The effect of gesture congruency on the cognitive learning process and learning outcomes.

Master Thesis 2018-2019

Supervisor: F. Kirschner

Second assessor: J. Janssen

Gitte van Helden (4244389)

Utrecht University

Gitte van Helden, Faculty of Social and Behavioural Sciences, Utrecht University.

Correspondence concerning this article should be addressed to Gitte van Helden, Faculty of Social and Behavioural Sciences, Utrecht University: g.vanhelden@students.uu.nl

THE EFFECT OF GESTURE CONGRUENCY ON THE COGNITIVE LEANING PROCESS AND LEARNING OUTCOMES

Abstract

It is known meaningful gestures made by the learner can lead to higher learning outcomes, there is still little known about the effect of gesturing on human cognitive processes. This study adds to this knowledge by comparing the effect of not gesturing, performing congruent (meaningful) gestures, and performing incongruent (meaningless) gestures on learning process and learning outcomes. Participants in this study were 94 primary school students between the age of 10 and 12 years old. Learning outcomes were measured with a post-task consisting of five recall questions and five comprehension questions. Cognitive load was measured by a dual task during both the learning-phase and the test-phase. A MANOVA showed no significant effect of gesture congruence on learning outcomes and on cognitive processes. These results are contradicting with previous research, but could be explained by a ceiling effect (i.e, the materials were not sufficiently complex), low internal consistency of the post-task and a low power. Future research should explore whether the results that were found are the consequence of unreliable and invalid measurement, or if there is indeed no effect of gesture congruency on the learning process and learning outcomes.

Keywords: gestures, congruency, cognitive load, learning outcomes, learning process.

The effect of gesture congruence on the cognitive learning process and learning outcomes.

In daily life, people gesture while speaking. It is known these gestures help listeners to process information (e.g. Beattie & Shovelton, 1999; Valenzo, Alibali, & Klatzky, 2002). There is also evidence that indicates gesturing serves the speaker as well. It was shown that gesturing makes it easier to recall information (Stevanoni & Salmon, 2005) and makes it possible to express implicit knowledge (Broaders, Wagner Cook, Mitchell, & Goldin-Meadow, 2007). If making gestures indeed helps people in everyday life to get information across and process information, gesturing could be a very useful tool in education for making learning more efficient and effective for learners. Gestures made by the learner could be used deliberately in classroom practice, for instance to foster processing of information when solving complex tasks.

Attention for the use of gestures made by the learner in education fits an emerging body of research arguing physical action enhances learning. This research revolves around the idea of embodied cognition, which states there is no strict distinction between body and mind when learning. According to this view all cognition is grounded in the bodies interaction with the world (Wilson, 2002). From the perspective of embodied cognition the use of physical action in the classroom, including gesturing, is promoted (Alibali & Nathan, 2012).

There is already a large body of research that shows a positive effect of gesturing on learning, mostly conducted within the context of primary schools (e.g. Broaders et al., 2007; Wagner Cook & Golding Meadow, 2006). However, there is not much cognitive empirical evidence about why gesturing leads to better learning. That is, there is not a lot known about the effect of gesturing on human cognitive processes. A brief start has been made by Goldin-Meadow, Nusbaum, Kelly, and Wagner (2001) and Wagner Cook, Yip, & Goldin-Meadow (2012) at investigating the relation between gestures and cognitive processes by incorporating measurements of cognitive load when empirically researching the effects of use of gestures. Nevertheless, in those studies cognitive load was only measured during the retrieval of information and not during the act of learning.

The aim of the current study is to provide more insight in the effect of gesturing on the learning process in combination with learning outcomes. It has already been shown that gestures can increase the learning outcomes of primary school students. This study will add to this existing knowledge by not only investigating learning outcomes, but also the perceived

cognitive load during both the learning phase and the test phase, in order to provide more insight in the effect of gesturing on the cognitive learning process.

Embodied cognition: the use of meaningful gestures

From the perspective of embodied cognition, it is assumed that our brain is not the only cognitive resource available for problem solving. Instead, cognition is deeply rooted within the body's interaction with the world and therefore our motions can replace the need for complex mental representations (Wilson, 2002; Wilson & Golonka, 2013). From an embodied perspective all learning, knowing and reasoning is inherently connected to sensorimotor activity. Every thought, even those derived from highly abstract reasoning, is grounded in physical embodiment (Abrahamson & Bakker, 2016; Glenberg, 2010).

Learning from an embodied perspective therefore typically involves physical movement, including gesturing (Alibali & Nathan, 2012; Hostetter & Alibali, 2008). There is empirical research that proves embodied learning tasks involving gestures made by the learner lead to better achievement compared to traditional learning tasks, within the context of primary school education. For example, Wagner Cook and Goldin-Meadow (2006) found that children could reproduce problem solving strategies for mathematical equivalence problems more easy when they were allowed to gesture during their explanation. Also, Broaders et al. (2007) found that children who were unable to perform mathematical equivalence problems benefitted more from instruction and performed better on a post-test compared to children who had not gestured.

However, making gestures does not necessarily lead to higher learning outcomes. From an embodied perspective the actions of a learner should be structurally aligned with the key features of a task (Alibali & Nathan, 2018). Only well aligned gestures are meaningful in relation to a task and foster learning. This claim is illustrated in a study of Golding-Meadow, Wagner Cook and Michell (2009) in which the effect of performing aligned gestures, partially aligned gestures and no gestures were compared on a task involving mathematical equivalence problems. Results indicated that children executing aligned gestures performed best on a post-test. The group that executed partially aligned gestures performed better than the group executing no gestures. This shows that the degree to which gestures are aligned with a task influences learning outcomes. Therefore, a distinction should be made between congruent gestures, which are well aligned with a task, and incongruent gestures, which are not aligned with a task

An explanation for the positive effect on learning outcomes when making congruent gestures during a task might be found in the effect gesturing has on human cognitive processes. There is evidence showing that performing congruent gestures lightens the cognitive load imposed on students when explaining mathematical equivalence problems, compared to performing incongruent gestures (Wagner Cook et al. 2012) and compared to not performing gestures (Goldin-Meadow et al., 2001).

Human Cognitive Architecture and Cognitive Load Theory

Human cognitive architecture (HCA) helps understanding cognitive processes (Sweller, 2004). An important facet of HCA is the interaction between unlimited long term memory (LTM) and limited working memory (WM) (Sweller, van Merriënboer, & Paas, 1998; van Merriënboer & Sweller, 2005). WM can temporarily hold about seven elements and actively process two or three elements. This amount is lower when elements are interacting with each other. LTM organizes and stores elements in cognitive schemata, in which multiple elements are 'chunked' into one element. WM can treat existing schemata as a single element, therefore prior knowledge can reduce the load imposed on WM.

Cognitive load theory (CLT) uses this knowledge of HCA as a fundament for instructional design principles, by making optimal use of the available cognitive resources (Sweller et al., 1998). Limitations of WM should be taken into consideration, because the load imposed on a learner should not exceed the capacity. Different types of cognitive load can be distinguished, which are additive. First, intrinsic load, which is embedded within the nature of material. Complex material, with many elements, is accompanied with a higher intrinsic load. The second type of load, extraneous load, is determined by features of instructional design that do not contribute to learning. Some distinguish a third type of load that is used for features in instructional design that foster cognitive processes, called germane load (e.g. Sweller et al., 1998; van Merriënboer & Sweller, 2005). There is criticism on the use of germane load because it is empirically insufficiently proven and conceptually indistinguishable from the other types of load (e.g. de Jong, 2010; Kayluga, 2011). Following the view of Kayluga (2011), it was decided to not use the concept of germane load in the current study. Instead, germane load was included as a part of intrinsic load.

There is no exact value that indicates which level of cognitive load is acceptable, because this depends on knowledge stored in LTM (Sweller et al., 1998). Existing schemata reduce the load imposed on WM. In sum, according to CLT, instruction is effective when it fosters creation of schemata while making optimal use of existing cognitive resources.

Therefore, it could be that performing congruent gestures makes instruction more effective by fostering these cognitive processes, which leads to higher learning outcomes.

Gesturing and cognitive load

Combining knowledge on the effect of gesturing on learning processes and learning outcomes, could add to the existing knowledge. It is already known congruent gestures can lead to higher learning outcomes (e.g. Broaders et al., 2007; Goldin-Meadow et al., 2009; Wagner Cook & Goldin-Meadow, 2006). Adding knowledge on the amount of cognitive load that is imposed on the learner, would provide more insight in the cognitive learning processes that accompany gesturing. Previous studies on the effect of gesturing on cognitive load, only incorporated measurement of cognitive load during the retrieval of information (Goldin-Meadow et al., 2001; Wagner Cook et al., 2012). Not much is known of the effect of gesturing on cognitive load during the learning phase.

Still, previous research provides insight in possible mechanism that explain the effect of gesturing on cognitive load and learning. It is suggested gestures can foster cognitive processes in two possible ways. First, from the perspective of embodied cognition it is suggested that meaningful physical action can replace the need for complex mental representations (Wilson, 2002; Wilson & Golonka, 2013). So, gesturing might create the possibility to process information through other systems than WM which would reduce the intrinsic load imposed on WM and therefore foster learning (Paas & Sweller, 2012; Wagner Cook et al., 2012). Second, gesturing possibly enriches the way information is represented, which fosters the processing of information and makes iFt easier to create schemata (Golding-Meadow et al., 2001; Paas & Sweller, 2012). The additional load imposed by the gestures can be classified as intrinsic because it fosters learning. However, it is required that gestures are aligned with the task in order to facilitate cognitive processes and thus reduce the cognitive load and foster learning (Goldin-Meadow et al, 2009; Wagner Cook et al., 2012). Incongruent gestures have no meaning in relation to a task and can be viewed as extraneous load, which can hamper learning when performing a complex task.

This study will add to existing knowledge, by comparing effects of congruent gestures, incongruent gestures, and no gestures on cognitive load during both the learning-phase and test-phase. The following research question will be answered: What is the effect of gesture congruency on the cognitive learning process and learning outcomes?

Three hypotheses are formulated. First, it is expected that performing congruent gestures leads to the lowest cognitive load during the learning-phase when compared to not

gesturing and performing incongruent gestures. Performing incongruent gestures is expected to lead to the highest cognitive load. This is expected because meaningful gestures foster information processing and the construction of schemata (Golding-Meadow et al., 2001; Wagner Cook et al., 2012), while meaningless gestures impose an extraneous load on the learner and therefore can hamper the construction of schemata (Wagner Cook et al., 2012).

Second, it is expected that performing congruent gestures leads to the lowest cognitive load during the test-phase when compared to not gesturing and performing incongruent gestures. Performing incongruent gestures is expected to lead to the highest cognitive load. During the test-phase students have to recall information from the schemata they constructed during the leaning-phase. Because it is expected that congruent gestures foster the creation of schemata, it will be easier to recall information (Golding-Meadow et al., 2001; Wagner Cook et al., 2012). Good schemata reduce the load imposed on a student (Sweller et al., 1998). Performing incongruent gestures is expected to hamper the creation of schemata, which will make it harder to recall information and therefore imposes a higher load on the student (Sweller et al., 1998).

Third, it is expected that performing congruent gestures leads to a higher performance on learning recall and learning comprehension (i.e., learning outcomes), compared to not gesturing and performing incongruent gestures. Performing incongruent gestures is expected to lead to the lowest score on learning recall and learning comprehension. Previous studies showed that performing gestures lead to higher learning outcomes (Broaders et al., 2007; Wagner Cook & Goldin-Meadow, 2006). However, gestures need to be aligned with the key features of a task to increase learning outcomes (Golding-Meadow et al., 2009). Incongruent gestures might lead to a higher extraneous load and therefore lead to lower learning outcomes when executing a complex task (Sweller, et al., 1998).

Method

Participants

The participants of this study were 94 primary school students (50 male, 44 female), with a mean age of 11.26 years (*SD*=0.45). Participants were found via three primary schools that were willing to let their classes with fifth grade students enrol in the research programme. Data were gathered in six different classes. All parents of the participants gave their active informed consent. Students with extreme reading difficulties and students who had already read the book ‘Ronja de Roversdochter’ (Lindgren, 1981) were excluded from this study.

Participants were distributed over three conditions: a control group (*n*=30), a congruent gesturing group (*n*=32) and an incongruent gesturing group (*n*=32). Per class students were randomly assigned to one of the conditions.

Instruments

Gestures. Three types of gesturing conditions were distinguished. The control group was instructed to keep their hands on the table while thinking about what they had read, to prevent them from making gestures. They were asked to sit up straight, while placing the palm of their hand in front them on the table. The incongruent gesturing group was instructed to move their hands in circles while thinking about what they had read, similar to the study of Wagner Cook et al. (2012). They were asked to sit up straight and hold their hand up in front of themselves while moving both hands in small circles. The congruent gesturing group was instructed to perform meaningful gestures that represented what they had read, while remaining seated. In order for the gestures to be meaningful for the individual, the students had to make up gestures themselves.

Learning task. A text was selected from 'Ronja de Roversdochter' (Lindgren, 1981) that was complex regarding the expected skill level of the participants. The story takes place in an imaginary world, so participants could not have prior knowledge regarding the book. A short introduction was included, to provide essential background knowledge for understanding the text. The text was divided in five story parts, under which the symbol of an orange circle was placed. The orange circle indicated students had to perform incongruent, congruent or no gestures, depending on the condition they were in. The text can be found in Appendix A. The learning-task was paper-based.

Post-task. The post-task was a test existing of ten questions, that measured to what extend students could recall information and to what extend students comprehended the storyline. Five questions were formulated for both recall and comprehension (Appendix B). Recall was tested with a statement for each story part (e.g., Ronja takes along a kettle to the cave). Students had to recall whether they were 'true' or 'not true'. To increase reliability and validity, statements could be directly found in the text. Comprehension was tested with multiple choice questions (e.g. What do Ronja and Birk find worst about the unwanted visit from the goblins? a. The goblins ate their bread, b. They are afraid the goblins will return., c. They are afraid the goblins will reveal to the bird witches where they live.). To increase reliability and validity, questions were based on three of the four skills for reading comprehension as described by CITO (2016): understanding (question 7, question 10),

interpreting (question 6; question 8) and extracting (question 9). The fourth skill, 'looking up', was excluded, because students were not allowed to view the text during the post-task. The post-task was paper-based.

Performance measurement. Recall was measured by five statements (question 1 – 5). Comprehension was measured by five multiple choice questions (question 6-10). Participants received a score between 0 and 5 (0= *no correct answer* and 5= *five correct answers*) for both recall and comprehension.

Cognitive-load measurement. A dual task was used to measure cognitive load. Remembering letters from paper was chosen as secondary task (Appendix C), because it meets the criteria of Brünken, Plass and Leutner (2003) for applying a dual-task approach. First, remembering the letters from paper is a simple task, that does not suppress the simultaneous leaning process. Second, remembering letters requires the same cognitive resources as the primary task, namely visual working memory. Third, remembering letters is proven to be a valid measurement of the cognitive load imposed by a task (Logan, 1979). Similar to Goldin Meadow et al., (2001) the number of letters presented to participants was 6 and participants were allowed to look at the letters for 5 seconds. The participants had to write down the letters on a blanc square on a paper, that was included in the post-task. The cognitive load was measured during the learning task and during the post-task, by assigning a score between 0 and 6 (0= *highest cognitive load* and 6= *lowest cognitive load*). The score was computed by assigning one point for every correct letter a student wrote down. When a student wrote down an incorrect letter, a half point was subtracted from the total score until the minimal score of 0 was reached.

Practice material. Practice material was designed to make participants acquainted with performing no gestures, congruent gestures and incongruent gestures. For this purpose five story parts similar to the learning task were selected from the book 'De Koning van Katoren' (Terlouw, 1971) (Appendix D).

Validity and Reliability. Validity and reliability were assured in various ways. First the comprehension questions were based on questions of the CITO test for learning comprehension, which is considered a reliable and valid test (Tomesen, Engelen, & Hiddink 2018). Recall questions could be directly traced back to a feature of every story part and were therefore a valid indication of the facts students could recall from the story parts.

In order to increase the reliability, the researcher followed a detailed protocol during explanation of the experiment and practice sessions in all three conditions (Appendix E). The

learning task, post-task and protocol were reviewed by five primary school teachers, who combined their job with an Educational Sciences master's programme at Utrecht University, during a feedback session. The learning task and post-task were judged on appropriate complexity. In response to this feedback session, some small changes were made in the post-task. For example the answer a. for question 6 was changed from 'amusing story' in to 'fantasy story', because students were possibly not familiar with the term 'amusing'. The protocol was judged on practical attainability. One change that was made in response to the feedback session was the addition of the preparation preliminary to the practice sessions.

Furthermore, a pilot of the experiment was executed for the congruent and incongruent condition. Both trials included two students that were representative for the sample. The pilot gave no rise to adaptions in the learning task, post-task and protocol. During the pilot it was decided to make use of an observation form (Appendix F), to verify whether all students showed the desired behaviour. Students who did not perform the appropriate behaviour, thus performing congruent, incongruent or no gestures depending on the condition they were in, were excluded from the study.

Despite these efforts the internal consistency of the post-task can be considered low, with a Cronbach's alpha of .08 for recall and a Cronbach's alpha of .20 for comprehension (Tavakol & Dennick, 2011). This is possibly due to the fact the post-task existed of a limited amount of multiple-choice items. When the number of test items is too small, the assumption of tau-equivalence is violated, which leads to an underestimation of the reliability (Tavakol & Dennick, 2011).

Procedure

Per class students were randomly divided in groups and equally distributed over three conditions. For all six participating classes three separate sessions were planned: one for every condition. A session consisted of three elements: (1) a practice for reading the text (2) reading-phase including measurement of cognitive load (3) the test-phase including measurement of cognitive load. Every session was led by the same researcher. Every session took place in a quiet room within the school. Every participant had an individual table and enough space to execute gestures without distracting other participants.

Practicing session. To make sure the students were acquainted with the procedure of their condition a practicing session was arranged. The practicing text (Appendix D) was used in all three conditions. The required behaviour was modelled by the researcher, by using the think aloud method (e.g. Silvén & Vauras, 1992; van Someren, Barnard,& Sandberg, 1994).

The researcher followed a protocol (Appendix E). Participants were allowed to ask questions during the entire practicing session.

Reading-phase. In all three conditions the reading task and the post-task were distributed. Participants were asked to fill in their name, school, class, gender and day of birth on the front page of the post task and to put the post task aside afterwards. Then they were asked to open the reading task and read the introduction. They had the opportunity to ask questions when they encountered unclarities during the introduction. After the introduction it was no longer allowed to ask questions. Then the researcher presented a card with 6 letters for 5 seconds at the front of the class. The letters were big enough for all participants to see. Participants were asked to remember these letters. Subsequently, participants were asked to carefully read the text until they encounter the first orange circle. When all students read the first story part, the researcher invited them to process the text with respectively no gestures, congruent gestures or incongruent gestures. The researcher registered on the observation form (Appendix F) whether students showed the appropriate behaviour. After 30 seconds, students were asked to move to the next story part. This procedure was repeated for the next four story parts. After completing the reading exercise, the students were asked to note the letters they remembered the blanc squares that were included in the post-task and to hand in the reading task. Again, they were asked to put the post-task aside.

Test-phase. The test phase procedure was similar for all three conditions. First the reading tasks were collected and returned to the researcher. Before starting the test, the researcher presented a card with 6 letters for 5 seconds at the front of the class. The letters were big enough for all participants to see. Students were asked to remember these letters. Then students were asked to individually complete the test questions in the post-task as accurate as possible. The maximal time available for completing the post test was 12 minutes. When all students completed the test, they were asked to write down the letters they remembered on the second blanc square on the last page of the post-task. The post-task was collected by the researcher.

Analysis

A multivariate analysis of variance (MANOVA) was used to examine the effect of gesture congruence on learning effectivity of primary school students (N=94). Before conducting the MANOVA, assumptions regarding normality, multicullinearity, linearity and homogeneity were checked. The Sharpiro-Wilk tests showed the assumption of univariate normality was violated. However, MANOVA is fairly robust against violations of the

normality assumption when groups exists of 30 or more participants (Allen & Bennet, 2012). Therefore the decision was made to continue the analyses. Multicollinearity was not of concern, because the correlations between the dependent variables were not excessive. Linearity could be assumed, because the relationships between dependent variables were roughly linear. Finally, homogeneity of variance-covariance matrices could be assumed, since Box's *M* was non-significant at $\alpha = .001$.

The dependent variables were also analysed individually at a Bonferroni adjusted alpha level of .13. Since previous analyses did not show necessity for analysing the effect of the three conditions, no gestures, congruent gestures and incongruent gestures, on the dependent variables separately, it was decided to not execute a Bonferroni post-hoc test.

Results

Analyses were focussed on the effect of gesture congruency on the learning process and learning outcomes. No registration problems occurred, so there was no need for excluding participants from the analyses.

Mean scores with standard deviation were computed for all dependent variables within the three gesturing conditions (Table 1). It was shown the means of all dependent variables were high. For this reason boxplots for al dependent variables per condition were computed (Figure 1). This showed that the scores are not distributed equally over the range of possible scores. Most students earned high scores on all dependent variables within all three conditions, causing a negatively skewed distribution.

Table 1.
Means and Standard Deviations of the Dependent Variables as a Function of Gesture Congruency.

	No Gestures (n = 30)		Congruent Gestures (n = 32)		Incongruent Gestures (n = 32)	
	M (*SD*)	95% CI	*M* (*SD*)	95% CI	*M* (*SD*)	95% CI
Letters learning-phase	4.67 (1.60)	[4.07, 5.26]	4.13 (1.63)	[3.53, 4.71]	3.69 (1.66)	[3.09, 4.29]
Letters test-phase	4.25 (1.61)	[3.65, 4.85]	3.22 (1.16)	[2.63, 3.80]	3.48 (1.68)	[2.88, 4.09]
Recall	3.90 (0.76)	[3.61, 4.18]	3.90 (0.76)	[3.31, 4.07]	3.81 (0.82)	[3.52, 4.11]
Comprehension	3.40 (1.07)	[3.00, 3.80]	3.60 (1.13)	[3.19, 4.00]	3.81 (0.90)	[3.49, 4.13]

Note. CI = confidence interval.

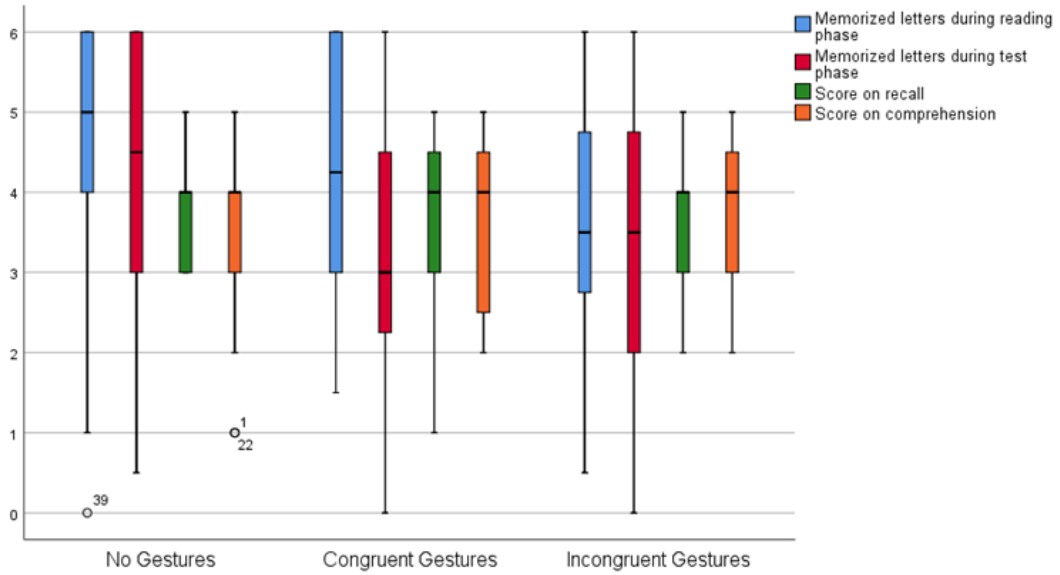

Figure 1. Boxplots showing the negatively skewed distribution of the dependent variables as a function of gesture congruency.

A MANOVA was used to examine the effect of gesture congruence on learning effectivity of primary school students (*N*=94). Findings showed that there was no significant effect of gesture congruency on the combined dependent variables, F (8, 178) = 1.79, p = .081.

Also, the four dependent variables were analysed separately. During the learning-phase perceived cognitive load was measured. Findings showed that there was no significant effect of gesture congruency on the experienced cognitive load during the learning-phase, F (2, 91) = 2.80, p = .066. During the test-phase perceived cognitive load and learning outcomes for recall and comprehension were measured. Findings showed no significant effect of gesture congruency on perceived cognitive load during the test-phase, F (2, 91) = 3.30, p = .041. Also, there was no significant effect of gesture congruency on both recall, F (2, 91) = 0.45, p = .642, and comprehension, F (2, 91) = 1.23, p = .297.

Conclusion

Previous research showed that performing gestures led to better learning outcomes, when gestures were aligned with a task (e.g. Broaders et al., 2007; Goldin-Meadow et al., 2009; Wagner Cook & Goldin-Meadow, 2006). Although not much was known about the effect of gesturing on human cognitive processes, there was research indicating that making meaningful gestures led to a lower cognitive load during the test (Golding-Meadow et al.,

2001; Wagner Cook et al., 2012). The current study contributed to this knowledge by investigating the effect of gesture congruency on the learning process and on learning outcomes.

First, there was no effect of gesture congruency on the learning process during the learning-phase. These findings were not in line with the hypothesis that performing congruent gestures would lead to a lower cognitive load compared to not gesturing and performing incongruent gestures during the leaning-phase, with incongruent gestures leading to the highest cognitive load (H1). This indicates that meaningful gestures do not foster information processing, as was expected based on literature suggesting meaningful gestures improve the creation of schemata (Golding-Meadow et al., 2001; Wagner Cook et al., 2012). Also, it shows that performing incongruent gestures do not hamper information procession due to the extraneous load they impose on students.

Second, there was no effect of gesture congruency on the learning process during the test-phase. These findings were not in line with the hypothesis that performing congruent gestures would lead to a lower cognitive load compared to not gesturing and performing incongruent gestures during the test-phase, with incongruent gestures leading to the highest load (H2). It is also contradicting towards previous studies that had shown students perceived a lower cognitive load when performing congruent gestures (Golding-Meadow et al., 2001; Wagner Cook et al., 2012). The current findings suggest that performing congruent gestures during a reading task does not lead to a lower cognitive load when retrieving information during the test-phase. Also, performing incongruent gestures does not lead to a higher cognitive load when retrieving information during the test-phase, which indicates incongruent gestures do not impose a high extraneous load on a student.

Third, there was no effect of gesture congruency on learning outcomes for both recall and comprehension. This was not in line with the expectation that performing congruent gestures would lead to higher learning outcomes during the test-phase, compared to not gesturing and performing incongruent gestures, with incongruent gestures leading to the lowest learning outcomes (H3). These findings also contradict with previous findings indicating that performing congruent gestures leads to higher learning outcomes compared to not gesturing (e.g. Broaders et al., 2007; Wagner Cook & Golding Meadow, 2006), while performing incongruent gestures accounts for lower learning outcomes (Wagner-Cook et al., 2012). The current findings suggest that performing congruent, incongruent or no gesturing do not lead to significant differences in learning outcomes.

In sum, the current study indicates there was no effect of gesture congruence on the cognitive learning process and learning outcomes. These findings do not align with findings from previous studies, which showed that performing congruent gestures led to higher learning outcomes (e.g. Broaders et al., 2007; Wagner Cook & Golding Meadow, 2006) and a lower perceived cognitive load (Golding-Meadow et al., 2001; Wagner-Cook et al., 2012) when compared to not gesturing and performing incongruent gestures. Also, these findings do not fit with the claim of embodied cognition that all learning is grounded in the physical interaction from the body with the environment (Wilson, 2002; Wilson & Golonka, 2013). Therefore, learning should benefit from meaningful movement, such as congruent gestures (Alibali & Nathan, 2012; Hostetter & Alibali, 2008).

A possible explanation for these unexpected results can be found in the nature of the task that was used. It was found that the mean scores on memorized letters during the test- and learning-phase and learning outcomes for recall and comprehension were all high. The negatively skewed distribution, with most of the students earning high scores and only a few low scores within all three conditions, suggests there is a ceiling effect (Austin & Brunner, 2003). This means a substantial number of the participant achieved a high score, because the scale only discriminates among participations in the low to moderate range. The reading task and test might have been too easy for the students. This explains why no effects on learning process and learning outcomes were found. When executing a simple task a low cognitive load is imposed on a learner, because few elements need to be processed at one time. Extraneous load, for example making incongruent gestures, can therefore be handled without influencing the learning process or learning outcomes, because there is space left in working memory. Also, if performing congruent gestures would alleviate the load that is imposed on a learner, this would not have impact during the execution of a simple task. The learner was already able to process the information. A complex task imposes a high load on the learner, hence it is important to make optimal use of available cognitive resources. In this case, features of instruction that help students process information or put an extra burden on the student, would influence the learning process and learning outcomes.

A limitation of this study therefore is that reliability and validly were possibly hampered. Validity was impaired, because the materials possibly did not have a sufficient level of complexity. Reliability was hampered, because the post-task lacked internal consistency. Also, the number of participants that enrolled in this study was relatively small. A larger sample would have enlarged the reliability and validity. Future research should

address these issues regarding reliability and validity in various ways. In first place it is important to develop valid and reliable materials that can be used in similar studies. Although the learning task and test were based on skills formulated by CITO (2016) and were judged as sufficiently complex by five primary school teachers who combined their job with a master Educational Sciences, a ceiling effect emerged. This can be prevented in the future by making the material more complex, for example by choosing tasks in which students should relate a high amount of elements to each other, because this imposes a higher load on the learner (Sweller et al., 1998). Second, reliability can be assured in future research by developing a post-task with more questions, which will probably increase the internal consistency. Third, a larger sample should be selected, in order to provide a better representation of the population.

A second limitation is that the results of this study do not provide insights in the long-term effects of the intervention. Students were only tested directly after reading the test, so there is no evidence for potential positive or negative long-term effects on learning outcomes and experienced cognitive load. In future research this limitation could be prevented by adding follow-up test that will take place after a longer period of time.

Lastly, it is a limitation that the congruent and incongruent gesture condition differed from the type of instruction students were used to in their normal classroom environment. A practice session was arranged to let students get familiar with the different conditions. However, it could be that performing gestures was still too new distractive for students. Therefore, future research should be covering a larger period of time, including multiple sessions. Students then would have time to get acquainted with this new kind of instruction.

If the findings in this study are not the consequence of unreliable and invalid measurement, they would have implications for educational practices. It was shown that performing gestures, regardless of their congruence with the learning task, did not significantly influence scores of students on both recall and comprehension. Based on these findings gestures should not be used to increase learning outcomes. However, performing congruent or incongruent gestures did not enlarge the cognitive load that was imposed on students. Therefore performing gestures also does not hamper the cognitive learning process, by causing a cognitive overload. Gestures therefore can be used in educational practice for other purposes than increasing learning outcomes. For instance, gestures can be used for motivational purposes.

Despite of its limitations, this study addresses an important gap in existing literature. It is the first study to connect perceived cognitive load during both the learning-phase and test-

phase to gesture congruency. Future research should explore whether the results that were found are the consequence of unreliable and invalid measurement, or if there is indeed no effect of gesture congruency on the learning process and learning outcomes.

References

Abrahamson, D. & Bakker, A. (2016). Making sense of movement in embodied design for mathematics learning. *Cognitive Research: Principles and Implications, 1*(1),1-33. doi:10.1186/s41235-016-0034-3

Alibali, M. W., & Nathan, M. J. (2018). Embodied cognition in learning and teaching: action, observation and imagination. Fischer, F., Hmele-Silver, C. E., Goldman, S. R., Reumann, P. (Eds.). *International Handbook of the Learning Sciences* (pp.75-85). New York, NY: Taylor & Francis.

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from leaners' and teachers gestures. *Journal of the learning sciences, 21*(2), 247-286. doi:10.1080/10508406.2011.611446

Allen, P. & Bennett, K. (2012). *SPSS a practical guide. Version 20.* Melbourne: Cengage Learning.

Austin, P.C. & Brunner, L. J. (2003). Type I error inflation in the presence of a ceiling effect. *The American Statistician, 57*(2), 97-104. doi:10.1198/0003130031450

Beattie, G. W., & Shovelton, H. K. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica, 123*, 1-30. doi:10.1515/semi.1999.123.1-2.1

Broaders, S. C., Wagner Cook, S., Michell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology, 136*(4), 539-550. doi:10.1037/0096-3445.136.4.539

Brünken, R., & Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53-61. doi:10.1207/S15326985EP3801_7

CITO (2016). *Begrijpend lezen 3.0 Leerkrachtmap gr.7.* Arnhem: CITO

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science,* 38, 105-134. doi:10.1007/s11251-009-9110-0

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley*

Interdisciplinary Reviews: Cognitive Science, 1, 586– 596. doi:10.1002/wcs.55

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: gesturing lightens the load. *Psychological Science, 12*(6), 516-522. doi:10.1111/1467-9280.00395

Goldin-Meadow, S., Wagner Cook, S., & Michell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science, 20*(3), 267-272. doi:10.1111/j.1467-9280.2009.02297.x

Hostetter, A. B. & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review, 15*(3), 495-514. doi:10.3758/PBR.15.3.495

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review, 23,* 1-19. doi:10.1007/s10648-010-9150-7

Lindgren, A. (1981). *Ronja de Roversdochter*. Amsterdam, The Netherlands: Uitgeverij Ploegsma.

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance, 5*(2), 189-207. doi:0096-1523/79/0502-0189$00.75

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review, 24,* 27-45. doi:10.1007/s10648-011-9179-2

Silven, M., & Vauras, M. (1992). Improving reading through thinking aloud. *Learning and Instruction, 2,* 69-98. doi:10.1016/0959-4752(92)90024-G

Stevanoni E., & Salmon, K. (2005). Giving memory a hand: Instruction children to gesture enhances their event recall. *Journal of Nonverbal Behaviour, 29*(4), 217-233. doi:10.1007%2Fs10919-005-7721-y

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science, 32*(2), 9-31. doi:10.1023/B:TRUC.0000021808.72598.4d

Sweller, J., van Merriënboer, J. J. G., Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296. doi:1040-726X/98/0900-0251$15.00/0

Tavakol, M. & Dennick, F. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2,* 53-55. doi:10.5116/ijme.4dfb.8dfd

Terlouw, J. (1971). *De Koning van Katoren*. Rotterdam, The Netherlands: Uitgeverij

Lemniscaat

Tomessen, M., Engelen, R., & Hiddink, L. (2018). *Wetenschappelijke verantwoording Begrijpend lezen 3.0 voor groep 7*. Arhnem: CITO.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28*(2), 187-204. doi:10.1016/S0361-476X(02)00007-3

Wagner Cook, S., & Golding-Meadow, S. (2006). The role of gesture in learning: do children use their hands to change their minds? *Journal of Cognition and Development, 7*(2), 211-232. doi:10.1207/s15327647jcd0702_4

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: recent developments and future directions. *Educational Psychology Review, 17*(2), 147-177. doi:10.1007/s10648-005-3951-0

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The Think Aloud Method.* London: Academic Press.

Wilson, M. (2002). Six Views of Embodied Cognition. *Psychonomic Bulletin & Review, 9(4),*625-636. doi:10.3758/bf03196322

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology, 4*(58), 1-13. doi:10.3389/fpsyg.2013.00058

Appendix A. Learning Task

Ronja de Roversdochter.

Ronja is een roversdochter. Haar vader is de leider van een roversbende. De burcht waar ze wonen staat in een bos vol met duistere wezens, zoals aardmannen en vogelheksen. Op een dag ontmoet ze Birk, een jongen uit een vijandige roversstam. In het geheim sluiten Ronja en Birk vriendschap. Op een dag besluiten ze samen weg te lopen van huis en in een grot in het bos te gaan wonen.

Het duurde wel even voordat ze weer terug waren bij de grot: Ronja met haar emmer, Birk met zijn boog en wat andere dingen uit zijn verstopplaats. Hij legde alles op een rij voor de ingang van de grot neer om Ronja te laten zien wat hij bezat. Een bijl, een slijpsteen, een keteltje,

visgerei, vogelstrikken, pijlen voor bij zijn boog, een korte speer; allemaal dingen die onmisbaar waren om in het bos te kunnen leven.

‘Ik zie dat jij weet wat wij bosbewoners allemaal moeten kunnen,’ zegt Ronja. ‘We moeten zelf ons voedsel verzamelen en ons verdedigen tegen de vogelheksen en roofdieren.’

Natuurlijk,’ zegt Birk. ‘We zullen…’

Verder kwam hij niet, want Ronja pakte hem bij zijn arm en fluisterde verschrikt: ‘Stil! Er zit iemand in de grot.’

Met ingehouden adem zaten ze te luistern. Het was waar, er was iemand naar binnen geslopen terwijl zij er niet waren. Birk pakte zijn speer en bleef stil naast Ronja staan wachten. Ze hoorden geluiden binnen de grot en het was afschuwelijk niet te weten wie die geluiden maakte. Misschien zat de grot wel vol met vogelheksen die op de loer lagen en ieder ogenblik met hun scherpe klauwen naar buiten konden komen.

Ten slotte kregen ze genoeg van al dat luisteren en wachten. ‘Kom toch naar buiten, vogelheksen!’ riep Birk. ‘Dan krijgen jullie de scherpste speer uit het hele bos te zien!’

Maar er kwam niemand naar buiten. Er kwam alleen een kwaad gesis van binnen.

‘Mens in het bos van de aardmannen! Kom hier, alle aardmannen, en ga ze te lijf!’

Op dat ogenblik barstte Ronja bijna van woede.

‘Schiet op, aardmannen!’, schreeuwde ze. ‘Loop naar de bliksem en gauw ook! Anders trek ik jullie haar uit jullie koppen!’

Onder luid geblaas en gesis kwam er een menigte aardmannen uit de grot tevoorschijn. Ronja siste net zo hard terug en Birk liet hun zijn speer zien. Toen verdwenen ze in razende vaart de berg af. Op handen en voeten lieten ze zich langs de steile helling omlaag naar de rivier glijden. Maar velen van hen gleden uit en vielen piepend van woede in de bruisende waterval. Na een tijdje dreven hele trossen aardmannen stroomafwaarts. Maar ze slaagden er ten slotte in weer op het droge te klimmen.

'Ze kunnen goed zwemmen, die monstertjes,' zei Ronja.

'Brood eten kunnen ze ook', zei Birk toen ze de grot in kropen en zagen dat de aardmannen een heel brood uit hun voorraad naar binnen hadden gewerkt. Verder hadden ze niets aangericht, maar het was al erg genoeg dat ze er geweest waren.

'Eigenlijk is dit heel vervelend,' zei Ronja. 'Want nu zullen hun kletspraatjes zich als een lopend vuurtje door het bos verspreiden en binnenkort weten alle vogelheksen waar wij zitten.'

Appendix B. Post-Task

Nummer_____ Conditie_____

Naam: _____

Klas:_____

School:_____

Ik ben een: 0 Jongen

0 Meisje

Ik ben geboren op:

_____ --_____--_____

(dag) (maand) (jaar)

Schrijf hier alle letters die je onthouden hebt op:

Vragen over de tekst:

Zijn de onderstaande zinnen waar of niet waar? Omcirkel het goede antwoord.

1. Ronja bracht een keteltje mee naar de grot.
   Waar / Niet waar

2. Ronja merkt dat er iemand in de grot zit, doordat zij geluid hoort.
   Waar / Niet waar
3. Birk wil de aardmannen verjagen met een bijl.
   Waar / Niet waar
4. De aardmannen springen expres in het water om te ontkomen aan Ronja en Birk.
   Waar / Niet waar
5. De aardmannen hebben spullen die in de grot lagen kapot gemaakt.
   Waar/ Niet waar

Omcirkel het juiste antwoord. Let op: er is maar één antwoord goed.

6. Wat voor soort tekst is dit?
   a. Een fantasieverhaal
   b. Een betoog
   c. Een informatieve tekst
   d. Een instructieve tekst

7. Welk antwoord beschrijft de emoties van Ronja en Birk in de juiste volgorde.
   a. Eerst zijn ze bang omdat iemand in de grot zit, daarna zijn ze blij omdat er geen vogelheks in de grot zit.
   b. Eerst zijn ze bang omdat iemand in de grot zit, daarna zijn ze boos omdat aardmannen hun grot zijn binnengedrongen.
   c. Eerst zijn ze boos omdat er iemand in de grot zit, daarna zijn ze bang omdat het aardmannen blijken te zijn.

8. Wat vinden Ronja en Birk het ergste aan het ongewenste bezoek van de aardmannen?
   a. De aardmannen hebben al hun brood opgegeten.
   b. Ze zijn bang dat de aardmannen een volgende keer terug zullen komen.

c. Ze zijn bang dat de aardmannen aan de vogelheksen zullen vertellen dat er mensen in de grot wonen.

9. Wat is de kern van het verhaal?
    a. Ronja en Birk hebben spullen gehaald om voedsel te kunnen vinden en zichzelf te kunnen beschermen.
    b. Ronja en Birk vonden aardmannen in hun grot, deze hebben ze verjaagd.
    c. Ronja en Birk waren bang dat er vogelheksen in de grot zaten.

10. Wat weet je niet na het lezen van dit verhaal?
    a. Waar Ronja en Birk wonen.
    b. Wat de beste manier is om een vogelheks te verslaan.
    c. Welke spullen Ronja en Birk nodig hebben om te overleven in het bos.

Schrijf alle letters die je onthouden hebt hier op:

Appendix C. Measurement of cognitive load

S L C P X T
M D P Z W Y
L Q X T B L
Z P F N R K

Appendix D. Practicing material

Tekst: De Koning van Katoren.

Wat vooraf ging: Stach wil koning worden. Hiervoor moet hij eerst zeven onmogelijke opdrachten voltooien. Voor zijn eerste opdracht moet hij naar het plaatsje Decibel, dat jaren geplaagd wordt door vogels die zo hard

krijsen dat iedereen die het hoort meteen doof wordt. Iedereen in Decibel moet daarom verplicht oorkleppen dragen. Stach verzint een plan om de vogels te laten stoppen met krijsen. Hij leent een geluidsinstallatie, versterkers en een toongenerator bij meneer Holder en installeert deze in het burgemeestershuis.

Stach zit naast de toongenerator en wacht. Natuurlijk zul je net zien dat de vogels nu niet komen. De hele dag niet en ’s nachts niet en de volgende dag ook niet. Voor de derde dag is de weersverwachting regenbuien met opklaringen. Meneer Holder komt langs en zegt dat hij zijn dure spullen niet aan zulk weer wil blootstellen. Hij heeft het woord blootstellen nog niet uitgesproken of het lieve leven begint. De vogels vliegen met duizenden over de stad en schreeuwen hun schreeuw.

Resoluut haalt Stach de schakelaar om. Even duurt het voordat de versterkers warm zijn. Dan schalt er een geluid uit de luidsprekers, even hard als dat van de vogels en ongeveer even hoog. Stach draait aan de grote schijf van de toongenerator. Nu is zijn toon hoger dan die van de vogels. Even zwijgen de dieren, verbaasd lijkt het. Dan beginnen ze weer, en o wonder, nu hoger dan eerst. Weer draait Stach de knop van de toongenerator een slag. Zijn ogen schitteren. De vogels proberen boven zijn toon uit te komen. Nu moet de rest ook lukken. Steeds hoger worden de klanken uit de luidsprekers, steeds hoger gaan ook de vogels.

‘Nu dan,’ zegt Stach.

Een fikse draai aan de schijf. Een vreselijk hoog piepend geluid. De vogels proberen er bovenuit te komen. En dan… pang… pang…pang, pang. Met venijnige tikjes knappen de stembanden van de kwelgeesten van Decibel. Na een minuut kan geen vogel meer een geluid uitbrengen. Stach schakelt de stroom uit. Een onwaarschijnlijke stilte valt over de stad.

De mensen in hun huizen hebben gehoord dat het geluid steeds hoger is geworden, ze hebben het gepang gehoord, en nu horen ze niets meer, terwijl de vogels nog steeds rondvliegen. De meesten denken dat ze zelf doof zijn worden, want de vogels zien en niets horen, dat kan niet.

Zij die aan het plein wonen zien hoe die vreemde jongen, die niet van hier is, het burgemeestershuis uitkomt, naar boven kijkt en dan zijn oorkleppen afneemt. Daarna komt de burgemeester het gemeentehuis uitstormen, ze zien hoe hij zijn oorkleppen op de grond smijt en vertrapt en hoe hij die jongen omarmt en een rondedans met hem maakt. Dan holt iedereen het plein op en alle oorkleppen worden op een grote hoop gegooid en de mensen praten met elkaar, ze verstaan elkaar zonder moeite. Ze juichen en dansen en ze proberen te zingen, maar ze kunnen niet zingen – wie heeft er ooit gezongen in Decibel?

Bron: De Koning van Katoren, Jan Terlouw.

Appendix E. Protocol instruction

Protocol Practicing session condition 1: No gestures (control group)

Voorbereiding: papieren met de oefentekst uit het boek De Koning van Katoren (Terlouw, 1971) wordt op de tafels van de leerlingen neergelegd.

De onderzoeker vraagt de studenten plaats te nemen achter een van de tafels en de tekst onaangeroerd te laten liggen. Wanneer alle studenten plaats hebben genomen kan de instructie worden gestart.

- De onderzoeker begint met een korte uitleg over het onderzoek. *'Hallo mijn naam is ... en ik voer een onderzoek uit naar begrijpend lezen bij basisschool leerlingen. Jullie klas is in verschillende groepen verdeeld en iedere groep gaat dezelfde tekst op een andere manier lezen. Daarna gaan alle groepen dezelfde toets maken en ga ik bekijken welke manier van lezen het beste werkt. We gaan eerst een oefening doen waarbij ik jullie de manier leer waarop jullie groep de tekst gaat lezen. Voordat we daarmee beginnen mogen jullie de introductie van de tekst lezen. Als jullie vragen hebben over het verhaal mogen jullie deze stellen.'*
- De kinderen lezen de introductie en stellen eventuele vragen. De onderzoeker beantwoordt eerst verhalen over de tekst. Daarna legt hij/zij uit hoe de leesopdracht zal verlopen. *'We gaan dadelijk de tekst lezen totdat we een oranje stip tegen komen.*

Lees zo zorgvuldig mogelijk. Zodra je bij de oranje stip bent aangekomen, ga je met je armen over elkaar zitten zodat ik kan zien wie klaar is met lezen. Wanneer iedereen van uit de groep klaar is met lezen, gaan we een denkoefening doen. Deze zal ik zo uitleggen. Jullie mogen beginnen met lezen.'

- Wanneer alle leerlingen klaar zijn legt de onderzoeker de denkoefening uit. *'We leggen allemaal onze handen op tafel en blijven stil zitten. We gaan nu 30 seconden nadenken over het verhaal dat we gelezen hebben. In die 30 seconden proberen we het stukje dat we gelezen hebben in ons hoofd aan onszelf te vertellen. Ik zal het voordoen: Stach heeft een toongenerator in het burgemeestershuis neer gezet en hij wacht tot de vogels komen, zodat hij zijn plan kan uitvoeren. De vogels komen alleen maar niet. Net wanneer het zijn plan dreigt te mislukken, omdat meneer Holder zijn apparaten terug wil komen de vogels gelukkig toch.* De wekker gaat na 30 seconden.
- De onderzoeker vraagt de leerlingen het tweede stuk tekst te lezen tot aan de oranje cirkel. Wanneer alle kinderen klaar zijn vraagt de onderzoeker aan de leerlingen welke dingen uit het verhaal zij zouden noemen in hun samenvatting. *'Als jullie het stuk dat jullie gelezen hebben moeten samenvatten in je hoofd, waar begin je dan mee'* Een kind krijgt de beurt. De onderzoeker vult het antwoord waar nodig aan. *'Hoe zou de samenvatting verder gaan?'* Een ander kind krijgt de beurt. De onderzoeker vult waar nodig aan. De onderzoeker geeft kinderen de beurt totdat de samenvatting volledig is. Ten slotte somt de onderzoeker de volledige samenvatting nog op: '*Stach haalt de schakelaar om, waardoor er een geluid uit de speakers komt dat lijkt op het geluid van de vogels. Daarna draait hij de toon omhoog, de vogels doen hem na. Hij blijft de toon omhoog draaien, de vogels blijven hem nadoen. Hij gelooft dat zijn plan gaat werken.'*
- De leerlingen mogen opnieuw lezen tot de oranje cirkel. Wanneer alle leerlingen klaar zijn vertelt de onderzoeker dat zij nu zelf in hun hoofd mogen samenvatten. *'Ik denk dat jullie klaar zijn om de opdracht helemaal zelf uit te voeren. Ik zet een wekker op 30 seconden. Jullie mogen in die tijd allemaal in proberen het verhaal samen te vatten in je hoofd, op de manier waarop we geoefend hebben. Het is belangrijk dat jullie je handen op tafel houden en geen bewegingen maken tijdens het nadenken.'*
- De onderzoeker vraagt een aantal leerlingen hun samenvatting te delen met de rest. De groep mag indien nodig aanvullen. Aan de hand hiervan kan hij beoordelen of de opdracht goed is begrepen door de groep. Indien de opdracht nog niet goed begrepen

is door de groep, kan de onderzoeker er voor kiezen ook de laatste twee delen uit de oefentekst te behandelen. Indien de opdracht begrepen is kan de onderzoeker beginnen met het experiment.

Protocol practicing session condition 2: congruent gestures

Voorbereiding: papieren met de oefentekst uit het boek De Koning van Katoren (Terlouw, 1971) wordt op de tafels van de leerlingen neergelegd.

De onderzoeker vraagt de studenten plaats te nemen achter een van de tafels en de tekst onaangeroerd te laten liggen. Wanneer alle studenten plaats hebben genomen kan de instructie worden gestart.

- De onderzoeker begint met een korte uitleg over het onderzoek. *'Hallo mijn naam is ... en ik voer een onderzoek uit naar begrijpend lezen bij basisschool leerlingen. Jullie klas is in verschillende groepen verdeeld en iedere groep gaat dezelfde tekst op een andere manier lezen. Daarna gaan alle groepen dezelfde toets maken en ga ik bekijken welke manier van lezen het beste werkt. We gaan eerst een oefening doen waarbij ik jullie de manier leer waarop jullie groep de tekst gaat lezen. Voordat we daarmee beginnen mogen jullie de introductie van de tekst lezen. Als jullie vragen hebben over het verhaal mogen jullie deze stellen.'*
- De kinderen lezen de introductie en stellen eventuele vragen. De onderzoeker beantwoordt eerst verhalen over de tekst. Daarna legt hij/zij uit hoe de leesopdracht zal verlopen. *'We gaan dadelijk de tekst lezen totdat we een oranje stip tegen komen. Lees zo zorgvuldig mogelijk. Zodra je bij de oranje stip bent aangekomen, ga je met je armen over elkaar zitten zodat ik kan zien wie klaar is met lezen. Wanneer iedereen van uit de groep klaar is met lezen, gaan we een denkoefening doen. Deze zal ik zo uitleggen. Jullie mogen beginnen met lezen*.'
- Wanneer alle leerlingen klaar zijn legt de onderzoeker de denkoefening uit. *' We gaan nu 30 seconden nadenken over het verhaal dat we gelezen hebben. In die 30 seconden proberen we het stukje dat we gelezen hebben in ons hoofd aan onszelf te vertellen. Daarbij gaan we ook gebaren maken die passen bij het verhaal. Tijdens de gebaren mag je geen geluid maken en mag je enkel je hoofd, armen en handen gebruiken. Je moet op je stoel blijven zitten. Ik zal het voordoen: Stach heeft een toongenerator in*

*het burgemeestershuis neer gezet *gebaar: zwaar voorwerp neerzetten* en hij wacht *gebaar: armen over elkaar, ongeduldig wachtend* tot de vogels *gebaar: vliegen* komen, zodat hij zijn plan *gebaar: wijst naar hoofd* kan uitvoeren. De vogels *gebaar: vliegen* komen alleen maar niet *gebaar: nee schudden*.. Net wanneer het zijn plan dreigt te mislukken *gebaar: handen in het haar*, omdat meneer Holder zijn apparaten terug wil *gebaar: voorwerp terugpakken* komen de vogels *gebaar: vliegen* gelukkig toch.* De wekker gaat na 30 seconden.

- De onderzoeker vraagt de leerlingen het tweede stuk tekst te lezen tot aan de oranje cirkel. Wanneer alle kinderen klaar zijn vraagt de onderzoeker aan de leerlingen welke dingen uit het verhaal zij zouden noemen in hun samenvatting in combinatie met de gebaren die zij zouden maken. *'Als jullie het stuk dat jullie gelezen hebben moeten samenvatten in je hoofd en bijpassende gebaren moeten maken, waar begin je dan mee'* Een kind krijgt de beurt. De onderzoeker vult het antwoord waar nodig aan. *'Hoe zou de samenvatting verder gaan? En welk gebaar past daar dan bij?'* Een ander kind krijgt de beurt. De onderzoeker vult waar nodig aan. De onderzoeker geeft kinderen de beurt totdat de samenvatting volledig is. Ten slotte somt de onderzoeker de volledige samenvatting nog op: '*Stach haalt de schakelaar om *gebaar: schakelaar omhalen*, waardoor er een geluid *gebaar: wijst naar oren* uit de speakers komt dat lijkt op het geluid van de vogels *gebaar: vliegende beweging*. Daarna draait hij de toon omhoog *draaiende beweging*, de vogels doen hem na *gebaar: vliegende beweging*. Hij blijft de toon omhoog draaien, de vogels blijven hem nadoen *draaiende beweging*. Hij gelooft dat zijn plan gaat werken *gebaar: twee duimen omhoog*.'*
- De leerlingen mogen opnieuw lezen tot de oranje cirkel. Wanneer alle leerlingen klaar zijn vertelt de onderzoeker dat zij nu zelf in hun hoofd mogen samenvatten. *'Ik denk dat jullie klaar zijn om de opdracht helemaal zelf uit te voeren. Ik zet een wekker op 30 seconden. Jullie mogen in die tijd allemaal in proberen het verhaal samen te vatten in je hoofd, op de manier waarop we geoefend hebben. Het is belangrijk dat jullie je handen op tafel houden en geen bewegingen maken tijdens het nadenken.'*
- De onderzoeker observeert of de leerlingen adequate bewegingen maken. Aan de hand hiervan kan hij/zij beoordelen of de opdracht goed is begrepen door de groep. Indien de opdracht nog niet goed begrepen is door de groep, kan de onderzoeker er voor

kiezen ook de laatste twee delen uit de oefentekst te behandelen. Indien de opdracht begrepen is kan de onderzoeker beginnen met het experiment.

Protocol Practicing session condition 3: incongruent gestures

Voorbereiding: papieren met de oefentekst uit het boek De Koning van Katoren (Terlouw, 1971) wordt op de tafels van de leerlingen neergelegd.

De onderzoeker vraagt de studenten plaats te nemen achter een van de tafels en de tekst onaangeroerd te laten liggen. Wanneer alle studenten plaats hebben genomen kan de instructie worden gestart.

- De onderzoeker begint met een korte uitleg over het onderzoek. *'Hallo mijn naam is ... en ik voer een onderzoek uit naar begrijpend lezen bij basisschool leerlingen. Jullie klas is in verschillende groepen verdeeld en iedere groep gaat dezelfde tekst op een andere manier lezen. Daarna gaan alle groepen dezelfde toets maken en ga ik bekijken welke manier van lezen het beste werkt. We gaan eerst een oefening doen waarbij ik jullie de manier leer waarop jullie groep de tekst gaat lezen. Voordat we daarmee beginnen mogen jullie de introductie van de tekst lezen. Als jullie vragen hebben over het verhaal mogen jullie deze stellen.'*
- De kinderen lezen de introductie en stellen eventuele vragen. De onderzoeker beantwoordt eerst verhalen over de tekst. Daarna legt hij/zij uit hoe de leesopdracht zal verlopen. *'We gaan dadelijk de tekst lezen totdat we een oranje stip tegen komen. Lees zo zorgvuldig mogelijk. Zodra je bij de oranje stip bent aangekomen, ga je met je armen over elkaar zitten zodat ik kan zien wie klaar is met lezen. Wanneer iedereen van uit de groep klaar is met lezen, gaan we een denkoefening doen. Deze zal ik zo uitleggen. Jullie mogen beginnen met lezen*.'
- Wanneer alle leerlingen klaar zijn legt de onderzoeker de denkoefening uit. *'We maken allemaal een draaiende beweging met onze handen, zoals ik nu voordoe. We gaan nu 30 seconden nadenken over het verhaal dat we gelezen hebben. In die 30 seconden proberen we het stukje dat we gelezen hebben in ons hoofd aan onszelf te vertellen. Ik zal het voordoen: Stach heeft een toongenerator in het burgemeestershuis neer gezet en hij wacht tot de vogels komen, zodat hij zijn plan kan uitvoeren. De vogels komen alleen maar niet. Net wanneer het zijn plan dreigt te mislukken, omdat meneer Holder*

zijn apparaten terug wil komen de vogels gelukkig toch. De wekker gaat na 30 seconden.

- De onderzoeker vraagt de leerlingen het tweede stuk tekst te lezen tot aan de oranje cirkel. Wanneer alle kinderen klaar zijn vraagt de onderzoeker aan de leerlingen welke dingen uit het verhaal zij zouden noemen in hun samenvatting. *'We maken allemaal opnieuw de draaiende beweging met onze armen. Als jullie het stuk dat jullie gelezen hebben moeten samenvatten in je hoofd, waar begin je dan mee'* Een kind krijgt de beurt. De onderzoeker vult het antwoord waar nodig aan. *'Hoe zou de samenvatting verder gaan?'* Een ander kind krijgt de beurt. De onderzoeker vult waar nodig aan. De onderzoeker geeft kinderen de beurt totdat de samenvatting volledig is. Ten slotte somt de onderzoeker de volledige samenvatting nog op: '*Stach haalt de schakelaar om, waardoor er een geluid uit de speakers komt dat lijkt op het geluid van de vogels. Daarna draait hij de toon omhoog, de vogels doen hem na. Hij blijft de toon omhoog draaien, de vogels blijven hem nadoen. Hij gelooft dat zijn plan gaat werken.'*
- De leerlingen mogen opnieuw lezen tot de oranje cirkel. Wanneer alle leerlingen klaar zijn vertelt de onderzoeker dat zij nu zelf in hun hoofd mogen samenvatten. *'Ik denk dat jullie klaar zijn om de opdracht helemaal zelf uit te voeren. Ik zet een wekker op 30 seconden. Jullie mogen in die tijd allemaal in proberen het verhaal samen te vatten in je hoofd, op de manier waarop we geoefend hebben. Het is belangrijk dat jullie met je handen de draaiende beweging maken die we geoefend hebben.'*
- De onderzoeker vraagt een aantal leerlingen hun samenvatting te delen met de rest. De groep mag indien nodig aanvullen. Aan de hand hiervan kan hij beoordelen of de opdracht goed is begrepen door de groep. Indien de opdracht nog niet goed begrepen is door de groep, kan de onderzoeker er voor kiezen ook de laatste twee delen uit de oefentekst te behandelen. Indien de opdracht begrepen is kan de onderzoeker beginnen met het experiment.

Appendix F. Observation form

Klas: School:
Conditie Aantal Leerlingen:

Taakgericht gedrag per verhaaldeel						Opmerkingen
	1	2	3	4	5	
Student ___						
Student ___						
Student ___						
Student ___						
Student ___						
Student ___						
Student ___						
Student ___						
Student ___						
Student ___						
Overige opmerkingen :						

Appendix G. FETC Form

APPLICATION FORM FOR THE ASSESSMENT OF A RESEARCH PROTOCOL BY THE FACULTY ETHICS REVIEW BOARD (FERB) OF THE FACULTY OF SOCIAL AND BEHAVIOURAL SCIENCES

General guidelines for the use of this form

1. This form can be used for a single research project or a series of related studies (hereinafter referred to as: "research programme"). Researchers are encouraged to apply for the assessment of a research programme if their proposal covers multiple studies with related content, identical procedures (methods and instruments) and contains informed consent forms and participant information, with a similar population. For studies by students, the FERB recommends submitting, in advance, a research programme under which protocol multiple student projects can be conducted so that their execution will not be delayed by the review procedure. The application of such a research programme must include a proper description by the researcher(s) of the programme as a whole in terms of the maximum burden on the participants (e.g. maximum duration, strain/efforts, types of stimuli, strength and frequency, etc.). If it is impossible to describe all the studies within the research programme, it should, in any case, include a description of the most invasive study known so far.
2. Solely the first responsible senior researcher(s) (from post-doctoral level onwards) may submit a protocol.
3. Any approval by the FERB is valid for 5 years or until the information to be provided in the application form below is modified to such an extent that the study becomes more invasive. For a research programme, the term of validity is 2 years and any extension is subject to approval. The researcher(s) and staff below commit themselves to treating the participants in accordance with the principles of the Declaration of Helsinki and the Dutch Code of Conduct for Scientific Practices as determined by the VSNU Association of Universities in the Netherlands (which can both be downloaded from the FERB site on the Intranet[1]) and guarantee that the participants (whether decisionally competent or incompetent and/or in a dependent relationship vis-a-vis the researcher or not) may at all times terminate their participation without any further consequences.
4. The researcher(s) commit themselves to maximising the quality of the study, the statistical analysis and the reports, and to respect the specific regulations and legislation pertaining to the specific methods.
5. The procedure will run more smoothly if the FERB receives all the relevant documents, such as questionnaires and other measurement instruments as well as literature and other sources on studies using similar methods which were found to be ethically acceptable and that testify to the fact that this procedure has no harmful consequences. Examples of studies where the latter will always be an issue are studies into bullying behaviour, sexuality, and parent-child relationships. The FERB asks the researcher(s) to be as specific as possible when they answer the relevant questions while limiting their answers to 500 words maximum per question. It is helpful to the FERB if the answers are brief and to the point.
6. **Our FAQ document that can be accessed through the Intranet provides background information with regards to any questions.**
7. The researcher(s) declare to have described the study truthfully and with a particular focus on its ethical aspects.

[1] See: https://intranet.uu.nl/facultaire-ethische-toetsingscommissie-fetc

THE EFFECT OF GESTURE CONGRUENCY ON THE COGNITIVE LEANING PROCESS AND LEARNING OUTCOMES

Signed for approval[2]:

Date:

[2] The senior researcher (holding at least a doctoral degree) should sign here.

THE EFFECT OF GESTURE CONGRUENCY ON THE COGNITIVE LEANING PROCESS AND LEARNING OUTCOMES

A. **GENERAL INFORMATION/PERSONAL DETAILS**

1.

a. a. Name(s), position(s) and department(s) of the responsible researcher(s):

Supervisor: Femke Kirschner, Education Development Assistant, Faculty of Social and Behavioural Sciences.

b. Name(s), position(s) and department(s) of the executive researcher(s):

Gitte van Helden, Master Student, Faculty of Social and Behavioural Sciences. (4244389)

2. Title of the study or research programme - Does it concern a single study or a research programme? Does it concern a study for the final thesis in a bachelor's or master's degree course?:

It concerns a study for the final thesis in a master's degree course.

3. Type of study (with a brief rationale):

- **experimental**
- observational
- otherwise:

4. Grant provider:

-

5. Intended start and end date for the study:

Start: January 24

End: June 11

6. Research area/discipline:

Educational Sciences

7. For some (larger) projects it is advisable to appoint an independent contact or expert whom participants can contact in case of questions and/or complaints. Has an independent expert been appointed for this study?[3]:

No

8. Does the study concern a multi-centre project, e.g. in collaboration with other universities, a GGZ mental health care institution, a university medical centre? Where exactly will the study be conducted? By which institute(s) are the executive researcher(s) employed?:

The study will be conducted at primary schools that are willing to let their students enrol in this research project. The researcher is hired as students assistant at the Freudenthal Institute for Mathematics and Science Education.

9. Is the study related to a prior research project that has been assessed by a recognised Medical Ethics Review Board (MERB) or FERB?

No

If so, which? Please state the file number:

B. SUMMARY OF THE BACKGROUND AND METHODS

Background

1. What is the study's theoretical and practical relevance? (500 words max.):

[3] This contact may, in principle, also be a researcher (within the same department, or not) who is able to respond to the question or complaint in detail. Independent is to say: not involved in the study themselves. The FERB upholds that an independent contact is not obligatory, but will be necessary when the study is more invasive.

Theoretical relevance.

We know from previous research that learning achievement can be increased by letting learners make gestures (Broaders et al., 2007; Goldin-Meadow et al., 2009; Wagner Cook & Goldin-Meadow, 2006). Less is known about the cognitive processes that might explain this increase. To effectively incorporate gesturing in instructional design, it is important to understand the effect of gesturing on cognitive processes. Knowledge of the human cognitive architecture and cognitive load theory (Sweller, van Merrienboer, & Paas, 1998) can be used to provide these insights. The current research will add to existing literature by comparing achievement and perceived cognitive load between three conditions: no gestures, congruent (meaningful) gestures and incongruent (not-meaningful) gestures. Therefore the theoretical aim of this research is to expand the knowledge of gesturing on the cognitive learning process and learning outcomes.

Practical relevance.

If gesturing indeed leads to a more effective and efficient way of learning, this could provide important opportunities for education. However, to effectively incorporate gestures in education we must know under what conditions gesturing is effective. Therefore the practical aim of the current study is to provide more insight in how gesturing can be used in educational setting in order to make learning more efficient and effective

2. What is the study's objective/central question?:

What is the effect of gesture congruency on the cognitive learning process and learning outcomes?

3. What are the hypothesis/hypotheses and expectation(s)?:

Three hypotheses are formulated. First, it is expected that performing congruent gestures leads to the lowest cognitive load during the learning-phase when compared to not gesturing and performing incongruent gestures. Performing incongruent gestures is expected to lead to the highest cognitive load. This is expected because meaningful gestures foster information processing and the construction of schemata (Golding-Meadow et al., 2001; Wagner Cook et al., 2012), while meaningless gestures impose an extraneous load on the learner and therefore can hamper the construction of schemata (Wagner Cook et al., 2012).

Second, it is expected that performing congruent gestures leads to the lowest cognitive load during the test-phase when compared to not gesturing and performing incongruent gestures. Performing incongruent gestures is expected to lead to the highest cognitive load. During the test-phase students have to recall information from the schemata they constructed during the leaning-phase. Because it is expected that congruent gestures foster the creation of schemata, it will be easier to recall information (Golding-Meadow et al., 2001; Wagner Cook et al., 2012). Good schemata reduce the load imposed on a student (Sweller et al., 1998). Performing incongruent gestures is expected to hamper the creation of schemata, which will make it harder to recall information and therefore imposes a higher load on the student (Sweller et al., 1998).

Third, it is expected that performing congruent gestures leads to a higher performance on learning recall and learning comprehension (i.e., learning outcomes), compared to not gesturing and performing incongruent gestures. Performing incongruent gestures is expected to lead to the lowest score on learning recall and learning comprehension. Previous studies showed that performing gestures lead to higher learning outcomes (Broaders et al., 2007; Wagner Cook & Goldin-Meadow, 2006). However, gestures need to be aligned with the key features of a task to increase learning outcomes (Golding-Meadow et al., 2009). Incongruent gestures might lead to a higher extraneous load and therefore lead to lower learning outcomes when executing a complex task (Sweller, et al., 1998).

Design/procedure/invasiveness

4. What is the study's design and procedure? (500 words max.):

The independent variables: (1) no gestures, (2) congruent gestures and (3) incongruent gestures.

The dependent variables: (1) cognitive load during learning phase (2) cognitive load during test phase (3) achievement on recall (4) achievement on reading comprehension.

Per class students were randomly divided in groups and equally distributed over the three conditions. For all participating classes three sessions were planned: one for every condition. The group size per session varied from 7 to 10. A session consisted of three elements: (1) a practice for reading the text (2) reading the text including measurement of cognitive load (3) conducting a post-test including measurement of cognitive load. Every session was led by the same researcher. Sessions took place one by one in a quiet room within the school. Every participant had enough space to execute gestures without distracting other students.

Practicing session. To make sure the students were acquainted with the procedure of their condition a practicing session was arranged. The practicing text was used in all three conditions. The control group was instructed to keep their hands on the table while thinking about what they had read. The incongruent gesturing group was instructed to move their hands in circles while thinking about what they had read. The congruent gesturing group was instructed to perform gestures that represented what they had read. The required behaviour was modelled by the researcher, by using the think aloud method (e.g. Silvén & Vauras, 1992; van Someren, Barnard,& Sandberg, 1994). A protocol for instruction during the practice was used.

Reading the text. Then students were asked to read the introduction of the text. They had the opportunity to ask questions when they encountered unclarities during the introduction. All three conditions started with the researcher presenting a card with 6 letters for 5 seconds. Students were asked to remember these letters. Then students were asked to carefully read the text until they encounter the first orange circle. When all students read the first story part, the researcher invited them to process the text according to the method they were taught during the instruction sessions. The control group did not move and thought about what they had read. The incongruent gesturing group moved their hands in circles while thinking about what they had read. The congruent gesturing group performed gestures that represented what they had read. After 30 seconds, students were asked to move to the next story part. This procedure was repeated for the next four story parts. After completing the reading exercise, the students were asked to note the letters they remembered on a paper.

Post-test. The post-test procedure was similar for all conditions. Before starting the test, the researcher presented a card with 6 letters for 5 seconds. Students were asked to remember these

letters. Then students were asked to individually complete the post-test questions. The maximal time available for completing the post test was 12 minutes. When all students completed the test, they were asked to write down the letters they remembered on paper.

5.

a. Which measurement instruments, stimuli and/or manipulations will be used?[4]

Gestures. *Three types of gesturing conditions were distinguished. The control group was instructed to keep their hands on the table while thinking about what they had read, to prevent them from making gestures. They were asked to sit up straight, while placing the palm of their hand in front them on the table. The incongruent gesturing group was instructed to move their hands in circles while thinking about what they had read, similar to the study of Wagner Cook et al. (2012). They were asked to sit up straight and hold their hand up in front of themselves while moving both hands in small circles. The congruent gesturing group was instructed to perform meaningful gestures that represented what they had read, while remaining seated. In order for the gestures to be meaningful for the individual, the students had to make up gestures themselves.*

Learning task. *A text was selected from 'Ronja de Roversdochter' (Lindgren, 1981) that was complex regarding the expected skill level of the participants. The story takes place in an imaginary world, so participants could not have prior knowledge regarding the book. A short introduction was included, to provide essential background knowledge for understanding the text. The text was divided in five story parts, under which the symbol of an orange circle was placed. The orange circle indicated students had to perform incongruent, congruent or no gestures, depending on the condition they were in. The text can be found in Appendix A. The learning-task was paper-based.*

Post-task. *The post-task was a test existing of ten question, that measured to what extend students could recall information and to what extend students comprehended the storyline. Five questions were formulated for both recall and comprehension (Appendix B). Recall was tested with a statement for each story part (e.g., Ronja takes along a kettle to the cave). Students had to recall whether they were 'true' or 'not true'. To increase reliability and validity, statements could be directly found in the text. Comprehension was tested with multiple choice questions (e.g. What do Ronja and Birk find worst about the unwanted visit from the Goblins? a. The goblins ate their bread, b. They are afraid the goblins will return., c. They are afraid the goblins will reveal to the bird witches where they live.). To increase reliability and validity, questions were based on three of the four skills for reading comprehension as described by CITO (2016): understanding (question 7, question 10), interpreting (question 6; question 8) and extracting (question 9). The fourth skill, 'looking up', was excluded, because students were not allowed to view the text during the post-task. The post-task was paper-based.*

Performance measurement. *Recall was measured by five statements (question 1 – 5). Comprehension was measured by five multiple choice questions (question 6-10). Participants received*

[4] Examples: invasive questionnaires; interviews; physical/psychological examination, inducing stress, pressure to overstep important standards and values; inducing false memories; exposure to aversive materials like a unpleasant film, video clip, photos or electrical stimulus; long-term of very frequent questioning; ambulatory measurements, participation in an intervention, evoking unpleasant psychological or physical symptoms in an experiment, denial, diet, blood sampling, fMRI, TMS, ECG, administering stimuli, showing pictures, etc. In case of the use of a device (apparatus) or administration of a substance, please enclose the CE marking brochure for the relevant apparatus or substance, if possible.

a score between 0 and 5 (0= no correct answer and 5= five correct answers) for both recall and comprehension.

Cognitive-load measurement. *A dual task was used to measure cognitive load. Remembering letters from paper was chosen as secondary task (Appendix C), because it meets the criteria of Brünken, Plass and Leutner (2003) for applying a dual-task approach. First, remembering the letters from paper is a simple task, that does not suppress the simultaneous leaning process. Second, remembering letters requires the same cognitive resources as the primary task, namely visual working memory. Third, remembering letters is proven to be a valid measurement of the cognitive load imposed by a task (Logan, 1979). Similar to Goldin Meadow et al., (2001) the number of letters presented to participants was 6 and participants were allowed to look at the letters for 5 seconds. The participants had to write down the letters on a blanc square on a paper, that was included in the post-task. The cognitive load was measured during the learning task and during the post-task, by assigning a score between 0 and 6 (0= highest cognitive load and 6= lowest cognitive load). The score was computed by assigning one point for every correct letter a student wrote down. When a student wrote down an incorrect letter, a half point was subtracted from the total score until the minimal score of 0 was reached.*

Practice material. *Practice material was designed to make participants acquainted with performing no gestures, congruent gestures and incongruent gestures. For this purpose five story parts similar to the learning task were selected from the book 'De Koning van Katoren' (Terlouw, 1971) (Appendix D).*

Validity and Reliability. *Validity and reliability were assured in various ways. First the comprehension questions were based on questions of the CITO test for learning comprehension, which is considered a reliable and valid test (Tomesen, Engelen, & Hiddink 2018). Recall questions could be directly traced back to a feature of every story part and were therefore a valid indication of the facts students could recall from the story parts.*

In order to increase the reliability, the researcher followed a detailed protocol during explanation of the experiment and practice sessions in all three conditions (Appendix E). The learning task, post-task and protocol were reviewed by five primary school teachers, who combined their job with an Educational Sciences master's programme at Utrecht University, during a feedback session. The learning task and post-task were judged on appropriate complexity. In response to this feedback session, some small changes were made in the post-task. For example the answer a. for question 6 was changed from 'amusing story' in to 'fantasy story', because students were possibly not familiar with the term 'amusing'. The protocol was judged on practical attainability. One change that was made in response to the feedback session was the addition of the preparation preliminary to the practice sessions.

Furthermore, a pilot of the experiment was executed for the congruent and incongruent condition. Both trials included two students that were representative for the sample. The pilot gave no rise to adaptions in the learning task, post-task and protocol. During the pilot it was decided to make use of an observation form (Appendix F), to verify whether all students showed the desired behaviour. Students who did not perform the appropriate behaviour, thus performing congruent, incongruent or no gestures depending on the condition they were in, were excluded from the study. Despite these efforts the internal consistency of the post-task can be considered low, with a Cronbach's alpha of .08 for recall and a Cronbach's alpha of .20 for comprehension (Tavakol & Dennick, 2011). This is possibly due to the fact the post-task existed of a limited amount of multiple-choice items. When the number of test items is too small, the assumption of tau-equivalence is violated, which leads to an underestimation of the reliability (Tavakol & Dennick, 2011).

b. What does the study's burden on the participants comprise in terms of time, frequency and strain/efforts?:

A session will take approximately one hour. Every participant will participate in one session.

c. Will the participants be subjected to interventions or a certain manner of conduct that cannot be considered as part of a normal lifestyle?:

No, the task is closely related to task students execute within the school context.

d. Will unobtrusive methods be used (e.g. data collection of uninformed subjects by means of observations or video recordings)?:

All participants and their parents will be informed about the collection of data.

e. Will the study involve any deception? If so, will there be an adequate debriefing and will the deception hold any potential risks?:

The study will not involve deception. Participants and their parents will be informed about the purpose of the research.

6. Will the participants be tested beforehand as to their health condition or according to certain disorders? Are there any inclusion and/or exclusion criteria or specific conditions to be met in order for a participant to take part in this study?:

Students with reading disabilities will be excluded from this study. Also, students who had already read the book Ronja de Roversdochter will be excluded, because their prior knowledge could influence the results.

7. Risks for the participants -

a. Which risks does the study hold for its participants?:

The study holds no risk for participants.

b. To what extent are the risks and objections limited? Are the risks run by the participants similar to those in daily life?:

8. How does the burden on the participants compare to the study's potential scientific contribution (theory formation, practical usability)?:

The burden on the participants is relatively small. The potential value of the study is advantageous, because it provides insight in cognitive mechanisms that are involved when a learner makes gestures during a task. Also, result provide practical guidelines for the use of gestures within the context of primary school.

9. Will a method be used that may, by coincidence, lead to a finding of which the participant should be informed?[5] If so, what actions will be taken in the case of a coincidental finding?:

[5] For instance: dementia, dyslexia, giftedness, depression, extremely low heartbeat in an ECG, etc. If coincidental findings may be found, this should be included in the informed consent, including a description of the actions that will be taken in such an event.

No

Analysis/power

10. How will the researchers analyse the data? Which statistical analyses will be used?:

A MANOVA will be conducted, for making a comparison between the three conditions (dependent variables) on (1) cognitive load during learning phase, (2) cognitive load during test phase, (3) achievement on recall and (4) achievement on reading comprehension.

11. What is the number of participants? Provide a power analysis and/or motivation for the number of participants. The current convention is a power of 0.80. If the study deviates from this power, the FERB would like you to justify why this is necessary:

A power analysis presented the desired amount of participants is 126. This comes down to 42 participants per condition. It was be attempted to reach the desired amount of participants, but due to the limited time available only 94 participants enrolled in this research.

C. PARTICIPANTS, RECRUITMENT AND INFORMED CONSENT PROCEDURE

1. The nature of the research population (please tick):

1. General population without complaints/symptoms

2. General population with complaints/symptoms

3. Patients or population with a diagnosis (please state the diagnosis)

2. Age category of the participants (please tick):

- 18 years or older
- 16-17 years
- 13-15 years
- **12 years or younger**

3. Does the study require a specific target group? If so, justify why the study cannot be conducted without the participation of this group (e.g. minors):

The study requires primary school students as participants. The relevance of this research is to gain insight in cognitive processes that accompany gesturing. Previous research on the effects of gesturing on learning outcomes took place within primary school context. A similar target group must be chosen in order to add to this existing research.

It was chosen to let 5th grade students (groep 7) enrol in the program, while the students should not be too young. Measuring cognitive load with very young children is intrusive, while they have few prior knowledge. Therefore processing information is harder for them and general assumptions of information processing do not apply to them.

4. Recruitment of participants -

a. How will the participants be recruited?:

Primary schools will be contacted and asked whether they want their students to enrol in this research project. If the school agrees to participate, parents of potential participant will be informed via e-mail and a letter. Distribution of the letter and e-mail will be done by the student's teachers. The letter and email will contain a form, which parents have to return. On this form parents have to choose whether they want they child to enrol in this study.

b. How much time will the prospective participants have to decide as to whether they will indeed participate in the study?:

Parents of students will have at least three weeks to let the teacher know whether want their child to participate in this research, by filling in the form.

5. Does the study involve informed consent or mutual consent? Clarify the design of the consent procedure (who gives permission, when and how). Does the study involve active consent or passive consent? If no informed consent will be sought, please clarify the reason:

The school gives active permission to let students enrol in the research and provides a place where the research can be executed. Parents receive an informed consent letter. They can return this letter with or without consent for letting their child enrol in the research (active consent).

6. Are the participants fully free to participate and terminate their participation whenever they want and without stating their grounds for doing so?:

The parents of participants can decide to leave the research at any time.

7. Will the participants be in a dependent relationship with the researcher?:

No

8. Compensation

a. Will the participants be compensated for their efforts? If so, what is included in this recompense (financial reimbursement, travelling expenses, otherwise). What is the amount?

No

b. Will this compensation depend on certain conditions, such as the completion of the study?

D. PRIVACY AND INFORMATION

THE EFFECT OF GESTURE CONGRUENCY ON THE COGNITIVE LEANING PROCESS AND LEARNING OUTCOMES

1.

a. Will the study adhere to the requirements for anonymity and privacy, as referred to in the Faculty Protocol for Data Storage[6]?:
   - anonymous processing and confidential storage of data (i.e. storage of raw data separate from identifiable data): **yes**/no
   - the participants' rights to inspect their own data: **yes**/no*
   - access to the data for all the researchers involved in the project: **yes**/no

If not, please clarify.

*Students' parents also have the right to inspect their children's data if desired.

b. Has a Data Management Plan been designed?

Yes

2.

a. Will the participant be offered the opportunity to receive the results (whether or not at the group level)?:

Participating schools will receive a summary of the results. These results will be communicated to students and their parents.

b. Will the results of the study be fed back to persons other than the participants (e.g. teachers, parents)?:

Yes, the participating schools may choose to communicate the results to all parents in their newsletters or on their website.

If so, will this feedback be provided at the group or at the individual level?

Results will be provided at group level.

[6] This can be found on the Intranet: https://intranet.uu.nl/wetenschappelijke-integriteit-facultair-protocol-dataopslag

3.

a. Will the data be stored on the faculty's data server?: **yes**/no

b. Will the data that can be traced back to the individual be stored separately on the other faculty server available for this specific purpose?:

Yes

If not, please clarify where will the data be stored instead?:

E. ADDITIONAL INFORMATION

F. FORMS TO BE ENCLOSED (CHECKLIST)

- Text (advert) for the recruitment of participants
- Information letter for participant
- Informed consent form for participant

Beste ouder/ verzorger,

Even voorstellen
Mijn naam is Gitte van Helden en ik volg de master Educational Sciences aan Universiteit Utrecht. Voor mijn masterthesis doe ik onderzoek naar het gebruik van gebaren in de klas. Middels deze brief vraag ik toestemming voor deelname van uw kind aan dit onderzoek.

Wat is het doel van het onderzoek?
We maken allemaal gebaren wanneer wij tegen iemand praten. Uit eerder onderzoek is gebleken dat het maken van gebaren ook degene die vertelt helpt om informatie te verwerken. Ik wil onderzoeken hoe wij dit op basisscholen in kunnen zetten om kinderen beter te laten leren. Helpt het maken van gebaren hen om lesstof te verwerken of is het toch te afleidend?

Wat houdt het onderzoek in?
Het onderzoek zal bestaan uit het lezen van een tekst, het verwerken van de tekst met behulp van gebaren en het maken van de bijbehorende vragen. De activiteit zal maximaal 1 uur duren en zal op school plaatsvinden. Er worden geen video- of geluidsopnamen gemaakt.

Privacy en vertrouwelijkheid
Alle gegevens worden vertrouwelijk behandeld en anoniem verwerkt. De docent krijgt de antwoorden van de leerlingen niet te zien. De gegevens worden alleen voor opleidings- en onderzoeksdoeleinden gebruikt. Leerlingen kunnen zelf ook aangeven of ze wel of niet mee willen doen.

Mogelijkheid tot vragen, informatie en toestemming
Heeft u naar aanleiding van deze brief nog vragen en/of wilt u op de hoogte gehouden worden van dit onderzoek, neem dan contact op met g.vanhelden@students.uu.nl.

Omdat wetenschappelijk onderzoek naar leren zeer afhankelijk is van vrijwillige deelname, is dit onderzoek enorm gebaat bij de deelname van uw kind. Uw kind kan alleen deelnemen aan dit onderzoek wanneer u hier toestemming voor geeft. Wilt u daarom dit formulier ondertekenen en meegeven aan uw kind? Alvast bedankt!

Met vriendelijke groet,
Gitte van Helden

✂----------

Ik geef mijn kind wel / geen toestemming om deel te nemen aan dit onderzoek.

Uw naam: Datum:

Naam kind: Handtekening:

School/klas kind:

THE EFFECT OF GESTURE CONGRUENCY ON THE COGNITIVE LEANING PROCESS AND LEARNING OUTCOMES

Signature(s):[7] Date and place:

Name, position:

[7] The senior researcher (holding at least a doctoral degree) should sign here.