Does Montessori Education enhance Executive Functions?
– A Pilot Study –

E.J.M. Persoon
I6052480

Student Master Developmental Psychology
Faculty of Psychology and Neuroscience – Maastricht University

1st supervisor: Dr. L. Jonkman
2nd supervisor: Dr. P. Sins

Word count: 11.457 (without title-page, tables and references)

Abstract

Previous studies concluded that Montessori education may enhance executive functioning, but these studies only looked at general executive functioning and used limited executive functions (EFs) tasks. This pilot study was set out to see if there were any differences between Montessori and Regular education on the three separate core EFs (working memory, inhibition and attention switching) in school aged (7-8 years of age) children (N=48) by the use of multiple standard computerized EF-tasks and questionnaires. This pilot study did not find any difference between the effect of Montessori and Regular education on EF-scores and behavior regulation. Future research with a larger sample size is highly recommended.

"All unnecessary help is an obstruction for the development of natural forces"

Montessori, The Method 1966

1. Introduction

Life in the digital age demands us to become faster and more flexible in our behavior. We shift from one app to the other, calculate how much everybody has to pay when you split a bill and while being in a restaurant with your friend you answer a text or two from that one person you met the other night. In these examples, we use our high cognitive functions to switch attention, be flexible in our thinking strategies, inhibit prepotent responses and manipulate information in our memory. These higher-order cognitive functions are recognized in current theoretical models as executive functions. Three core executive functions are attention shifting or flexibility, working memory capacity and inhibition (Miyake et al., 2000). As we see in our example above executive functions are very important for daily functioning and they also play an important part in learning important and complex academic skills like mathematics and speech during childhood (Fuhs, Nesbitt, Farran & Dong, 2014). Specifically, Clark, Pritchard and Woodward (2010) concluded that differences in executive functioning in children accounted for differences in mathematic performance. Before discussing the importance of executive functions any further, we will first explain what executive functions are and give the definitions for each of the three core executive functions.

Executive functions in adulthood are general, top-down control processes that monitor, regulate and guide our cognitive functions and our behavior. Miyake et al. (2000) studied executive functions in a large group of adults who performed multiple executive function tasks. They confirmed the existence of a three-factor model in adult executive functioning consisting of three core executive functions: shifting, updating and inhibition. Shifting is the ability to shift your attention and focus between tasks, updating refers to your capability to update and monitor representations in your working memory, and inhibition points to the ability to inhibit your behavior regarding prepotent responses (Miyake et al., 2000). These different types of executive functions can operate separate from each other, fitting the proposed three-factor model, but they can also co-occur in a complex task (Miyake et al., 2000). While adult executive functions can be regarded as three separate, independent factors the early development of the executive functions may not be as differentiated. Lee, Bull and Ho (2013) found evidence that the three core executive functions (which they call: updating, inhibition and shifting) are relatively incorporated with each other early in life. Lee et al. (2013) discussed that, through development, shifting seemed to build on working memory and inhibitory processes. This finding points to an initial two-factor executive function model in childhood (Lee et al., 2013). When reaching adolescence, the three core executive functions can be distinguished as separated functions (Lee et al., 2013). In conclusion, the three core executive functions are subject to development during childhood and in the next paragraph we will discuss some other factors that influence executive functions and their development.

While there is substantial evidence that executive functions have an important impact on school performance and the development of cognitive skills, like mathematics (Fuhs et al., 2014), executive functions can be trained as well. A well-known example of a preschool intervention in which executive function are trained is “Tools of the Mind”. Tools of the Mind is an education program used in an early or preschool setting with the primary goal to enhance executive function capabilities in young children (Diamond, 2011; Moreno, Shwayder & Friedman, 2016). Tools of the Mind is based on the ideas by Vygotsky that cognitive development is shaped by scaffolding. Sanders and Sugg Welk (2005) explain scaffolding as a technique for the child to reach higher cognitive levels. Thus, scaffolding means that the teacher or adult support children when they try new activities they cannot do on their own (yet). They will let the child practice these new skills and help him or her when necessary. After a while, the child can perform the activities on his own (Diamond, 2011). Diamond, Barnett, Thomas and Munro (2007) compared Tools of the Mind as an executive function training for preschoolers with another curriculum set up by the school district. They found that children aged 4-5 years who had the Tools of the Mind executive function training had higher executive function scores than the other group. Although the primary goal of Tools of the Mind is the training of the executive functions, the training also has generalizing effects on academic achievement (Diamond, 2011). All in all, we may conclude that differences in education and educational programs can influence and even promote executive functions in children. In the next section we will discuss the Montessori educational system and explain a possible enhancing effect of Montessori education on executive functioning.

Looking at the proven training-effect by “Tools of the Mind”, different primary school programs may have a similar effect on executive function development. Around the world there are several different educational programs and systems implemented by primary schools. One specific educational system is Montessori education. Montessori education teaches their students to work on their own, plan their own time, and to regulate their own behavior. In short, it emphasizes the independence of the child and self-directed learning (Peng & Md-Yunus, 2014). The higher level of independence and self-directed learning in the Montessori classroom might make a stronger and earlier call upon executive functions compared to regular educational systems which are often based on more direct teacher-instruction. Before we go any deeper into the specific aspect of Montessori education that may have more positive effects on children's executive functions than other forms of education we will first explain the most important aspects of Montessori and her educational method.

Montessori education was founded by one of the first female medical doctors in Italy, Maria Montessori, who, after her graduation as a doctor, developed an interest in pedagogy (Schwegman,1999). Montessori constructed an educational program, called "The Method", which she used to teach children in impoverished Roman neighborhoods (Montessori, 1966; Schwegman, 1999). In her method, Montessori emphasized the phenomenon of the growth of a "helpless child" into an independent being. She stated that if we want children to actually learn something, we must not "serve" them by doing an activity for them (like dressing or feeding them) but we must teach them how to dress and feed themselves. This way, children will be raised to live an independent life and help themselves (Montessori, 1966). Regarding the educational activities, "The Method" teaches children in heterogeneous groups (meaning that children of various ages are taught in one classroom), uses special Montessori materials for educational purposes and gives instruction which is individualized or given to children within the same level groups (Lillard & Else-Quest, 2006). In short, the most important difference between Montessori and overall regular education is that in the Montessori system the student choses their own work and works individually and in regular education the student relies more on teacher instruction (Lillard et al., 2006; Peng et al., 2014).

So how may Montessori education specifically enhance executive functions? What elements may contribute to enhancing and training the executive functions within Montessori education? There are some suggestions on how this effect may come about. Initially, Montessori education may have an effect on executive function development through its basic principles. Educational systems such as Montessori in which independent learning is stimulated at an early age likely puts higher and earlier demands on children's executive function abilities and this might in turn lead them to be better at planning and regulating their own behavior. This idea is supported by a book published in 2015 by the Dutch Montessori Association that proposes a total of seven reasons to choose Montessori education for your child (Wasmann-Peters, 2015). One of these reasons is that children choose and plan their own activities and it is their own responsibility to do this planning correctly. Implicitly this suggests that children must regulate their own behavior and think about how they plan their day and activities. This aspect of Montessori education may enhance executive functioning because the children are stimulated to plan their own schedule and work for themselves.

There are some previous studies that have studied a possible difference in executive function scores between children in Montessori school programs and regular school programs. For example, Lillard (2012) studied children's executive functioning directly in Montessori school programs. She used two measures for executive functioning in her research: the "head-toes-knees-shoulders"-task and a "delay of gratification"-task. Preschool children coming from the classic Montessori program showed higher executive functions gains within a year when compared to children coming from a conventional education program. The higher gains in executive function scores were thought to be coming from the use of the specific educational materials in Montessori education (Lillard, 2012). So the Montessori educational materials may have an enhancing effect on executive functions..

While Lillard (2012) looked at the effect of Montessori education in preschool children, limited research is devoted to the effect of Montessori education on executive functions in primary school children. One study researched executive functioning in primary school children around the age of 10 with the use of a questionnaire called the Behavior Rating Inventory of Executive Functioning, or BRIEF (Bagby, Barnard-Brak, Sulak, Jones & Walter, 2012). The study used both the parent and the teacher version of this questionnaire. There were three different education types who participated in the study, a Montessori school, a Classical school and a Catholic school. The results show that although there was no difference between the scores of the three schools on the parental questionnaire, the teacher questionnaire did show a difference between the groups. The teacher scores indicated that the children in the Montessori and Catholic school had higher executive functioning skills than the Classical school. The results were based on one general total score of executive functioning (Bagby et al., 2012). So while both Lillard (2012) and Bagby et al. (2012) used tests that involve executive functioning, they did not measure the separate core executive functions (inhibition, attention flexibility and working memory). It is important to look at the possible effect of Montessori and regular education on all three separate executive functions in school-aged children because many children are currently enrolled in both Montessori as regular educational programs and gaining information about how these programs influence the development of the three separate executive functions in childhood might improve educational systems all over the world. It could be the case that a specific type of education enhances only one or two of the core executive functions so it is important to study the executive functions separately. We will also take in account the children's emotion regulation skills. Emotion regulation skills and executive functions influence each other and Schmeichel and Tang (2015) concluded that emotion regulation even builds on some specific executive functions. Therefore, the main aim of this current study will be to investigate if, due to higher demands on children's self-directed learning and the use of special, specific Montessori materials, the Montessori educational program leads to advanced development of children's core executive functions (being: inhibition, attention flexibility and working memory) and emotion regulation skills compared to children following regular education depending more on teacher and group instruction. This will be investigated by measuring and comparing the level of development of the three core executive functions of inhibition, attention flexibility and working memory (by a computer battery of tests) and emotion regulation skills (by a self-report questionnaire) between two groups of 7-8 year old children following either a Montessori or regular educational school program. For this study, we will also measure and if necessary control for potential differences in the level of educational achievement of the child and the highest academic level of the parents as an indication for parental socioeconomic status (SES).

2. Methods

2.1. Participant inclusion and recruitment

For this study we recruited a total of sixty-six 7-to 8-year old children during the academic year 2015-2016 from different schools in The Netherlands. The schools were recruited with the help of the Dutch Montessori Association and the University of Applied Sciences Saxion. In Amsterdam we tested 10 children from one classroom, in Haarlem we tested 31 children from 4 different classrooms and in Vlaardingen we tested 25 children from two different classrooms. When tested the children came from the same classroom so all children were familiar with each other. Overall, 41 of the participants came from a Dutch Montessori school (Amsterdam and Haarlem) and 25 participants had a regular Dutch education background (Vlaardingen). All tests and procedures in this research were approved by the ethical board of the Faculty of Psychology and Neuroscience at the University of Maastricht before we started recruiting schools and participants. On the first hand we tried to find a regular school in the same region as the Montessori schools (so in the neighborhood of Amsterdam and Haarlem) but unfortunately we were not able to do so in time. Luckily we found a school in time with a regular educational curriculum in Vlaardingen, which is roughly 80 kilometers from Amsterdam and Haarlem. We will further explain the procedure we used for recruiting the schools and participants in the procedure section.

2.2. Procedure

We first contacted the schools via e-mail and had further contact via telephone. If a school gave permission for participation they received a letter with full explanation of the study and the school-principal was asked to sign an informed approval for participation. This study was conducted under a "passive informed consent" meaning that the school accepted the study as a part of their school curriculum and all children could participate without separate parental permission for each child. However, to inform the parents and give them an option to withdraw their child from participation for whatever reason, the parents were sent an information letter about the study. The letter explained the study and what it entailed for the children and provided an e-mail address to which the parents could send a short e-mail to withdraw their child from participation. In total, 4 parents decided to withdraw their child from the study and those children did not participate in any of the test-sessions.

We initially tested the Montessori children and afterwards matched the regular school group with the Montessori school group on age (date of birth), gender and time of day for testing (morning or afternoon). All children first filled in the questionnaire (FEEL-KJ; for more information of this questionnaire see below) in groups of 4 to 6 children with the use of the Qualtrics computer program and afterwards performed the six computerized executive function tests (for a more detailed description of the tests, see below) in groups of 3-5 children. The questionnaire and executive function tests were either performed on the same day (with a break in between) or on separate days. We were careful to construct the groups, taking in account time of measurement and group size, in such a way that they were as equal as possible for each educational program.

For the questionnaires and tests the children were seated at one large table with the laptops placed in front of them. In all three schools the room in which the large table was situated was a "public" area for all the students and teachers of the school and was relatively calm during test-sessions. The children could see each other but not each other's laptop-screens. Although the room was calm we did ask the children to wear noise-reducing headphones in an effort to reduce whatever noise there was as much as possible. The students took about 30-50 minutes to fill in the questionnaire on a laptop with open internet connection and if they had any questions a researcher was present to answer them. If a child finished the questionnaire he or she was thanked and send back to their classroom. For the executive function tests the children were invited to the room in groups of 3 to 5 children and seated behind a laptop with a button-box. The tests were introduced and explained, and all children started all the tests at the same time. If a child finished a task before the rest of the group they were asked to take a little break and quietly wait for the rest to finish the task so they could continue to the next task together. After all the tasks were completed the children waited for each other and walked back to their classroom in a group. In total, the six executive function tasks took between 1 hour and 1 hour and 20 minutes which varied because of the amount of instruction time and the reading and response speed of the children during different tasks.

The parents received the BRIEF-questionnaire (for more information about this questionnaire, see below) via mail from the school-principal containing a short explanation about the questionnaire and the link to the digital version of the questionnaire. Filling in the BRIEF-questionnaire was valued at 15 to 20 minutes.

2.3. Scores, questionnaires and tests

School performance scores and parental SES were considered possible confounders and we asked the schools to provide us with the school performance scores of the children and asked the parents to provide us with their highest academic achievement level. We studied the emotion/self-regulation abilities of the children with the use of a self-report questionnaire. The executive functions were measured with a task battery developed at the University of Maastricht.

2.3.1. School performance scores – AVI and mathematic achievement

To assess school performance we asked the teachers and board of the schools to provide us with the reading and mathematics scores of the children. All schools tested the reading-level of the children with the use of AVI, a Dutch test. Two schools used CITO-mathematic tests to assess the mathematic scores and one school used a school method-based test. We used these scores to divide the children into groups of 1)above average, 2)average, and 3)below average scores on the mathematic- and AVI-test. We compared the Montessori participants with the regular education participants on these scores with the use of a Mann-Whitney U Test. If school performance scores significantly differed between the Montessori and Regular education group this variable would be considered a confounder for the questionnaires and EF-tasks analyses.

2.3.2. Parental SES

To assess the socioeconomic status of the parents we asked the parents to fill in with the BRIEF questionnaire what their highest academic level is. There were six options the parents could choose from, the first option reflected the lower Dutch Academic level, and the sixth reflected the highest Dutch Academic level. We compared the Montessori participants with the regular education participants on this score with the use of a Mann-Whitney U Test. If this measure of parental SES significantly differed between the Montessori and Regular education group, this variable would be considered a confounder for the questionnaires and EF-tasks analyses.

2.3.3. Emotional self-regulation and executive function questionnaires

To measure their emotion regulation skills the children filled in the Dutch version of the *“Der Fragebogen zur Erhebung der Emotionregulation bei Kindern und Jugendlichen”* (FEEL-KJ) questionnaire (Braet, Cracco, Theuwis, Grob & Smolenski, 2013). To measure the executive functions of the children in daily life, the parents of the children were asked to fill in the *“Behavior Rating Inventory of Executive Function”* (BRIEF) questionnaire (Smidts & Huizinga, 2009). All questionnaires were filled in by the use of the Qualtrics program on computers with internet-access.

2.3.3.1. FEEL-KJ – emotional regulation questionnaire

The FEEL-KJ is a questionnaire that measures the emotion regulation capacities in children (Braet et al., 2013). The children filled in the questionnaire themselves. The questionnaire asked the children how they coped with situations in which they feel angry, fearful or sad. There are 90 items in the questionnaire, 30 on how they coped when feeling angry, 30 when they felt fearful and 30 when they felt sad. In the end, the FEEL-KJ showed which emotion regulation strategies the children used to handle their negative feelings in the different situations. The FEEL-KJ differentiated three categories of emotion regulation strategies: adaptive strategies, maladaptive strategies and external regulation strategies. Using the adaptive strategies the child tried to cope in negative situations by trying to change the negative situation or divert their attention from the negative situation. When using maladaptive strategies the child did not try to give a different, more positive spin on the situation but they acted on their negative feelings by confronting others or going away to be alone. If children used external regulation strategies they relied on social support, expressing their feelings so others can see how they felt or by trying to control their emotions themselves (Braet et al., 2013). We calculated three normed scores for each child, an adaptive strategies-score, a maladaptive strategies score and an external regulation strategies score. Norming was based on the age of the child. We compared the Montessori group with the regular education group on these scores with the use of multiple Independent Samples T-Tests.

2.3.3.2. BRIEF – executive functions questionnaire

The BRIEF questionnaire is a questionnaire containing 75 items regarding the executive functioning of the child. There is both a parent as a teacher questionnaire (Smidts et al., 2009). In this current study the BRIEF was only filled in by the parents of the children. The questionnaire had several subscales: inhibition, cognitive flexibility, emotion regulation, taking the initiative, working memory, planning and organizing, orderliness and neatness, and behavior evaluation. Apart from these subscales that measured various constructs related to executive functioning, two indexes were also calculated: a behavior regulation index and a metacognition index. (Smidts et al., 2009). We calculated three age and gender-normed scores for each of the participants, one for the behavior regulation index, another for the metacognition index and lastly a total norm score. We compared the Montessori group with the regular education group on these three scores, with the use of multiple Independent Samples T-Tests.

2.3.4. Computerized Executive Functions task battery

This computerized task battery consisted of six different executive function tests. The battery took about 1 hour and 15 minutes to complete. Each of the tests will be explained below. All tests were accompanied with simple instructions and multiple breaks. To prevent possible fatigue effects there were six different orders in which the tests were presented. One group of children all received the same order of presentation and in both Montessori and the regular education group all six orders were used. All tasks were programmed in the stimulus presentation program “Presentation” and were performed by the children on a laptop. One task was performed with the computer keyboard and for the other tasks the children used a button-box with two buttons, a left button colored blue and a right button which was yellow.

2.3.4.1. Dots-task – attention switching

This task was derived from a study by Huizinga, Dolan and Van der Molen (2006). It was used to measure attention switching as a core executive function, separated from the other executive functions. The current task has multiple adaptations from the original task used by Huizinga et.al. (2006). These adaptations were: 1) all three conditions were preceded by picture instructions and 2) a practice trial with feedback (“correct” when the answer is correct, “wrong” when the answer is incorrect and “too fast” if the response is too fast), 3) after the practice trial there was no feedback anymore, 4) there was always a minimal of two dots or triangles difference between the left/right or lower/upper section of the grid during a trial, 5) the task was shorter than the one by Huizinga et.al. (2006) and 6) in the final block where the dots and triangle conditions switch, they did so every 3 trials instead of 4. During this task a 4 by 4 grid was presented on the computer screen in all conditions (three conditions, two non-switch and one switch).

In the first non-switch condition block, called the “Triangle-game”, the grid was filled with a semi-random amount of small triangles and the children had to choose which side had the most triangles, the left or the right, and had to push the left or right button on their button box. In the next non-switch condition block, the “Dots-game”, the grid was filled with small dots and now the children had to decide if the top or the bottom of the grid had more dots. The left button corresponded with the answer “top” and the right button with the answer “bottom”. In both non-switch conditions there were 5 practice trials with feedback and 30 test-trials. Before we continue our explanation of this task we should mention that though the Dots-condition is presented as a non-switch condition, prior experience with this task showed us that this condition is experienced by the participants as a switch-block. Therefore, we only used the Triangle-condition as non-switch condition. Because the Dots-condition was already present in the task, we still let the participants perform it, but we did not take the results from this condition into further account.

The last block was a switch condition, in which the grid was filled with either triangles or dots and the participants had to push the button corresponding to the correct answer in the triangle- or dot-condition. This switch-condition consisted of 12 practice trials and 93 test-trials from which 63 are non-switch and 30 are switch trials. Each trial started with the presentation of an empty screen for 1000 milliseconds (ms), followed by the presentation of an empty grid for a duration of 900-1100ms. After this the grid was filled by a number of dots or triangles and this image stayed on the screen until an answer was recorded. Between each trial there was a 500ms inter stimulus interval in all blocks (so both the practice as experimental blocks). For later analyses the mean percentage of hits and mean reaction times in the non-switch and switch conditions were calculated for each participant. We performed a total of two two-by-two mixed analysis of variance (ANOVA) to check for a possible effect of Education type (Montessori versus regular education) on the difference in the percentage of hits and in reaction times (only the reaction times recorded during the correctly answered trials in the Dots-task). We compared the non-switch condition with the switch condition to look for differences in switch costs between the tasks (this is the within-subjects factor “Switchcosts”). Also, a comparison was made between Montessori and the regular education group to see if there was an effect of Education type on the Dots-task performance (this is the between-subjects factor "Education type"). Lastly, we tested if there was an interaction effect between Switchcosts and Education type. One two-by-two analysis of variance studied the possible difference in the percentage of hits on the DOTS task, and the other looked into the possible difference in reaction times between both groups on the DOTS task.

2.3.4.2. Attention switching task (AST) – attention switching

The AST task was the second task in this study that attempts to measure attention switching as a core executive function. This task has previously been used in a study by Markus and Jonkman (2007), who based it on the task used by Cepeda, Kramer and Gonzalez de Sather (2001). The current task was an adapted version of the tests used by Cepeda et al. (2001) and Markus et al. (2007). The AST consisted of three conditions: two non-switch consisting of 32 trials each, and one switch condition consisting of 64 trials. This task was also performed with the use of the button box. In the first condition the phrase "WHICH NUMBER" was presented on the screen, after which the numbers 1 or 3 were presented. The numbers could be presented in single form (1 or 3) or in a row of three (111 or 333) and in all three conditions each stimulus had a 25% chance to appear. The left button on the button box corresponded to the answer "1" and has to be pressed when either the stimulus "1" or "111" was presented. The right button corresponded to the answer "3" and had to be pressed when either the stimulus "3" or "333" was presented. In the second condition the phrase "HOW MANY NUMBERS" was presented and the children now had to push the left button corresponding to the answer "one digit" if they detected a single digit on the screen (1 or 3) and the right button, which corresponded to the answer "three digits", if they detected three digits (111 or 333). In the third and final condition the phrases "WHICH NUMBER" and "HOW MANY NUMBERS" were presented in a random order and the children had to respond accordingly to the condition. All three conditions were preceded by practice-trials and during the whole task feedback was given when a wrong button was pushed ("wrong") or when the response was too slow ("faster"). In the case of a correct response, no feedback was given. In all three conditions each trial started when the instruction cue appeared ("WHICH NUMBER" or "HOW MANY NUMBERS") and after 200ms the actual stimulus was presented. Both the instruction cue and the stimulus remained on the screen until a response was recorded. When no response was detected the program would interrupt the trial after 3 seconds. For later analyses the mean percentage of hits and mean reaction times in the non-switch (average score computed across the two non-switch AST blocks) and switch conditions were computed for each participant. A total of two two-by-two mixed ANOVAs were used to check for a possible effect of Education type (Montessori versus regular education) on the difference in the percentage of hits and in reaction times (only the reaction times recorded during the correctly answered trials in the AST-task). We compared the means of the AST non-switch conditions with the AST switch conditions means to look for differences in switch costs between the tasks (this is the within-subjects factor “Switchcosts”). Also, a comparison was made between Montessori and the regular education group to see if there was an effect of Education type on Dots-task performance (this is the between-subjects factor “Education Type”). Lastly, we tested a possible interaction effect between Switchcosts and Education type. One two-by-two analysis of variance studied the possible difference in the percentage of hits on the AST task, and the other looked into the possible difference in reaction times between both groups on the AST task.

2.3.4.3. N-Back task – working memory

Based on a study by Schleepen and Jonkman (2010) the N-Back is a task that sets out to measure the core executive function “working memory”. Different from the task used by Schleepen et al. (2010) the current task consisted of two conditions instead of 3. In both conditions letters were presented (A, B, F, G, H, K, L, S, T, W, X, Z) on the screen one at a time. In the “1-back”-condition the children had to push the right button on their button box if they detected two of the same letters in a row (e.g. W-W). In the second condition, the “2-back”-condition, the children had to push the right button if they detected two of the same letters in a row but only if there was another letter in between them (e.g. Z-W-Z). Both conditions were preceded by a practice block consisting of 20 practice stimuli. Because the “2-back”-condition was more challenging to understand it was practiced twice (2 practice blocks) if the percentage correct in the first 2-back practice trial was below 70%. During practice, feedback was given. The word “wrong” appeared on the screen when an error was detected, and the word “missed” was presented when a target-event was not answered. After that there were 2 test-blocks of each conditions alternating each other (test-trial order: 1-back, 2-back, 1-back, 2-back). The test-blocks did not give feedback after a wrong answer or a missed target. Each experimental test-block consisted of 50 stimuli, adding up to 200 experimental test-stimuli in total. All experimental task-blocks had the same target frequency of 36% and each trial lasts 2000ms. In the 2000ms the stimulus was presented for 500ms followed by a inter stimulus interval of 1500ms. All blocks, both practice and experimental, were preceded by an instruction screen with information about the condition (1-back or 2-back). For later analyses the mean percentage of hits and mean reaction times in the 1-back and 2-back conditions were computed for each participant. A total of two two-by-two mixed ANOVAs were used to check for a possible effect of Education type (Montessori versus regular education) on the difference in the percentage of hits and in reaction times (only the reaction times recorded during the correctly answered trials in the N-Back task). We also compared the 1-back condition with the 2-back condition to look for differences in working memory load between the tasks (this is the within-subjects factor “Working memory load”). And a comparison was made between the Montessori and the regular education group to see if there was an effect of Education type on N-back performance (this is the between-subjects factor “Education Type”). Lastly, we tested for an interaction effect between Working memory load and Education type. One two-by-two analysis of variance studied the possible difference in the percentage of hits on the N-back task, and the other looked into the possible difference in reaction times between both groups on the N-back task.

2.3.4.4. Flanker task – response inhibition and interference control

The Flanker task was developed by Eriksen & Eriksen (1974). The task attempts to measure response inhibition and interference control. In this study it was used to measure the single executive function “inhibition”. In the task three letters were presented on the screen. The participants were asked to only respond to the letter in the middle, known as the “target letter”. If that target letter was a “B” or “H” they had to press the left button on their button-box and if the target letter was a “F” or “T” they had to press the right button. The target letter was always accompanied by two flanking letters: the “B”, “H”, “F”, or “T”. The flanking letters were always identical to each other. There were three possible conditions during this task. The first was the stimulus-response congruent (SRC) condition in which the center, target letter was identical to the flanking letters. The second was the stimulus incongruent (SI) condition in which the target letter was surrounded by flanking letters which corresponded to the same button (so “BHB” or “FTF” etc.) so while there is no response conflict, the flanking letters are incongruent to the target letter. The third and last condition was both a stimulus and response incongruent (SRI) condition in which the flanking letters were 1) different from the target letter in the center and 2) corresponding to the other button (so “FHF” or “BTB”). In this last condition there was both a stimulus as a response conflict. In total, 144 trials were presented in three blocks (each block consisted of 48 trials). In each block all three conditions were presented in a random order. Every trial started with the appearance of the two flanking stimuli for 200ms. After this the target stimulus appeared and all three letters were presented on the screen for another 700 ms. A 500ms inter-stimulus interval separated each trial from the next. The possible response window was 1400ms in total for each trial. The three experimental blocks were preceded by one practice block with 18 trials with feedback ("correct" when the answer is correct; "wrong" when the answer is not correct; "faster" if the answer is too slow). During the three experimental blocks, no feedback was provided anymore. For later analyses the mean percentage of hits and mean reaction times in the stimulus response congruent and stimulus response incongruent conditions were computed for each participant. A total of two two-by-two mixed ANOVAs were used to check for a possible effect of Education type (Montessori versus regular education) on the difference in the percentage of hits and in reaction times (only the reaction times recorded during the correctly answered trials in the Flanker-task). We compared the results from the SRC condition with the results from the SRI condition to look for differences in between the tasks (this is the within-subjects factor "response conflict"). Also, a comparison was made between Montessori and the regular education group to see if there was an effect of Education type on Flanker-task performance (this is the between-subjects factor "Education Type"). Lastly, we tested if there was an interaction effect between response conflict and Education type. One two-by-two analysis of variance studied the possible difference in the percentage of hits on the Flanker task, and the other looked into a possible difference in reaction times between both groups on the Flanker task.

2.3.4.5. Running span – working memory

The running span is a task that tries to measure the sole executive function "working memory", and more specifically, the working memory process of "updating". The task that we used in this study resembled the digit spans used in a study by Jahanshahi, Saleem, Ho, Fuller and Dirnberger (2008). Jahanshahi et al. (2008) aimed to investigate the running digit span as a measure of working memory capacity. The task in this current study was conducted with the use of the actual computer key board. The participants were presented with a row of digits, each digit was presented one at a time for 1 second followed by an interval of 500ms between all numbers. The sequences varied in length, there were a total of three set sizes: 5 digits, 6 digits and 7 digits. The participants had to remember the exact sequence of the digits. When all the digits of one set were presented the participants were shown the sequence with the last 3 digits missing. The participants than had to fill in the last 3 digits of the sequence and press the "enter"-button as soon as possible. In between the sequences there were breaks and before the start of each sequence the participants were presented with a text: "ready for the next sequence? Press the space-key". After they pressed the space-key, the sequence immediately started. Before the experimental-trials there were two instruction texts presented that explained the task, followed by two practice trials with feedback (either “Good!” when the sequence was completed correctly, or “wrong” followed by the correct completed sequence when the answer was not correct). During the experimental trials there was no feedback when a sequence was completed. In total, 12 sequences were presented. Four trials consisting of 5 digits, four consisting of 6 and again four consisting of 7 digits. The sequences were intermixed in a random order so the participants did not know the length of each sequence beforehand. For later analyses the mean proportion of correct recalled series of digits (in total and for each individual set size) were computed for each participant. A comparison was made with the use of multiple Independent Samples T-Tests, between the Montessori group and the regular education group on the four scores (the total score and the three separate set size scores) to see if there was a difference between the groups on working memory span measured by the running span.

2.3.4.6. Go-NoGo task – response inhibition

A task resembling the Go-NoGo task of this current study has been used previously in an fMRI-study by Casey et al. (1997). The Go-NoGo task is frequently used to measure response inhibition. During the task in this current study letters were presented in the middle of the computer screen. The letters were presented one by one and in a random order. The participants had to push the right button on their button-box as fast as they could each time they saw a letter (the go-stimulus) with the exception of one, the letter “X” (the no-go stimulus). So, they had to inhibit themselves and not push the button when they saw the letter “X”. There was one practice block consisting of 15 trials with feedback (the word “wrong” was presented if the participant failed to inhibit a response). After this practice block there were two experimental blocks without feedback and each consisted of 100 trials (so 200 experimental trials in total). Between each block there was a pause. Each stimulus was presented for 500ms, followed by a inter stimulus interval of 700ms thus making each trial 1200ms in total. For later analyses the mean percentage of hits and mean reaction time over all the correctly answered trials were computed for each participant. A comparison was made, with the use of multiple Independent Samples T-Tests, between the Montessori group and the regular education group on the two scores (the percentage of hits and mean reaction time) to see if there was a difference between the groups on inhibition measured by the CPT-X.

2.4. Statistical analyses

The Statistical Package for the Social Sciences (SPSS) version 21 was used to analyze the collected data. First we calculated all the subscale-scores for the FEEL-KJ and BRIEF questionnaires in SPSS. Then all the raw data from all the EF-tests was transferred to SPSS for further analyses. For all tests a significance level of p<.05, two tailed was adopted

To see if there were any differences between the groups on academic achievement scores and parental SES, three Mann-Whitney U-tests were used. For the FEEL-KJ and BRIEF we planned to perform three Independent Samples T-Tests for each questionnaire.

For all EF-tasks statistical analyses will be performed on the dependent measures mean percentage of hits (reflecting accuracy) and mean reaction time (on correctly responded trials), except for the running span task where only the span measure will be analyzed as dependent variable (which is expressed in the proportion of correctly responded trials). For the CPT-X the dependent variable is the mean proportion of correctly inhibited responses on NoGo-trials. For four of the EF-Tasks a total of two 2 x 2 ANOVAs were used per task. These four EF tasks were: the DOTS-task, the AST-task, the N-Back task and the Flanker Task. To analyze the results from the two remaining EF-tasks, the Running Span and CPT-X task, multiple Independent Samples T-Tests were used to compare the group-means.

.

3. Results

3.1. Group formation and pre-analyses

To create equally sized groups and prevent age and gender differences that could have influenced EF-scores, we matched 24 Montessori education children with 24 regular education children on age and gender. An independent samples T-Test showed us that there was no significant difference in age in months between the groups (Montessori: M=95,17, SD=3,45; regular: M=95,26, SD=3,32); t(46)= -0,10, p= 0.93. Gender was matched one on one, so there was no difference in gender between the groups.

Next we checked our data for below chance level performers. For this study, below chance level performers are defined as participants who had an accuracy score of below 50% in the easiest conditions of the EF tasks. During these easiest conditions all children should have easily reached a score above chance level if they understood all the instructions and were seriously performing the task. These easiest conditions were: the 1-back condition in the B-Back task, the congruent condition in the Flanker Task, the non-switch condition in the attention switching tasks (for the DOTS task and AST task) and Go-accuracy in the CPT-X task. The Running span was not checked for under chance level performers because this task measures capacity and not hit-rate. Identified below chance performers on these five EF-tasks were excluded from further analyses for that specific task because this indicated that, despite thorough instruction and effort, the child either did not understand the task or was not motivated enough to perform sufficiently. Two EF tasks had below chance performers: 11 in the Flanker task and 1 in the N-back.

After this, all EF and questionnaire data was checked for outliers. We used the "IQRx1,5 rule" to determine if extreme scores were in fact outliers (Moore, McCabe & Craig, 2012). The "IQRx1,5 rule" is the interquartile range (the difference between the first and third quartile score in the sample) times 1,5 and the result of this calculation is added to the third quartile score and subtracted from the first quartile score. An extreme score below this "first quartile – IQRx1,5" is considered an outlier in the lower range, and an extreme score above the "third quartile + IQRx1,5" is considered an outlier in the upper range (Moore et al., 2012). For the EF-tasks, we again checked the easiest conditions of each task for outliers (see above). This time we also checked the Running Span accuracy for outliers. The CPT-X task had one outlier, the DOTS task had four outliers and the AST task had two outliers. Next we checked the questionnaires for outliers. In the FEEL-KJ we checked the total norm score of each of the strategies for outliers with the "IQRx1,5 rule". The adaptive strategies scores had no outliers, but the maladaptive and external strategies scores each had one. We also checked the BRIEF norm scores for the behavior regulation index, the Metacognitive index and the total score for outliers with the use of the "IQRx1,5 rule". None of these scores were labeled as an outlier according to this rule. Further analyses for the EF tasks and questionnaires was performed without the specific outliers.

Next we performed the different analyses for the possible confounders (school performance scores and parental SES), the questionnaires and each of the EF-task. The number of children included in the different questionnaires and EF-analyses are given in the various summary tables (1.1, 1.2, 1.3 and 1,4). The amount of children and parents included in the possible confounders are given in the text below. Next we will give the results for each possible confounder, questionnaire and EF-task in separate sections.

3.2. School performance scores and Socioeconomic Status (SES)

Before performing each analysis, the assumptions for each of the used parametric analysis were checked. Though most of the assumptions were met, our sample size is very small. We performed multiple Mann-Whitney U tests to see if there is any difference between the Montessori and regular education group on school performers scores and parental SES. If there were significant differences present, we would take these variables into account for further analyses.

3.2.1. School performance scores – AVI and mathematic achievement

As described we divided the children into groups of 1)above average, 2)average and 3)below average AVI and mathematic scores. We obtained the AVI-score of all the 24 Montessori and 24 regular education children from the schools, but for the mathematic scores one matched participant from the regular education group had no test result. So for the mathematic scores we obtained the scores for 24 Montessori children and 23 regular education children. To see if there was a difference between the groups on school performance scores we used two Mann-Whitney U tests. The results on the AVI scores show that there was no difference between the Montessori (MD=1,00) and the regular education group (MD=2.00), U=228,0, *p*=0.18. The results for the mathematic scores also showed that there was no difference between the groups (Montessori MD=1,00, regular MD=1,00; U=256,0, *p*=0.61). Because there were no significant differences between the groups on school achievement scores, we did not include these variables as covariates in our analyses.

3.2.2. SES of the parents

To assess the socioeconomic status of the parents of the children we asked the parents who filled in the BRIEF to put down their highest academic level. There were six possible answers, reflecting the Dutch academic levels. Twelve parents from the Montessori group and nine parents from the regular education group filled in the questionnaire. A Mann-Whitney U test was used to see if there was any difference in the highest academic level of the parents. The results show that there was no significant difference between Montessori (MD=6,00) and regular (MD=5,00) education participants on SES, U=32,0, *p*=0.09. Because there were no significant differences between the groups on parental SES, we did not use this scores as a covariate in our analyses.

3.3. Results Questionnaires

Before performing each analysis, the assumptions for each of the used parametric analysis were checked. Though most of the assumptions were met, our sample size is very small. So normality of distributions was not always completely met, but data transformation was not performed due to robustness of ANOVA to slight deviances from normality An overall summary of the results from the questionnaires can be found in table 1.1 (FEEL-KJ) and 1.2 (BRIEF).

3.3.1. FEEL-KJ – emotion regulation questionnaire

Since not all children completed the entire FEEL-KJ questionnaire, in table 1.1 we indicated the number of Montessori and regular education participants that completed all the items of the questionnaire that were necessary to compute the different strategy scores. We analyzed the age-appropriate norm score depicted from the FEEL-KJ manual (Braet et al., 2013). Norming was based on age (8-12 years of age). Afterwards we performed a total of three Independent Samples T-Test to check for a significant difference between the scores of the children coming from a Montessori education background and those from a regular education background. For each of the three Independent Samples T-Test the Levene's Test for Equality of Variances was not significant, thus we assumed equal variances when interpreting the results. No significant differences between the two educational groups were found for the adaptive strategies norm score, the maladaptive strategies norm score and the external strategies norm score (see table 1.1. for the means, SDs and t-test statistics).

Table 1.1. The Mean-scores, standard deviations (SD) and T-test results of both groups on the dependent subscales of the FEEL-KJ questionnaire.

Score	**Mean Montessori**	**SD Montessori**	**Number of participants Montessori**	**Mean regular**	**SD regular**	**Number of participants Regular**	**df**	**t-score**	***p*-value**
Adaptive strategies norm score	50,21	11,93	19	50,57	8,65	14	31	-0,096	0.924
Maladaptive strategies norm score	49,50	6,40	20	51,60	6,75	15	33	-0,939	0.355
External strategies norm score	50,30	8,25	20	52,06	6,61	17	35	-0,707	0.484

3.3.2. BRIEF – executive functions questionnaire

When checking the data, it became apparent that two questions were not included in the Qualtrics-version of the questionnaire. Luckily this was not a big problem because the BRIEF allowed a total of two missing items for each of the subscales. The two missing items belonged to the subscale "inhibition" and the BRIEF-manual described that these missing scores can be replaced by a score of “1”, which resembles the answer “never”. We did this for all participants. Also, one of the participant’s parents did not answer one question from another subscale (“taking the initiative”) and this missing score was also replaced by a score of “1”. Afterwards we performed a total of three Independent Samples T-Tests to check for any differences between the educational types on the different scores. Because all of the performed Independent Samples T-Test had a non-significant Levene’s Test for Equality of Variances value, we assumed equal variances when interpreting the BRIEF-results. No significant differences between the two educational groups were found for the behavior regulation norm score, the metacognition index norm score and the total norm score (see table 1.2. for the means, SDs and t-test statistics).

Table 1.2. The Mean-scores, standard deviations (SD) and T-test results of both groups on the dependent subscales of the BRIEF questionnaire.

Score	Mean Montessori	SD Montessori	Number of participants Montessori	Mean regular	SD regular	Number of participants Regular	df	t-score	*p*-value
Behavioral regulation index normed score	37,08	5,60	12	41,0	7,53	9	19	-1,370	0.187
Metacognition index normed score	39,50	6,43	12	44,67	9,58	9	19	-1,481	0.155
Total normed score	37,33	5,99	12	42,56	8,58	9	19	-1,647	0.116

3.4. Results from the Computerized Executive Functions task battery

In the sections below we will give the results for each EF task separately. Before each analysis, the assumptions for each of the used parametric analysis was checked. Though most of the assumptions were met, we are analyzing only a small sample size. For a summary of the results from all EF tasks see tables 1.3 and 1.4.

3.4.1. Dots-task – attention switching

The 2 x 2 ANOVA performed on the percentage of hits (performance accuracy) in the DOTS task revealed a significant main effect of switch costs on the difference between the non-switch and switch percentage of hits, indicating that the percentage of hits was higher in the non-switch than the switch condition, signifying the presence of switch costs. There was no significant main effect of Education type on the percentage of hits in the DOTS task. We also found no significant interaction effect between Education type and Switching for the percentage of hits. For reaction time recorded during the correctly answered trials only, the analysis showed a significant main effect Switching, indicating that the reaction time was higher in the switch than in the non-switch condition. There was no significant main effect of Education type on the reaction times, and also the interaction between Education type and Switching on reaction time was not significant. See table 1.3. for the means, standard errors and 2 x 2 ANOVA statistics.

3.4.2. Attention switching task (AST) – attention switching

The 2 x 2 ANOVA on percentage of hits showed a significant main effect of Switching indicating that the percentage of hits was higher in the non-switch than the switch condition, signifying the presence of switch costs. We found no significant main effect of Education type on the percentage of hits from the AST task. As for an interaction effect between switch costs and Education type we found no significant result. The reaction time analysis also revealed a main effect of Switching, indicating that the reaction time was higher in the switch than the non-switch conditions. There was no significant main effect of Education type. A marginal significant Switching x Education-type interaction effect was found, indicating that during the non-switch condition, the Montessori group was faster, but during the switch condition, the regular education group performed faster. See table 1.3. for the means, standard errors and 2 x 2 ANOVA statistics.

3.4.3. N-Back task – working memory

The 2 x 2 ANOVA on difference in percentage of hits between the 1-back and 2-back task showed a significant main effect for working memory load. There was no significant main effect of Education type on the percentage of hits in the N-Back. And there was also no significant interaction effect between working memory load and Education type for the percentage of hits. For reaction time the 2 x 2 ANOVA analysis showed that there was a main effect of working memory load on reaction time. There was no significant main effect of Education type on reaction time, and also no significant interaction effect between the working memory load and Education type. See table 1.3. for the means, standard errors and 2 x 2 ANOVA statistics.

3.4.4. Flanker task – response inhibition and interference control

The 2 x 2 ANOVA for percentage of hits showed that here was a main effect of response conflict on the percentage of hits. There was no significant main effect of Education type on the percentage of hits in the Flanker task, and also no significant interaction between the response conflict and Education type. The 2 x 2 ANOVA for the differences in reaction time found a significant main effect of response conflict on the reaction time. There was no main effect of Education type on the reaction time in the Flanker task. There was a borderline significant response conflict x Education-type interaction effect (p=.051), indicating that the Montessori group was faster during the SRC condition, but the regular education group was faster during the SRI condition in this study's sample. See table 1.3. for the means, standard errors and 2 x 2 ANOVA statistics.

3.4.5. Running span – working memory

We performed four Independent Samples T-Tests to check for differences in performance on the running span between the Education groups. We checked the Levene's Test for Equality of Variances before interpreting the results for each Independent Samples T-Test, and all Levene's tests were not significant so we assumed equal variances. We performed an Independent Samples T-Test to check for differences on the entire task and this analysis showed that there was no significant difference between the groups on the complete task. After this analysis we also performed three separate Independent Samples T-Tests to check for any differences between the groups on the different running span set sizes of 5, 6 or 7 numbers. These three Independent Samples T-Tests showed us that there was no significant difference between the groups on the trials with different set sizes. There was a marginally significant effect for set-size six (p= .08), but again this was not a statistically significant difference between the groups. See table 1.4. for the means, SDs and t-test statistics.

3.4.6. Go-NoGo task – response inhibition

We performed three separate Independent Samples T-Tests to check for any differences between the groups on the Go-NoGo (CPT-X) task. First we checked the Levene's Test for Equality of Variances before interpreting the results for each independent samples T-Test. All the Levene's tests were not significant so we assumed equal variances. The first Independent Samples T-Test compared both groups on their ability to inhibit a prepotent response and showed no difference in the scores for Montessori and regular education. We also performed an Independent Samples T-Test to check for difference between the groups on the percentage of hits on the Go-stimuli and again there was no difference between the groups. The last Independent Samples T-Test was performed to check for any differences between the groups on the reaction time of the correctly answered “Go”-stimuli and there was no significant difference between the groups on this reaction time. See table 1.4. for the means, SDs and t-test statistics.

Table 1.3. The two-by-two ANOVA results for the Dots, AST, N-Back and Flanker tasks with the Mean-scores and standard errors (Mon.= Montessori group; Reg.= Regular education group).

Task	**Type of analysis**	**Amount of Participants Mon.**	**Amount of Participants Reg.**	**Effect**	**Mean**	**Standard error**	**df**	**F-score**	***p*-value**
Dots	Dots task main effect Switchcosts – percentage of hits	21	23	Non-switch	92,64	0,89	(1, 42)	132,21	<0.001
				Switch	70,09	2,14			
	Dots task main effect Education type – percentage of hits	21	23	Mon.	83,73	1,90	(1, 42)	488,80	0.08
				Reg.	79,01	1,81			
	Dots task interaction effect “Switchcosts x Education type” – percentage of hits	21	23	Non-switch x Mon.	94,13	1,29	(1, 42)	0,80	0.38
				Switch x Mon.	73,32	3,09			
				Non-switch x Reg.	91,16	1,23			
				Switch x Reg.	66,85	2,96			
	Dots task main effect Switchcosts – reaction time correct trials	21	23	Non-switch	885,80	38,75	(1, 42)	132,28	<0.001
				Switch	1336,96	36,51			
	Dots task main effect Education type – reaction time correct trials	21	23	Mon.	1122,71	46,46	(1, 42)	11283, 88	0.73
				Reg.	1100,04	44,40			
	Dots task interaction effect “Switchcosts x Education type” – reaction time correct trials	21	23	Non-switch x Mon.	873,38	56,03	(1, 42)	1,59	0.21
				Switch x Mon.	1373,05	52,79			
				Non-switch x Reg.	899,22	53,54			
				Switch x Reg.	1300,87	50,44			
AST	AST task main effect Switchcosts – percentage of hits	22	24	Non-switch	87,35	1,13	(1, 44)	87,59	<0.001
				Switch	70,39	2,12			
	AST task main effect Education type –	22	24	Mon.	79,31	2,07	(1, 44)	0,09	0.76

	percentage of hits			Reg.	78,43	1,99			
	AST task interaction "Switchcosts x Education type" – percentage of hits	22	24	Non-switch x Mon.	86,90	1,63	(1, 44)	0,98	0.33
				Switch x Mon.	71,73	3,06			
				Non-switch x Reg.	87,81	1,56			
				Switch x Reg.	69,06	2,93			
	AST task main effect Switchcosts – reaction time correct trials	22	24	Non-switch	773,59	19,26	(1, 44)	73,94	<0,001
				Switch	1315,07	68,82			
	AST task main effect Education type – reaction time correct trials	22	24	Mon.	1059,84	57,09	(1, 44)	0,15	0.70
				Reg.	1028,82	54,66			
	AST task interaction "Switchcosts x Educational type" – reaction time correct trials	22	24	Non-switch x Mon.	730,57	27,82	(1, 44)	3,46	0.07
				Switch x Mon.	1389,10	99,41			
				Non-switch x Reg.	816,61	26,64			
				Switch x Reg.	1241,04	95,18			
N-Back	N-Back main effect Working memory load – percentage of hits	23	24	1-back	0,83	0,02	(1, 45)	101,29	<0.001
				2-back	0,60	0,03			
	N-Back main effect Education type – percentage of hits	23	24	Mon.	0,73	0,03	(1, 45)	0,98	0.33
				Reg.	0,69	0,03			
	N-Back interaction "Working memory load x Education type" – percentage of hits	23	24	1-back x Mon.	0,84	0,03	(1, 45)	1,06	0.31
				2-back x Mon.	0,63	0,04			
				1-back x Reg.	0,82	0,03			
				2-back x Reg.	0,57	0,04			
	N-Back main effect Working memory load – reaction time correct trials	23	24	1-back	683,58	14,88	(1, 45)	41,52	<0.001
				2-back	801,94	20,31			
	N-Back main effect Education type –	23	24	Mon.	743,83	21,79	(1, 45)	0,01	0.94

	reaction time correct trials			Reg.	741,69	21,33			
	N-Back interaction “Working memory load x Education type” – reaction time correct trials	23	24	1-back x Mon.	681,87	21,26	(1, 45)	0,09	0.76
				2-back x Mon.	805,79	29,02			
				1-back x Reg.	685,29	20,81			
				2-back x Reg.	798,09	28,41			
Flanker	Flanker main effect Response conflict – percentage of hits	18	19	SRC	77,56	1,80	(1, 35)	31,88	<0.001
				SRI	67,36	2,06			
	Flanker main effect Education type – percentage of hits	18	19	Mon.	74,14	2,44	(1, 35)	0,97	0.33
				Reg.	70,79	2,38			
	Flanker interaction “Response conflict x Education type” – percentage of hits	18	19	SRC x Mon.	78,36	2,57	(1, 35)	0,94	0.34
				SRI x Mon.	69,92	2,96			
				SRC x Reg.	76,76	2,51			
				SRI x Reg.	64,81	2,87			
	Flanker main effect Response conflict – reaction time correct trials	18	19	SRC	618,17	13,17	(1, 35)	77,32	<0.001
				SRI	678,38	13,25			
	Flanker main effect Education type – reaction time correct trials	18	19	Mon.	646,38	18,29	(1, 35)	0,02	0.88
				Reg.	650,17	17,80			
	Flanker interaction “Response conflict x Education type” – reaction time correct trials	18	19	SRC x Mon.	609,35	18,88	(1, 35)	4,09	0.051
				SRI x Mon.	683,41	18,98			
				SRC x Reg.	626,99	18,38			
				SRI x Reg.	673,35	18,48			

Table 1.4. The Mean-scores, standard deviations (SD) and T-test results of both groups on the dependent subscales of the running span and CPT-X tasks (Mon.= Montessori group; Reg.= Regular education group).

Effect	**Mean Mon.**	**SD Mon.**	**Number of participants Mon.**	**Mean Reg.**	**SD Reg.**	**Number of participants Reg.**	**df**	**t-score**	***p*-value**
Running Span – complete task	0,38	0,24	24	0,29	0,2	24	46	1,477	0.15
Running Span – set size of 5 numbers	0,48	0,34	24	0,39	0,27	24	46	1,070	0.29
Running Span – set size of 6 numbers	0,44	0,36	24	0,27	0,29	24	46	1,770	0.08
Running Span – set size of 7 numbers	0,24	0,25	24	0,21	0,25	24	46	0,432	0.67
Go-NoGo – inhibition prepotent response	0,41	0,16	24	0,46	0,20	23	45	-1,030	0.31
Go-NoGo – Percentage of hits on Go-stimuli	0,93	0,048	24	0,92	0,044	23	45	0,799	0.43
Go-NoGo – Reaction time on Go-stimuli	455,19	58,24	24	451,04	53,47	23	45	0,254	0.80

4. Discussion

In this study the possible effect of Education type on executive functions in 7-8 year old primary school children has been analyzed. Two types of education were compared in this study, the Montessori education system and the Regular Dutch education system based on teacher instruction. No statistically significant differences were found between the groups on executive functioning and emotion-regulation behavior. We carefully matched the groups on age and gender and checked for potential group differences in school performance scores and parental SES (based on highest parental academic achievement) that could confound the results, and no statistically significant differences between the groups on these control factors have been found either. So these control factors could not have influenced our results. Looking a bit closer at our results, one marginally significant and one borderline significant interaction effect was found for the AST and Flanker Task respectively. The marginal significant effect came from the interaction between the Switchcosts reaction time and Education type in the AST task (p=0,07). In the non-switch condition, the Montessori group was faster. But in the switch condition, the regular education group was faster. The Flanker task showed an borderline significant interaction effect between response conflict reaction time and Education type (p=0,051). In the stimulus response congruent condition the Montessori group was faster. But in the stimulus response incongruent condition the regular education group was faster. Significant results could have implied that the different Education types had an effect on the reaction time in the AST and Flanker tasks. But for this study we can only conclude that these interactions were present in our studied sample and not in the population because they were not statistically significant.

So, overall our results did not confirm our main-hypothesis that children coming from a Montessori education background have better executive functioning capacities and emotion regulation skills than children coming from a regular education background. We will first look a bit closer at our results concerning emotion-regulation skills measured with the FEEL-KJ. In our introduction we concluded that emotion-regulation and executive functions build on each other (Schmeichel et al. 2015). We did not find a difference in FEEL-KJ performance between Montessori and regular education. This means that children from both educational backgrounds used the same coping strategies in situations in which they felt angry, fearful or sad. Because we discussed that emotion regulation and executive functions are linked (Schmeichel et al. 2015), it seems logical that when one is not significantly different between both groups, the other also isn’t. However, Montessori or regular education could have influenced emotion regulation in other ways that only via executive functioning. But we did not find any differences between the groups on this construct so this has not been the case. Our study's main focus is on the possible effect of Education type on executive functions, so we will now discuss our EF-results further and compare them to the results of previous studies.

We based our main EF-hypothesis on the studies from Lillard (2012) and Bagby et al. (2012) who did find better EF skills in children following Montessori education compared to the regular, classical school program. As discussed in the introduction, there were multiple differences between the current study and the study done by Lillard (2012), amongst others the used EF-tests and manner of EF-scoring. A difference worth mentioning is that Lillard (2012) tested the children at the beginning and end of the schoolyear, so a gain in executive functioning was measured, this current study was a single measurement of executive functioning and we did not take into account possible differences in EF-gains between the groups. However, the most important difference is that Lillard (2012) conducted her research on a group of preschool children and we studied older, school-aged children. An explanation for the differences in the outcomes of Lillard's (2012) and our study may be that the positive effect of Montessori education on executive functions is only visible in preschool children. It could be that with overall cognitive development and growth the noticeable differences in the effect of education type on executive functioning diminishes. Also, Lillard (2012) had a much larger sample size than the current study, so the study by Lillard (2012) has more power in general. Considering Lillard's (2012) finding that preschool Montessori children have higher gains in EF scores, the pre-school Montessori program could have a training effect like the "Tools of the mind" curriculum for preschool children as explained by Diamond (2011).

Other than Lillard (2012), Bagby et al. (2012) used an overlapping EF-measure as in this study: the BRIEF questionnaire. An important difference is that Bagby et al. (2012) used both the teacher and parental questionnaire, and we only used the parental questionnaire because of time constraints. As in the current study, Bagby et al. (2012) also studied school-aged children, but their sample size was much larger than our current study's sample size. Bagby et al. (2012) found a difference between Montessori education and Classical education on the teacher BRIEF-questionnaire. However, a Catholic education program scored equally to the Montessori program. Also, the differences were only visible in the teacher BRIEF-questionnaire, but comparable to the present study, not in the parental BRIEF-questionnaire. Perhaps the Montessori school and Catholic school had a lot of similarities in their educational program but it has to be noted that the differences were only visible in the teacher BRIEF-questionnaire. It could be that the teachers from the different schools had different standards. While Bagby et al. (2012) did find a difference in children's daily executive functioning in the classroom between Montessori and Classic education as shown by the teacher-BRIEF, the parents of the same children did not note such differences in the home environment. Considering these results, we must be cautious when concluding that the differences in the study by Bagby et al. (2012) are a possible "training effect" of Montessori education on executive functioning, especially because the parental-BRIEF questionnaire did not find a difference between the three groups and because the Catholic school program scored equally to the Montessori program.

From the studies by Diamond et al. (2007) and Diamond (2011) it became clear that executive functions can be trained. The main hypothesis of this current study implemented that Montessori education had a similar training effect on executive functioning like "Tools of the Mind". But as our results showed, there were no differences between Montessori and regular education, so there was no training effect of Montessori education on executive functioning.

Looking back at the Montessori education program, this program proposed different aspects that could have possibly supported executive functioning. To start with her own work, Montessori states in her book "The Method" that the main point of her educational ideas was that children grow up to become independent adults in the world that are able to find solutions on their own and support themselves in life (Montessori, 1966). As we see in our examples of executive functioning in everyday life in the introduction, a lot of problem solving depends on our executive functions. Because the ability to solve complicated problems on our own make us independent beings, Montessori education seems to build on executive functioning. Some specific aspects of Montessori education that could have enhanced EFs are: the ability of children to work independently, and the skill to plan their own work-schedule (Wasmann-Peters, 2015). These aspects of Montessori require some or all of the executive functions to be carried out correctly. But as the results show, none of the executive functions were better in Montessori school children than children following the regular education. Maybe the executive function "training" that Montessori education implements through stimulating children to work independently and plan their own schedule is not generalizable to overall executive functioning. So the children may very well work better independently and have better planning skills, but they may only have them in their own classroom setting. This may explain why Bagby et al. (2012) only found a difference in the teacher BRIEF-questionnaire (classroom environment) and not in the parental BRIEF-questionnaire (non-classroom environment). Instead, the "Tools of the mind" program has proven to increase executive function skills in preschool children and even have generalizing effects to academic skills (Diamond et al., 2007; Diamond, 2011). It may be very interesting to see in future research what the difference is between Montessori education and the “Tools of the mind” program and compare both groups on executive functioning. “Tools of the mind” is based in preschool education and because Lillard (2012) did find better executive functioning in Montessori preschool children compared to children coming from a regular education background it may be an interesting comparison.

In conclusion, this study did not find any differences between Montessori and regular education on executive functioning and emotion regulation. We may thus conclude that overall, 7-8 year old children following a Montessori curriculum have the same level of executive functioning as children who follow a regular education curriculum in The Netherlands. Because our sample size was very small and selected from a few schools only, our results might not be fully representative for the complete population of schools, so we must be careful with this conclusion. Future research on differences in executive functioning between Montessori and regular education groups with a larger sample size and the inclusion of more schools is necessary to find out if the results from this pilot study can be confirmed and were not in fact influenced by the small sample size or other potential confounding effects which we will explain next.

This study has multiple limitations which must be taken very seriously when interpreting the results. To start, our sample size is very small and differs in size between the groups on the different tasks and questionnaires. Because of the matching of the groups on age and gender, multiple participants from the Montessori group were not taken into account. There were missing variables from the FEEL-KJ questionnaire because of difficulties with the filling in of the questionnaire. Because the parents were asked to voluntarily fill in the BRIEF questionnaire, not all parents responded so only 21 parents in total of the matched groups filled in the BRIEF. It is also very important to take into account that the two Montessori schools were situated in the same region, but the regular education school was situated in another region of The Netherlands (roughly 80 kilometers apart). Although we found no statistically significant differences in school performance scores and parental SES, unforeseen effects caused by the difference in region of the schools could have led to the current marginal and borderline significant results. As we said, we were not able to control for this factor, so future research should take the possible effect of region into account and control for this. In our analyses we also controlled for two possible confounders, school performance scores of the children and parental SES based on the highest academic achievement level of the parents.

We compared the groups on these factors and the groups did not significantly differ. The parental SES level difference was marginally significant, but with a very small sample size because only half of the parents returned the BRIEF we cannot truly conclude that there was no difference between the participants on parental SES.

Finally, it is very important to mention that this study is a pilot study set out to see if there are any differences between the effect of Montessori and regular education on school-aged children's executive functioning. We had limited time and resources and therefor only a small sample was selected with participants that were carefully matched on age and gender. This way we could eliminate possible effects of age and gender on executive functioning. Although the results of this pilot study were not significant, some marginally significant results may point to a possible effect of different types of education on executive functioning. Thus, future research with a much larger sample size is advised to see if the current marginal significant effects point to actual significant differences between the groups that are the effect of education type or if they are caused by other factors like the region in which the school is situated.

References

Bagby, J., Barnard-Brak, L., Sulak, T., Jones, N., & Walter, M. (2012). The effects of environment on children's executive function: A study of three private schools. *Journal of Research in Childhood Education*, *26*(4), 418-426. doi: 10.1080/02568543.2012.711431

Braet, C., Cracco, E., Theuwis, L., Grob, A., & Smolenski, C. (2013). *FEEL-KJ: vragenlijst over emotieregulatie bij kinderen en jongeren*. Amsterdam: Hogrefe.

Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Giedd, J. N., Castellanos, F.X., Haxby, J.V., Noll, D.C., Cohen, J.D., Forman, S. D., Dahl, R.E. & Rapoport, J.L. (1997). A developmental functional MRI study of prefrontal activation during performance of a go-no-go task. *Journal of cognitive neuroscience*, *9*(6), 835-847. doi: 10.1162/jocn.1997.9.6.835

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. (2001). Changes in executive control across the life span: examination of task-switching performance. *Developmental psychology*, *37*(5), 715. doi: 10.1037/0012-1649.37.5.715

Clark, C. A., Pritchard, V. E., & Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental psychology*, *46*(5), 1176. doi: 10.1037/a0019672

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science (New York, NY)*, *318*(5855), 1387. doi: 10.1126/science.1151148

Diamond, A. (2011). Biological and social influences on cognitive control processes dependent on prefrontal cortex. *Progress in brain research*, *189*, 319. doi: 10.1016/b978-0-444-53884-0.00032-4

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, *16*(1), 143-149. doi: 10.3758/bf03203267

Fuhs, M. W., Nesbitt, K. T., Farran, D. C., & Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental psychology*, *50*(6), 1698. doi: 10.1037/a0036633

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017-2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Jahanshahi, M., Saleem, T., Ho, A. K., Fuller, R., & Dirnberger, G. (2008). A preliminary investigation of the running digit span as a test of working memory. *Behavioural neurology*, *20*(1-2), 17-25. doi: 10.1155/2008/195684

Lee, K., Bull, R., & Ho, R. M. (2013). Developmental changes in executive functioning. *Child Development*, *84*(6), 1933-1953. doi: 10.1111/cdev.12096

Lillard, A. S., & Else-Quest, N. (2006). The early years: evaluating Montessori education. *Science*, 313, 1893–1894. doi: 10.1126/science.1132362

Lillard, A. S. (2012). Preschool children's development in classic Montessori, supplemented Montessori, and conventional programs. *Journal of School Psychology*, *50*(3), 379-401. doi: 10.1016/j.jsp.2012.01.001

Markus, C. R., & Jonkman, L. M. (2007). Attention switching after dietary brain 5-HT challenge in high impulsive subjects. *Journal of Psychopharmacology*, *21*(7), 700-708. doi: 10.1177/0269881107077354

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49-100. doi: 10.1006/cogp.1999.0734

Montessori, M. (1966). *De methode*. Amsterdam: Van Holkema & Warendorf.

Moore, D. S., McCabe, G. P., & Craig, B. A. (2012). *Introduction to the Practice of Statistics.* New York: W.H. Freeman and Company.

Moreno, A. J., Shwayder, I., & Friedman, I. D. (2016). The Function of Executive Function: Everyday Manifestations of Regulated Thinking in Preschool Settings. *Early Childhood Education Journal*, 1-11. doi: 10.1007/s10643-016-0777-y

Peng, H. H., Md-Yunus, S. (2014). Do children in Montessori schools perform better in the achievement test? A Taiwanese perspective. *International Journal of Early Childhood*, *46*(2), 299-311. doi: 10.1007/s13158-014-0108-7

Sanders, D., & Welk, D. S. (2005). Strategies to scaffold student learning: Applying Vygotsky's zone of proximal development. *Nurse Educator*, *30*(5), 203-207. doi: 10.1097/00006223-200509000-00007

Schleepen, T. M., & Jonkman, L. M. (2010). The development of non-spatial working memory capacity during childhood and adolescence and the role of interference control: An N-back task study. *Developmental neuropsychology*, *35*(1), 37-56. doi: 10.1080/87565640903325733

Schmeichel, B. J., & Tang, D. (2015). Individual differences in executive functioning and their relationship to emotional processes and responses. *Current Directions in Psychological Science*, *24*(2), 93-98. doi: 10.1177/0963721414555178

Smidts, D., & Huizinga, M. (2009). *Handleiding van de BRIEF Executieve Functies Vragenlijst.* Amsterdam: Hogrefe Uitgevers BV.

Wasmann-Peters, E. (2015) *Kiezen voor Montessori.*’s-Gravenhage: Nederlandse Montessori Vereniging.