.psych.58.110405.085658

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, *23*(1), 151–169. https://doi.org/10.1146/annurev.publhealth.23.100901.140546

Martin, A. J., Kennett, R., Pearson, J., Mansour, M., Papworth, B., & Malmberg, L.-E. (2021). Challenge and threat appraisals in high school science: Investigating the roles of psychological and physiological factors. *Educational Psychology*, *41*(5), 618–639. https://doi.org/10.1080/01443410.2021.1887456

Micceri, T. (1989). The Unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, *105*(1), 156–166. https://doi.org/10.1037/0033-2909.105.1.156

Miller, P. J. (2003). The effect of scoring criteria specificity on peer and self-assessment. *Assessment & Evaluation in Higher Education*, *28*(4), 383–394. https://doi.org/10.1080/0260293032000066218

Okita, S. Y. (2014). Learning from the folly of others: Learning to self-correct by monitoring the reasoning of virtual characters in a computer-supported Mathematics Learning Environment. *Computers & Education*, *71*, 257–278. https://doi.org/10.1016/j.compedu.2013.09.018

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, *66*(4), 543–578. https://doi.org/10.3102/00346543066004543

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? the effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, *21*(2), 133–148. https://doi.org/10.1080/0969594x.2013.877872

Polman, E., & Vohs, K. D. (2016). Decision fatigue, choosing for others, and self-construal. *Social Psychological and Personality Science*, *7*(5), 471–478. https://doi.org/10.1177/1948550616639648

Putwain, D. W., Remedios, R., & Symes, W. (2015). Experiencing fear appeals as a challenge or a threat influences attainment value and academic self-efficacy. *Learning and Instruction*, *40*, 21–28. https://doi.org/10.1016/j.learninstruc.2015.07.007

Putwain, D., Sander, P., & Larkin, D. (2012). Academic self-efficacy in study-related skills and behaviours: Relations with learning-related emotions and academic success. *British Journal of Educational Psychology*, *83*(4), 633–650. https://doi.org/10.1111/j.2044-8279.2012.02084.x

Qualtrics. (2022, November 22). *Qualtrics XM // Krachtige experience management software*. Retrieved January 18, 2023, from https://www.qualtrics.com/nl/

Raaijmakers, S. F., Baars, M., Schaap, L., Paas, F., van Merriënboer, J., & van Gog, T. (2018). Training self-regulated learning skills with video modeling examples: Do task-selection skills transfer? *Instructional Science*, *46*(2), 273–290. https://doi.org/10.1007/s11251-017-9434-0

Ridgley, L. M., DaVia Rubenstein, L., & Callan, G. L. (2021). Are gifted students adapting their self-regulated learning processes when experiencing challenging tasks? *Gifted Child Quarterly*, *66*(1), 3–22. https://doi.org/10.1177/00169862211025452

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology*, *77*(3), 313–322. https://doi.org/10.1037/0022-0663.77.3.313

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin*, *137*(3), 421–442. https://doi.org/10.1037/a0022777

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

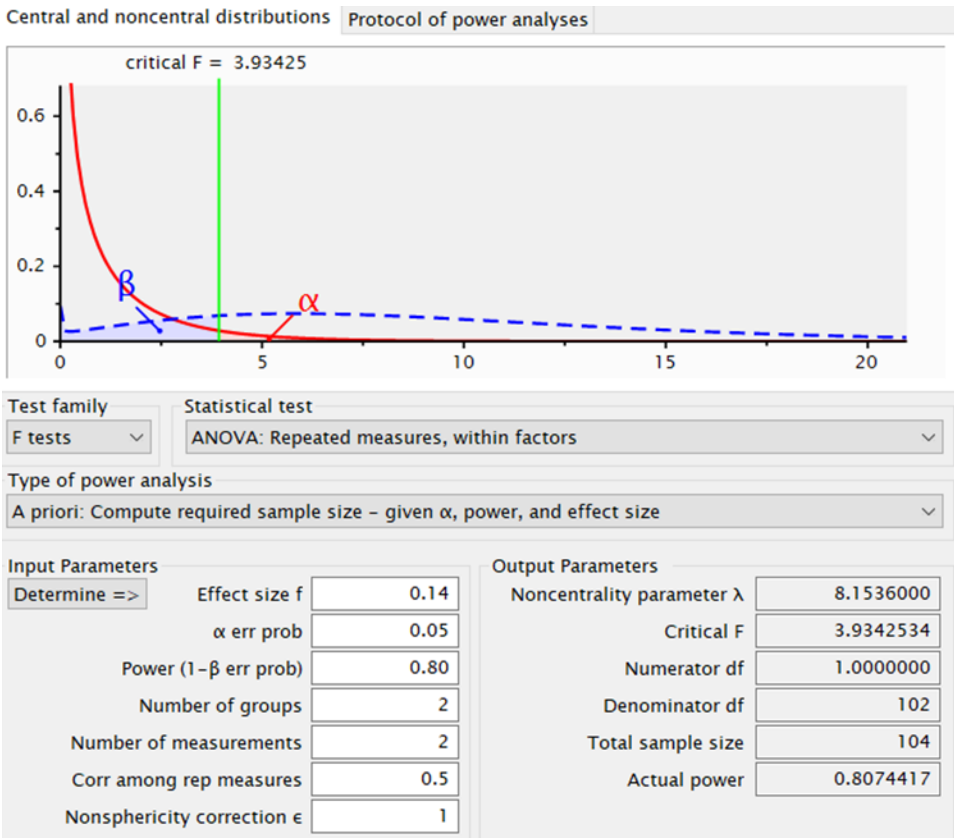
Utrecht University. (2020, September 7). *What is Yoda?* Retrieved January 27, 2023, from https://www.uu.nl/en/research/yoda/what-is-yoda

Van Merriënboer, J. J. G., & Kirschner, P. A. (2018). *Ten steps to complex learning: A systematic approach to four-component instructional design (3rd ed.)*. Routledge.

Appendix A

Calculations from G*power test from Faul et al. (2009)

<u>G*power power estimation, effect size and sample size</u>



- *Effect size = 0.14*
    - *We consulted a methods and statistics expert named Dr. Barbara Flunger. She works at the University of Utrecht, faculty of social sciences. Her recommendation was to have an effect size of 0.14.*

Appendix B

Self-Efficacy Rating Scale

The construct of "self-efficacy" was measured with the two items ‘*how confident are you that you can solve these biology problems?'* and *‘how confident are you that you can select a suitable next biology problem?'* Participants answer on a scale from (1) heel erg weinig, (3) weinig, (5) niet veel/niet weinig, (7) veel, and (9) heel erg veel. Questions are shown below.

Figure 2

How confident are you that you can solve these biology problems?


Hoeveel vertrouwen heb je er op dit moment in dat je in staat bent om de biologie opgaven op te lossen?

Heel erg weinig 1	2	Weinig 3	4	Niet veel / niet weinig 5	6	Veel 7	8	Heel erg veel 9
○	○	○	○	○	○	○	○	○

Figure 3

How confident are you that you can select a suitable next biology problem?


Hoeveel vertrouwen heb je er op dit moment in dat je in staat bent om een geschikte volgende opgave te kiezen?

Heel erg weinig 1	2	Weinig 3	4	Niet veel / niet weinig 5	6	Veel 7	8	Heel erg veel 9
○	○	○	○	○	○	○	○	○

Appendix C

Mental Effort Rating Scale

Participants used figure 4 scale to rate how much mental effort it took to complete the biology tasks and to choose a next best task.

Figure 4

How much effort did it took you to complete and choose a next biology problem?

Hoeveel moeite heeft het je gekost om al deze biologie opgaven te maken en volgende opgaven te kiezen?

	Heel erg weinig 1	2	Weinig 3	4	Niet veel / niet weinig 5	6	Veel 7	8	Heel erg veel 9
	○	○	○	○	○	○	○	○	○

Appendix D

Information Letter

Informatiebrief wetenschappelijk onderzoek naar zelf-regulerend leren

Dr. V. Hoogerheide, Florence & J. Prinsen

Datum

Dinsdag 4 april

Onderwerp

Informatie onderzoek 'Het verbeteren van zelfgestuurd leren'

Telefoon

-

E-mail

J.prinsen@students.uu.nl

Beste ouder/verzorger,

Binnenkort geven wij als onderzoekers van de Universiteit Utrecht een korte training over effectief zelfgestuurd leren op de school van uw kind, en doen we onderzoek om deze training verder te kunnen verbeteren. In deze brief willen we u hierover informeren en vragen om toestemming voor het gebruik van de antwoorden van uw kind voor dit onderzoek naar 'Het verbeteren van zelfgestuurd leren'.

Wat is het doel van het onderzoek?

Uit onderzoek blijkt dat zelfgestuurd leren heel erg lastig is voor veel leerlingen. De Universiteit Utrecht doet onderzoek naar hoe we leerlingen kunnen helpen om zelfgestuurd leren effectiever te maken. Dat doen we door middel van een korte training waarin leerlingen leren om hun eigen prestatie op een biologie opgave te beoordelen en hoe ze op basis daarvan het beste een nieuwe opdracht kunnen kiezen. Het doel van dit specifieke onderzoek is om beter inzicht te krijgen in hoe we die training nog verder kunnen verbeteren. Ook willen we weten hoe we daarbij rekening kunnen houden met verschillen tussen leerlingen (wat voor de één goed werkt, werkt voor een ander misschien niet zo goed). De uitkomsten van dit onderzoek, zijn heel erg belangrijk om de training verder te ontwikkelen zodat deze uiteindelijk gebruikt kan gaan worden op scholen.

Wat houdt het onderzoek in?

Op donderdag 13 april, maandag 17 april of woensdag 19 april komen wij bij uw kind in de klas om dit onderzoek uit te voeren. De leerlingen die deelnemen aan het onderzoek zullen in de klas het onderzoek uitvoeren op de computer. De resultaten van individuele leerlingen worden niet met de school gedeeld, er worden geen cijfers gegeven en de prestatie op de taak telt niet mee in de schoolresultaten.

Het gehele onderzoek duur ongeveer 75 minuten en vindt plaats onder schooltijd.

Privacy en vertrouwelijkheid

Omdat het onderzoek bestaat uit taken die kunnen helpen met het verbeteren van zelfgestuurd leren en de school dit ook van belang acht, doen alle leerlingen mee aan de onderzoekstaken onder schooltijd. Leerlingen mogen zelf aangeven of hun data (de antwoorden op de vragen en opdrachten) wel of niet voor het onderzoek gebruikt mogen worden. Ook u als ouder kunt het aangeven als u er bezwaar tegen heeft dat de data van uw kind gebruikt worden voor dit onderzoek.

I.v.m. het feit dat u bezwaar kunt maken tegen het gebruik van de data van uw kind, vragen wij naar de naam van uw kind. Verder worden er geen persoonsgegevens verzameld. De naam van uw kind wordt in een apart bestand bewaard, los van de antwoorden op de vragen en opdrachten, op een andere (goed beveiligde) server waar niemand behalve de onderzoekers toegang toe heeft. De antwoorden op de trainingsopgaven en vragen worden volledig anoniem verwerkt en opgeslagen. Deze onderzoeksgegevens (de antwoorden op de vragen en opdrachten) worden tot tenminste 10 jaar nadat de resultaten gepubliceerd zijn bewaard op een beveiligde server van de universiteit. Deze anonieme gegevens kunnen ook gedeeld worden met andere onderzoekers zodat ze gebruikt kunnen worden in toekomstig onderzoek.

Mogelijkheid tot vragen en informatie

Als u nog vragen hebt over het onderzoek, stuur dan een mail aan Jelmer Prinsen (j.prinsen@students.uu.nl)

Wilt u contact opnemen met een onafhankelijk onderzoeker over dit onderzoek? Mail dan naar (r.haen@uu.nl). Formele klachten over deze studie kunnen gericht worden tot klachtenfunctionaris-fetcsocwet@uu.nl. Vragen over privacy kunnen gesteld worden aan de onderzoeker, of via privacy@uu.nl.

Toestemming

Mocht u bezwaar hebben tegen het gebruik van de antwoorden van uw kind voor dit onderzoek, dan wil ik u vragen om dat uiterlijk zes weken na de onderzoek sessie via **de link op de volgende bladzijde** aan te geven. Als u geen bezwaar heeft, dan hoeft u niets te doen.

Met vriendelijke groet, namens het onderzoeksteam,

Jelmer Prinsen, Florence & Dr. Vincent Hoogerheide

Appendix E

Informed Consent (Objection Form)

BEZWAARFORMULIER

tegen gebruik van de anonieme antwoorden van uw kind voor onderzoek over het verbeteren van zelfgestuurd leren.

Via school heeft u de informatiebrief ontvangen over de deelname van uw kind aan dit onderzoek. Het onderzoek vindt plaats op school.

Als u <u>niet</u> wilt dat de anonieme antwoorden van uw kind worden gebruikt voor dit onderzoek, willen we u vragen om dit via onderstaande url aan te geven.

https://survey.uu.nl/jfe/form/SV_dg9r771nhfiafum

Als u geen bezwaar heeft tegen het gebruik van de anonieme antwoorden van uw kind voor dit onderzoek, **hoeft u verder niets te doen**.

Appendix G

Full disclosure of outliers for task-selection accuracy and self-efficacy

1.0 = z-score > 3.29

2.0 = z-score > 2.58

3.0 = z-score > 1.96

Normal range / 4.0 = z-score ≤ 1.95

Table 7

Outliers for task-selection accuracy of "choosing a task for yourself":

Zscore(Acuratesse_yourself_mean)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00000	3	1.2	2.6	2.6
	2.00000	1	.4	.9	3.5
	3.00000	4	1.7	3.5	7.0
	'Normal range	107	44.4	93.0	100.0
	Total	115	47.7	100.0	
Missing	System	126	52.3		
Total		241	100.0		

Table 8

Outliers for task-selection accuracy of "choosing a task for someone else":

Zscore(Acuratesse_other_mean)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00000	2	.8	1.7	1.7
	2.00000	1	.4	.9	2.6
	3.00000	2	.8	1.7	4.3
	'Normal range	110	45.6	95.7	100.0
	Total	115	47.7	100.0	
Missing	System	126	52.3		
Total		241	100.0		

Table 9

Outliers for self-efficacy variable "SE biology"

Zscore: Hoeveel vertrouwen heb je er op dit moment in dat je in staat bent om de biologie opgaven op te lossen? - -

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00000	1	.4	.9	.9
	2.00000	1	.4	.9	1.7
	3.00000	5	2.1	4.3	6.1
	4.00000	108	44.8	93.9	100.0
	Total	115	47.7	100.0	
Missing	System	126	52.3		
Total		241	100.0		

Table 10

Outliers for self-efficacy variable "SE task-selection"

Zscore: Hoeveel vertrouwen heb je er op dit moment in dat je in staat bent om een geschikte volgende opgave te kiezen? - -

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3.00000	2	.8	1.7	1.7
	4.00000	113	46.9	98.3	100.0
	Total	115	47.7	100.0	
Missing	System	126	52.3		
Total		241	100.0		

Appendix H

Results from analyses without the outliers

Research question one, task-selection accuracy

Table 11

Results Non-Parametric Friedman Test Without Outliers Having z-score > 3.29

	N	*df*	*M*	*SD*	χ^2	*p*	*Range*
Task-selection accuracy							
TS accuracy self*	110	1	0.57	0.83	5.71	.017	0 - ∞
TS accuracy other*	110	1	0.37	0.48	5.71	.017	0 - ∞

Note. * = results without outliers having z-scores > 3.29.

Research question two, does the task-selection accuracy effect depend upon self-efficacy?

Correlations:

Table 12

Correlations Between Task-Accuracy and Self-Efficacy Without Outliers Having a z-score > 3.29

		N = 114***		
1. TS accuracy self***	-			
2. TS accuracy other***	.28**	-		
3. SE biology***	-.38	-.04	-	
4. SE task-selection***	-.01	-.20*	.52**	-

Note. *p < .05. **p < .01. *** = without outlier having a z-score > 3.29.

Repeated measures ANCOVA:

Table 13

Repeated Measures ANCOVA Results for Task-Selection Accuracy and SE Biology

	N = 114*					
Variable	*df*	*SS*	*MS*	*F*	*p*	η^2
TS accuracy	1	0.296	0.296	0.529	.469	.005
TS accuracy X SE biology*	1	0.008	0.008	0.015	.903	.000
Error	112	62.634	0.559			

Note. * = Results without outlier having a z-score > 3.29.

Table 14

Repeated Measures ANCOVA Results for Task-Selection Accuracy and SE Task-Selection

	N = 114*					
Variable	*df*	*SS*	*MS*	*F*	*p*	η^2
TS accuracy	1	0.314	0.314	0.569	.452	.005
TS accuracy X SE task-selection*	1	0.851	0.851	0.217	.217	.014
Error	112	61.792	0.552			

Note. * = Results without outlier having a z-score > 3.29.

Appendix I

Assignment 4, thesis plan (19-1-2023)

Academic integrity

In order to be able to work as an educational scientist, it is not only important to be able to set up, carry out, and report research, but also to do so in such a way that the academic integrity of the research is ensured at all times. While academic integrity is closely linked to ethical conduct as a professional in general, there are also special points for attention during certain phases of research.

For this assignment, we ask you to reflect on possible issues, risks, and/or dilemmas and describe what measures you will take concerning the following aspects of your research plan:

- *Sample characteristics and consent procedures, together with data handling and storage*

As stated in my method section, the measures I will take to make sure that ethical procedures will be adhered to is that participants will first need to sign the informed consent letter. Participants will not be able to continue the experiment unless the informed consent is signed. The letter informing them of the relevance and importance of the study will be present hardcopy and they can, on request, receive and take an information letter with them.

I will collect and manage data confidentially, which means that participants will be anonymous, non-traceable, and data will be safely saved in YODA. Participants will be notified that they can leave the experiment at any time without being questioned, but that all data, complete or partial, will be saved. It is important to note that whenever a participant is 15 years or younger than they will still be able to participate in the experiment, but their data will not be stored. To make sure that their data will not be stored I will make a separate test in Qualtrics where participants will be send when they answer 15 years or younger button that erases their data at the end of the experiment.

- *Choice of instruments and possibly sensitive questions*

I do not foresee a problem with possible sensitive questions. Questions regarding sexual preference, religion, etc. will not be asked and/or stored. The questions regarding the pre-test/practice phase are about biology. However, the questions regarding self-efficacy can be a little confronting since participants will need to properly assess themselves and answer honestly.

I do however am concerned if it is wise to put a few more items into measuring the construct of self-efficacy. Because if reliability is low later on and I only have one or two items measuring the construct, it will be difficult to get the reliability to a higher level.

- *Effort required from participants and how this weighs against the relevance of the study.*

The time from the participants is valuable and will not be wasted. Therefore, I will try to minimize the time needed to conduct this experiment and opting to not go beyond the number of one and a half hour. This will also counter the experimental fatigue where participants just answer the questionnaire at random because they want to end the experiment. This experimental fatigue might cause participants to not honestly answer the questions anymore which can jeopardize the reliability and validity of this study. It is therefore important to minimize the time needed to conduct this experiment.

- *Any other issues concerning the academic integrity of your study.*

None