# Grounding Erosion: Are Congruent Gestures and Learning Related?

Ilse J.P. Koffijberg, 3684733

SEC - 45 ECTS

Utrecht University

Supervisors:

Arthur Bakker, Rosa A. Alberto and Nico P.G. Rutten

Utrecht University

Second Examiner:

Wouter R. van Joolingen

Utrecht University

## Author Note

Ilse J.P. Koffijberg, Freudenthal Institute, Utrecht University; Arthur Bakker, Freudenthal Institute, Utrecht University; Rosa A. Alberto, Freudenthal Institute, Utrecht University; Nico P.G. Rutten, Freudenthal Institute, Utrecht University; Wouter R. van Joolingen, Freudenthal Institute, Utrecht University.

Nico P.G. Rutten is now at Department Health & Wellbeing, Saxion University of Applied Sciences, Enschede.

Special thanks to Mathijs Booden, Department of Education, University of Amsterdam, for making available a mobile Augmented Reality Sandbox; the teachers and students at the participating schools; all supervisors, who spent a lot of time proof reading and guiding me throughout the whole process; and my brother who can charge my mental battery with just a smile.

Correspondence concerning this article should be addressed to Ilse Koffijberg, Bisschopssteeg 455, 3584SH Utrecht. E-mail: i.j.p.koffijberg@uu.nl

SEC requirement: kind of journal: Journal of Experimental Psychology: Learning, Memory and Cognition

## Abstract

A quasi-experimental design was set up to determine whether congruent gestures are related to learning, when learning about erosion facilitated by an Augmented Reality Sandbox. Congruent gestures are gestures that are analogically related to the meaning of a concept, in this case erosion. It is thought that by using congruent gestures, students ground erosion in the brain's modal system, thereby making erosion immediately meaningful to them. A total of 62 participants (84% male, 16% female, aged 11-15) from two secondary schools (VMBO and VWO) was randomly divided into the hands condition (Hands) or scoops condition (Scoops). Hands were able to manipulate sand with their hands, whereas Scoops received a sand scoop to manipulate sand with. It was assumed that these conditions stimulated either congruent gestures (hands condition) or incongruent gestures (scoops condition). Pre- and posttest scores were analyzed as well as video recordings. Results show that there is no significant difference in learning gain between Hands and Scoops ($t(60) = 1.42$, $p = .162$), but there was an effect size of $d = 0.36$. Furthermore, Scoops spend more time making congruent gestures than Hands. The discussion speculates about the possible negative impact that not being able to touch sand might have had on learning gain.


*Keywords:* grounding; embodied cognition; congruence; gestures; geography education

Science education has become increasingly important with the rise of new technological developments in society. Nowadays, it is hard to imagine a technology-free society. These new technologies may have great potential for overcoming complex (global) problems, however they are not without risk. Hence, an understanding of science is needed to make decisions about these problems and to fully understand our society (the cultural and democratic argument). Other reasons for science education are its (science's) practical use for individuals and the need of science to sustain our advanced industrial society (utilitarian and economic argument) (Osborne & Hennessy, 2003; Thomas & Durant, 1987). Given the rising importance of science education, it is worrisome that students' performances of science have declined between 2006 and 2015 in The Netherlands (OECD, 2016). Though this paper will not investigate specific causes for the decline over recent years, it is well known that science is perceived as hard to learn. Science education is abstract, fragmented and often not connected to daily-life experiences (Millar, 1991; OECD, 2016; Osborne & Collins, 2000).

In recognition of the aforementioned problems, there has been a shift from "chalk and talk" activities to activities that are authentic, involve real-world problems, need collaboration and are meaningful for students (Barron & Darling-Hammond, 2008; European Commision, High Level Group on Science Education, 2007). This approach is driven by the ideology of (social) constructivists who hold that cognition is actively (and socially) constructed. Scientific knowledge is no exception in this view. In the last decades however, another, promising field to understanding cognition has received increasing attention from scholars: the field of embodied cognition. Although (social) constructivists and embodied cognitivists share a rejection of behaviorism, (social) constructivists continued to see "the mind [. . . ] as an isolated information processor" (Marshall, 2016, p.246). In contrast, embodied cognitivists believe that cognition "arises from the interaction between a brain, a body and an environment, and cannot properly be studied as a product of the mind/brain alone" (Beer, 2003, p. 211). New technologies such as augmented, mixed and augmented reality promote physical movement and can therefore facilitate embodiment (Lindgren & Johnson-Glenberg, 2013).

Several attempts to implement embodiment in education have been made across different disciplines. Abrahamson and Lindgren (2014) discuss the Mathematical Imagery Trainer and MEteor. Whereas the former is a device that allows students to enact proportional progression by gestures, the latter is a mixed reality game that allows students to enact a meteor's motion through space with their whole body. Students who used the Mathematical Imagery Trainer or MEteor had a better understanding of mathematical/scientific concepts, than their peers who could not ground these concepts physically. From these two examples alone, it should already become clear that technologies, used to aid embodiment, differ in their abilities and limitations. Moreover, these technologies differ in *how* they help to embody concepts.

A helpful tool to assess the embodiment by technologies is the taxonomy provided by Johnson-Glenberg, Birchfield, Tolentino, and Koziupa (2014). This taxonomy is based on how much a technology affords motoric engagement, congruency and immersion. The higher these three aspects score, the higher the degree of embodiment. Not only the taxonomy for embodied learning recognizes the importance of congruency, but other studies also found that congruency is beneficial for learning (f.e. see Dijkstra, Kaschak, & Zwaan, 2007; Jang, Vitale, Jyung, & Black, 2017; Segal, Tversky, & Black, 2014). However, congruency often remains undefined, and so physical and conceptual congruency is accomplished in different ways. A study that does define congruency is Lindgren and Johnson-Glenberg (2013). Congruency, according to them, means that: "[...] actions must be structurally or analogically related to [...] symbols and their meaning..." (p. 466). A congruent motion in physics would be to move one's forearm with a fixed shoulder, when learning about a pendulum's motion (Figure 1)(Lindgren & Johnson-Glenberg, 2013).

The abovementioned studies almost all focus on solving educational problems in mathematics and physics. Educational research specific to geoscience is relatively young and therefore not as extensive compared to other sciences (King, 2008). Nevertheless, research in embodiment is carried out in this field as well. Kastens, Agrawal, and Liben (2008), who studied the role of gestures in geoscience, think gestures are important because they are able to show 3-dimensional properties, they can take into account time and because they can convey continuity. Kastens, Agrawal, & Liben recommend creating situations that foster student gesturing. Unfortunately this study does not look at congruency. Another example is the study of Woods, Reed, Hsi, Woods, and Woods (2016). They found that an Augmented Reality Sandbox can enhance students' spatial thinking and modeling abilities. An Augmented Reality Sandbox consists out of a computer or laptop that is connected to a projector and Kinect camera. Beneath the Kinect camera and projector a box with sand is placed (Figure 2). This enables the camera to register the distance between the sand and camera. This data is transformed into a contour map by the computer, which in turn is projected by the projector. The Augmented Reality Sandbox offers a tacit environment, which students can physically manipulate in the "real" world (the sand). In the meanwhile a 2-dimensional overlay of a topographical map is continuously projected on top of the sand. The 2-dimensional overlay is updated in real-time. In other words, when a user would, for example, create a mountain – the map will show warm colors indicating height. In contrast when valleys are created, cool colors are displayed. Though this study recognizes the positive effects of embodiment, it does not look at congruency.

The aim of this study is to explore whether congruent gestures, made by aid of a new technology, can be an effective way to ground concepts in geoscience. This study thus combines the insights of the two above mentioned studies. Not only using new technologies that facilitate embodiment, but also making use of congruent gestures. The idea is that by embodying a concept in this way, it becomes meaningful and less abstract for students. Consequently, this would result in a better science understanding. The concept that will be grounded in this study is erosion: the wearing down of soil by moving water or ice. This concept is chosen because not much research has been done in geoscience on students ideas about weathering and erosion, even though students experience difficulties with using these terms (Dove, 1997). Erosion is also an important term in the Dutch geography curriculum (College voor Toetsen en Examens, 2015). Finally, and most importantly, erosion is a concept that lends itself to be congruently embodied. This can be achieved by making gestures that represent moving water or ice[1] and would be in accordance with the definition of Lindgren and Johnson-Glenberg (2013).

The research question is formulated as follows: “Are congruent gestures and learning related, when learning about erosion facilitated by an Augmented Reality Sandbox?" Students are placed in an experimental group or a control group. Whereas the experimental group is stimulated to make congruent gestures, the control group is not. It is expected that there is a positive relation between congruent gestures and learning gain, because the concept that is being grounded would become more meaningful and less abstract to students. In this research design the relationship between congruent gestures and learning is measured by the relative amount of time spent on congruent gestures and learning gain scores. So, it is expected that the experimental group has higher learning gain scores than those in the control group. The rationale behind these expectations is presented in the Theoretical Background. If such a relation is found, it would provide designers and educators with concrete tools on how to embody tasks.

## Theoretical Background

Although the field of embodied cognition is broad and there are multiple approaches to embodiment, there is a shared conviction. The belief "that psychological processes are influenced by the body, including body morphology, sensory systems, and motor systems" (Glenberg, 2010, p. 586). This is in sharp contrast with the traditional (computational) cognitivists who place thought processes, and thus cognition, in the brain alone (Shapiro, 2011). A widely used teaching trainer textbook states that the most important assumption cognitivists hold is that "mental processes exist, [and] that they can be studied scientifically and that humans are active participants in their own acts of cognition" (Ashcraft, 2006 cited by Woolfolk, Hughes, & Walkup, 2013, p. 290). Thus, on this view, cognition is symbol manipulation by mental processes, which in

[1] Images of examples will be presented in the Methods

turn are computational. Note, that most contemporary approaches to overcome problems in science education, such as the (social) constructivists mentioned earlier, are based on this view.

However, seeing cognition as symbol manipulation is problematic. For if thought processes are isolated in the brain and disconnected from the world, and if thought processes are nothing more than manipulations of symbols, then how do these symbols acquire meaning? This problem is known as the 'symbol grounding problem' (Harnad, 1990).

Several scholars suggested solutions for this problem. One of these solutions is the idea of perceptual symbols: "records of the neural states that underlie perception" (Barsalou, 1999, p. 581). These created symbols are multimodal; they are created for experiences that arise from vision, audition, haptics, olfaction and gustation (Barsalou, 1999)[2]. Amodal symbols also exist. The difference between them is that modal symbols are self-referential entities, whereas amodal symbols are not. Because amodal symbols are not self-referential entities, they are arbitrary. Yet, according to the mainstream ideas about cognition rather than storing experiences in direct modal perceptual symbols, experiences get transduced into arbitrary amodal symbols. Barsalou (1999) argues that there is little empirical evidence for this and that we do not need these amodal symbols; "[c]ognition is inherently perceptual" (Barsalou, 1999, p. 577).

Phrased in another way this is called the process of grounding: "the environment, situations, the body, and simulations in the brain's modal systems ground the central representations in cognition" (Barsalou, 2010, p. 717). Symbols only become meaningful through our interaction with the world (Glenberg, 2010). This is not a new idea and it can be put in the tradition of empiricist thinkers (Hume, 1777; Kant, 2004).

The use of congruent gestures is one way to ground concepts in cognition. Segal et al. (2014) nicely phrase it as: "[w]hat we *do* can influence thought [. . . ]" (p. 124). They found that congruent actions enhanced children's performance in addition and number

[2] This is in line with the definition given by Harnad (1990): "A perceptual symbol is a record of the neural activation that arises during perception" (p. 583).

line tasks. In the addition task, students had to solve addition problems. Congruent actions consisted out of tapping the screen as often as the sum of the addition, plus one to select the correct answer. In the line number task, students had to guess where a number was located on a number line. Here the congruent action was to dragg a red line to the correct position. In another research participants either saw a video that explained the working of a four-stroke engine supported by *action gestures* or *structure gestures*. Whereas action gestures resembled actions of mechanical parts, structure gestures resembled the form of mechanical parts. Here again, congruency led to deeper understanding. Note that *watching* congruent motions led to this effect. A study by Jang et al. (2017) found that students who were able to rotate images of inner body organ structures, by using an object that was coupled with the image performed better than peers who watched the same images, but who did not rotate the images. They point out the beneficial learning effects of being able to make physically congruent motions (the rotating of the images). As these examples make clear, there are different ways to accomplish congruency. What unifies these examples is that in all cases the action and to be learned concept must be similar. As mentioned before, this research will make use of the definition of Lindgren and Johnson-Glenberg (2013): "[. . . ] actions must be structurally or analogically related to the symbols and their meaning. . . " (p. 446). In this study, it means that moving water or ice is resembled by a "shoving" motion of the hand.

In summary, embodied cognition holds that meaning of the world is created through interaction of body, brain and environment. There is no need for explaining cognition in terms of mental representations because interaction results in modal perceptual symbols, which are self-referential. As a consequence, what we do can influence thought. Thus it makes sense that cognition about concepts, such as erosion, can be grounded in congruent (action) gestures and may be beneficial for learning.

# Methods

To answer the research question “are congruent gestures and learning related, when learning about erosion facilitated by an Augmented Reality Sandbox?” a quasi-experimental design was used. Students were divided into either the hand condition (Hands) or the scoops condition (Scoops). A pre- and posttest was administered immediately before and after the intervention. The intervention entailed two tasks: one task was about a valley and the other about a meandering river. Students used an Augmented Reality Sandbox to carry out the tasks. However, whereas Hands could use their bare hands to manipulate sand in the Augmented Reality Sandbox, Scoops had to use a sand scoop. Next to collecting pre- and posttest results, video recordings were also made. Statistical analyses were done on the pre- and posttest scores, while gestures were coded by means of the video records.

## Participants

Schools were invited to participate in this study through a short informative article published on the Dutch online platform for Geography Pedagogics (Koffijberg, 2017). Two secondary schools enrolled. One of these schools, A, was not a traditional school institute, but an institute that is specialized in maritime and transport. It offers programs for both secondary and higher education. The 59 participants (85% male, 15% female, aged 11-14 based on students’ school year) used from this school were enrolled in the Dutch VMBO level (which is preparing students for vocational education). The other school, B, was a regular secondary school. It offered programs for all Dutch secondary school levels. The 12 participants (67% male, 33% female, aged 13-15 ($M$ = 14.17, $SD$ = 0.58)) used from this school were enrolled in the Dutch VWO level (which is preparing students for pre-university education). The total sample size was $N = 71$, however 9 students of school A were excluded (7 males and 2 females) because they did not complete the posttest, resulting in $N = 62$. All students were familiar with erosion, valleys and meandering rivers before participation in this research. It has not been investigated how or when these terms were taught. This means that all students had some preknowledge about erosion before enacting erosion, rather than being tabula rasa. Hence, their knowledge about erosion after the activity should be composed of different sources and not purely created by grounding the concept.

## Intervention

Groups of four students (made by the teacher) first made the pretest. Thereafter, students received a brief explanation on how the Augmented Reality Sandbox works and were assigned roles. The roles corresponded to the experiment and control condition (Hands or Scoops). They were told that students in the hands condition were allowed to manipulate the sand in the sandbox only by means of their hands, while students in the scoops condition were only allowed to manipulate the sand in the sandbox using a sand scoop. The sand scoop had to be attached to a glove by Velcro (Figure 3). After these general instructions, students received written instruction abouttask 1 and task 2 depending on their role. The tasks for both groups were the same, but the order they made the tasks in differed. For task 1, Hands first had to instruct Scoops, after five minutes Scoops had to instruct Hands. For task 2 this order was reversed: Scoops first had to instruct Hands and after five minutes Hands had to instruct Scoops (Figure 4). The order was reversed to prevent an effect that might arise due to either first instructing peers or manipulating the sand. After finishing the tasks, a posttest was administered. The choice to work with instructors and manipulators arose from the desire to keep the assignment as close as possible to a regular class activity, in which students work and discuss problems together, and to collect speech data. Speech can help to clarify students' intentions of gestures and their reasoning and thus interpreting video records.

## Tasks

In task 1, as well as in task 2, students had to simulate the process of erosion. The instructing party had to guide their co-students in simulating erosion by means of speech. The task sheets were build up in three sections: information, hints and logbook. The informational text explained the goal of the assignment (simulating the process), the hints consisted of keywords and pictures, and the logbook section asked students to write down answers to the questions: What did you instruct your peers to do?, What did your peers do?, and What was the effect of your instruction on the landscape? in the designated boxes upon hearing an alarm that went off every minute. Two small adjustments have been made in the tasks between data collection at school A and B. The first one is the addition of the following sentence: Make sure that your peers are not changing the landscape in a way that is impossible by nature. This sentence was added to emphasize the goal of simulating the processes rather than replicating the end stage. For the same reason, the second adjustment was made: replacement of the pictures (Figure 5 and Figure 6).

**Task: IJssel valley.** The informational text informed students that the IJssel valley was shaped by glaciers throughout thousands of years. It explained that they had to guide their co-students, by means of speech, to simulate the processes that were involved during the creation of the IJssel valley. To help them they were given the keywords: glacier, weathering, transport, and erosion. The two images were: one schematic representation of a glacier "pushing" away land (Figure 5a) and one contour map (without contour lines, but in color) of the IJssel valley in the present (Figure 5b). At school B these two pictures were replaced by a cross-section of land with a glacier and a cross-section of that land when the glacier had melted and the valley was shaped (Figure 5c). Lastly, they were explained that every minute an alarm would go off and asked to write down the answers to the questions in designated boxes.

**Task: A meandering river.** The informational text informed students that it can take up to thousands of years before meandering rivers are formed. It explained that they had to guide their co-students, by means of speech, to simulate the processes that were involved during the creation of meandering rivers. To help them they were given the keywords: current velocity, erosion, and sedimentation. One of the images was a schematic representation of a meandering river. This image showed the path of the fastest current flow, the places where erosion and sedimentation occur, and in what direction the path of the river moves (Figure 6a). The other image was a Google Maps screenshot of a meandering river near Doesburg (place in The Netherlands) containing the map view as well as the satellite view (Figure 6b). At school B these two pictures were replaced by a schematic formation of a meandering river in three steps. This image only showed the path of the fastest current flow (Figure 6c). Lastly, they were explained that every minute an alarm would go off and asked to write down the answers to the questions in designated boxes.

**Pre- and posttest**

Pre- and posttest consisted out of 11 questions. A selection of five items was made from the Science Assessment bank developed by the American Association for the Advancement of Science (AAAS) (`http://assessment.aaas.org/`). This data bank includes more than 600 items for middle and early high school students in the domains of life science, physical science, earth science, and nature of science. The five questions used in pre- and posttest were selected from the topic 'Weathering, Erosion and Deposition'. The remaining six items were selected from the Topographical Map Assessment developed by researchers of the Spatial Intelligence and Learning Center (Jacovina, Ormand, Shipley, & Weisberg, 2014). This assessment included questions such as: draw the route of a river between point A and B, given this map. Both the questions of the AAAS Science Assessment bank and the Topographical Map assessment are widely in the literature and thus validated (Atit, Weisberg, Newcombe, & Shipley, 2016; Gagnier et al., 2016).

The pre- and posttest were supposed to have different, yet similar, questions, in order to reduce a possible testing effect. Unfortunately a human error was made. Therefore three questions were in both the pre- and posttest handed out at school A. It concerned two items of the AAAS Science Assessment and one item of the topographical map assessment. For school B these three items have been replaced by different, yet similar questions. Appendix A provides an overview of the pre- and posttest given at school A and B.

Tests were scored using the original keys, students could score one point for each correct (sub)question. This corresponded with the original keys except for question TMA2 (question 5 in both pretests), originally students had to answer two questions but could only score one point. This has been changed to two points. The test scores were calculated as percentages: scored points divided by the maximum points times 100.

### Augmented Reality Sandbox

For school A and B different Augmented Reality Sandboxes were used because of logistic reasons. However, both Sandboxes made use of the software provided by UC Davis W.M. Keck Center for Active Visualization in the Earth Sciences (n.d.). The device used at school A contained a weaker graphical card, therefore it could not run water simulation. However, the water simulation was not needed for the tasks. The dimensions of the sandbox also differed. Figure 7 shows the sandboxes used at school A and B.

### Statistical Analyses

To assess the relation between congruency and learning gain, the pre-to-posttest scores were compared for Hands and Scoops, with the assumption that Hands mostly made congruent gestures and Scoops incongruent gestures. The comparison was performed with an independent samples $t$ test using learning gain means (posttest mean scores minus pretest mean scores). Next to $t$-test values, effect sizes of learning gain have also been computed and a Bayes factor. Bayes factors are a way to quantify relative support for a hypothesis over another hypothesis. Therefore, you need a way to quantify your alternative hypothesis, this is done with a Cauchy prior. This study used a Cauchy prior of 0.707 (which is the commonly accepted value)(for a short accessible introduction, see Lakens, 2016). A repeated measured ANOVA with time (mean of the pretest and mean of the posttest) was used to identify relations between learning gain, the conditions and school. For all statistical tests, an alpha level of .05 was used. The probability of Hands having a higher learning gain than Scoops was calculated by randomly matching the learning gains of Hands and Scoops 10.000 times, using a program written in C (Appendix B). The grading of the open questions in the pre- and posttest (question TMA2 and TMA10 in the pretest, and questions TMA11 and TMA12 in the posttest) were checked with a second rater to ensure reliability. The first and second rater substantially agreed about the correction of the open questions ($k_{pre}$ = .66 and $k_{post}$ = .77). A post-hoc power analysis, made with GPower 3.1, revealed the power of the study was 0.29.

## Video Analysis

Students' gestures were analyzed to test the assumption that Hands mostly made congruent gestures and Scoops incongruent gestures. For this purpose, the video recordings were analyzed with BORIS video/audio coding software. Only 6 out of 18 videos were analyzed due to time limitations. Three videos were randomly selected from school A, while the other 3 videos were randomly selected from school B. The gestures of each student in the videos were coded according to a scheme consisting of four codes: to undo, to scoop, to shove, and to form (Table 1). For example, form would be applied when students would "pet" the sand in order to make it smoother, or when they slightly pressed sand to get it in a specific shape. They were "forming" the landscape with their bare hands (or scoop). For a visual representation, see Figure 8, which is made by taking several stills from the video recordings. Only shoving was defined as a congruent motion. The coding was done using "state events" as opposed to "point events". State events take the duration of a gesture into account, whereas point events only count the use of a gesture. Since some gestures are brief and can quickly be made in repeating motions (think again of petting sand), they can drive up the counter quite fast, even though during the experiment a student may barely have used the gesture. Therefore the relative time spent on each gesture was calculated for Hands and Scoops and these percentages were compared. One fourth of the videos in which students' gestures were coded, were also coded by a second rater, the inter-rater agreement was $k = 0.69$.

# Results

## Learning Gain

To figure out whether congruency is related to learning, it was explored whether students in the hands condition learnt more than students in the scoops condition. This was done by comparing the difference in pre- and posttest scores (learning gain) in a $t$ test. Table 2 shows the mean and standard deviations of the pre- and posttest scores, as well as the mean and standard deviation of the learning gains. We found that there was no statistically significant difference between the groups ($t(60) = 1.42$, $p = .162$).

The setting of school A and B were quite different. Whereas school A was a non-traditional school, B was a typical school. Moreover students were enrolled in different levels of education and not in the same year. Therefore it could be possible that the intervention was better suited for one type of school. Thus we wanted to know whether Hands learnt more than Scoops at school A and B respectively. Table 2 also shows the mean and standard deviations of the pre- and posttest and learning gain for each school separately. Both Hands and Scoops scored higher at the pre- and posttest at school B than Hands and Scoops at school A. They also have a higher learning gain, especially Scoops. Nevertheless, no statistically significant difference in learning gain between Hands and Scoops at school A ($t(48) = 1.28$, $p = .208$), nor at school B ($t(10) = 0.62$, $p = .550$) was found.

Because significant levels depend on sample sizes, and the sample size in this data set was small, effect sizes were also computed to compare the differences in learning gain between Hands and Scoops. Effect sizes are independent of sample sizes. They were computed for all Scoops and Hands, and again for each school separately. In all three cases an effect size of $d = 0.36$ was found, though it cannot be ruled out that this effect is due to chance rather than following treatment in either the hands or scoops condition.

Next to frequentist statistics, Bayesian statics were also computed. Bayesian statistics are better suited for small group sizes. A Cauchy prior of 0.707 was used to compute a Bayes factor for the alternative hypothesis versus the null hypothesis, using the BayesFactor package for R. A Bayes factor of $BF_{10} = 0.59$ was found. $BF_{10}$

indicates how much more likely the alternative hypothesis is, given the data, compared to the null hypothesis. For example a Bayes factor of 1 indicates that the data is equally likely under the alternative and null hypothesis. In this case, with $BF_{10} = 0.59$, the null hypothesis is supported but not convincingly (Wetzels et al., 2011).

Lastly 10.000 simulations (see Appendix B for the code) were ran, to find out how often Hands scored a higher learning gain than Scoops on average, given the observed data. Therefore, the simulator randomly matched the learning gain of all Hands and Scoops and returned how often Hands scored a higher learning gain than Scoops. The probability of Hands outperforming Scoops was 58% (18.4 out of 32).

So far, only a comparison between the learning gain of Hands and Scoops was mad. Both frequentist, and Bayesian statistics support the null hypothesis. However, the Bayes factor only supports the null hypothesis slightly and frequentist statistics are not very suited for our small sample size. Given the effect sizes, other relations that may have had an effect on learning gain scores were also explored.

To identify the relations between learning gain, the conditions and school setting, a repeated measures ANOVA was conducted with time (mean of the pretest and mean of the posttest) as within subject factors and condition (Hands and Scoops) and school (A and B) as between subject factors. Results showed a significant main effect of Time, in contrast the interaction effect of condition and school, respectively, was not significant (Table 3). This suggests that the intervention (making of test + doing the learning activity) stimulated learning about erosion, though the reason is unclear. The results also showed a significant effect of school setting (Table 3). Thus, students at school B scored higher at the pretest and posttest, than students at school A. This means that test scores were influenced by the school setting. No effects were found between Hands and Scoops, the results therefore indicate that congruency is not related to learning gain.

**Gestures**

Differences in Hands and Scoops were explored to find out whether congruency is related to learning. The design was set up to provoke congruent gestures in the hands condition and incongruent gestures in the scoops condition. However comparing Hands and Scoops by test scores gives no insight into what students actually did. To discover whether Hands made congruent gestures and Scoops incongruent gestures, video recordings have been coded. 4 Types of gestures were distinguished: undo, scoop, shove, and form. Only shove was defined as congruent and it was expected that those in the hand condition would use this.

The results of the video analyses are shown in (Figure 9a). Counter to what was expected, Hands spent their time mainly forming, followed by scooping and least the congruent gesture of shoving. Scoops mainly scooped, followed by shoving and least forming. The amount of time spent on undoing is nigligible in both conditions. These results suggest that students in the hands and scoops condition adopt a different strategy for manipulating the sand.

Because school settings differed (traditional versus non traditional school, VMBO versus VWO), results of the video analysis were split out by gestures made at school A and school B. These results are shown in (Figure 9ab).

At school A, Hands spent most of the time forming (41%), followed by scooping (35%) and shoving (24%). Scoops used a different strategy, they spent the most time scooping (47%), followed by forming (31%) and shoving (22%). Neither Hands nor Scoops undid their actions. Hands made more congruent gestures than Scoops, but the difference is very small; only 2%.

Hands at school B prefered forming (42%). However, the time spent on shoving (39%) almost equals the time Hands spent on forming, this is at the expense of time spent on scooping (17%). Scoops at school B, suprisingly, spent most time shoving (53%), followed by scooping (25%) and forming (19%). Congruent gestures are made more often by Scoops than Hands. Moreover, for Scoops this is also the preferred gesture to use.

Based on the different gestures made at each school it becomes clear that students at school B made more congruent gestures than students at school A. Thus this is yet another difference in school setting, next to the type of school, and level and year of education of the students.

Notably, when looking at similarities, it becomes clear that all Hands (at school A and B) do have one thing in common. All Hands spent most of the time forming. Not only did they spent the most time on this gesture, they also always spent more time on it than Scoops. Therefore, this gesture rather than the congruent gesture of shoving might be related to learning.

## Discussion

This research was conducted to discover whether congruent gestures are related to learning, when students learn about erosion facilitated by an Augmented Reality Sandbox. This question was asked, because congruency and new technologies can play a role in effectively grounding concepts; thus in effectively embodying concepts.

To answer this question a quasi-experimental design was set up with two conditions: hands and scoops. Whereas Hands were able to perform assignments about erosion in an Augmented Reality with their bare hands, Scoops had to manipulate sand with a sand scoop. The reason for giving Scoops a sand scoop was to provoke incongruent gestures. It was assumed that students with a sand scoop would scoop sand, which is not analogically related to the meaning of erosion. Hands were thought to make shoving motions, pushing the sand away as they could freely use their hands. Hands and Scoops made the same assignments about valleys and meandering rivers. A pretest was given to students before working through the tasks, while a posttest was given upon finishing the tasks. The scores of the pre- and posttest were analyzed as well as video recordings.

The results showed that there was no significant difference in learning gain of students in the hands and scoops condition, however effect sizes of the learning gain are in favor of Hands. Suprisingly, it were not Hands that spent most time making the congruent gesture of shoving, but Scoops. A deeper analysis of the research conditions and what they provoked, in terms of gestures needs to be performed, in order to understand the relation between congruent gestures and learning about erosion.

A post-hoc power analysis revealed the low power of this study (0.29), which means that even if congruency is related to learning, the chances of finding it through this or a similar study is very low. Therefore, these results should be taken as explorative results.

A first pitfall of this study is that the learning gain cannot give any indication of why the posttest scores are significantly higher, because there was no condition that did not do the assignments provoking gestures. Therefore, making the pretest, or talking with peers about erosion might have been enough to find higher posttest scores, rather than the use of gestures. Also all students did already possess some pre-knowledge on erosion, so perhaps they did not learn anything, but only remembered more about erosion between the time of the pretest and posttest. Secondly, the learning gain measured at school A is most likely higher than in reality, because at school A some questions in the posttest were also in the pretest and thus a testing effect may have occured. It is not expected that this led to a difference in learning gain (and effect size) between Hands and Scoops because the same tests have been given to all Hands and Scoops at school A, therefore all Hands and Scoops could equally have "profited" from the repeated questions. It only matters in comparising test scores of (students of) school A with test scores of (students of) school B. Thirdly, the school setting of school A and B differed. It were different types of school (non traditional versus traditional), and the level and year of education of students differed. So it is not suprising that there was a significant difference in the scores at school A and B. Moreover, students at school B scored better in the tests. Most likely this is because they are enrolled in a higher level and year of education. Thus, they could have understood the assignment better, and/or learn faster. Fourthly, the sample size: not only was it very low, it also mostly included males, which means the results might not be representative for females.

It is interesting to see that, though not statistically significant, Hands appear to score higher on learning gain than Scoops, and appear to have some positive effect, while making the least congruent gestures. This is also the hardest part to explain. All Hands spent more time forming than Scoops. This is not completely surprising, as they were able to touch and manipulate sand with their hands and forming is easier done with hands than a rigid, harsh sand scoop. Moreover, Hands *touched* the sand. By forming they also used a big surface area of their hands to touch the sand, as they spread out their hands to manipulate the sand. Therefore, *touching* might be the factor rather than *congruency* that could explain the results, that seem to be favoring the hands condition. This means that for Scoops any positive relation that might exist between congruency and learning, might be lost because of the negative impact of not being able to touch the sand. Of course, assuming this as a factor is mere speculation but it shows the complexity of the relations under investigation.

In order to be able to provide designers and educators with better guidelines to embody concepts, future research should focus on why and when congruent gestures are made. Perhaps a top-down study could shed more light on the process. Rather than trying to provoke congruent gestures, search for strategies quick learners employ.

## References


Abrahamson, D., & Lindgren, R. (2014). Embodiment and embodied design. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 358–376). Cambridge, UK: Cambridge University Press.

Ashcraft, M. (2006). *Cognition* (4th ed.). Upper Saddles River, NJ: Prentice Hall.

Atit, K., Weisberg, S. M., Newcombe, N. S., & Shipley, T. F. (2016). Learning to interpret topographic maps: Understanding layered spatial information. *Cognitive Research: Principles and Implications*, *1*(2). doi: 10.1186/s41235-016-0002-y

Barron, B., & Darling-Hammond, L. (2008). Book excerpt teaching for meaningful learning: A review of research on inquiry-based and cooperative learning. In R. Furger (Ed.), *Teaching for meaningful learning: A review of research on inquiry-based and cooperative learning.* Retrieved from https://files.eric.ed.gov/fulltext/ED539399.pdf (Reprinted from *Powerful learning: What we know about teaching for understanding*, by Darling-Hammond, L., Barron, B., Pearson, P.D., Schoenfeld, A.H., Stage E.K., Zimmerman, T.D., Cervetti, G.N., Tilson, J., 2008, San Fransisco, CA: Jossey-Bass)

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 637-660.

Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science*, *2*(4), 716-724.

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, *11*(4), 209-243. doi: 10.1177/1059712303114001

College voor Toetsen en Examens. (2015). *Syllabus centraal examen 2017.* Retrieved from https://www.examenblad.nl/

Dijkstra, K., Kaschak, M. P., & Zwaan, R. A. (2007). Body posture facilitates retrieval of autobiographical memories. *Cognition*, *102*(1), 139-149.

Dove, J. (1997). Student ideas about weathering and erosion. *International Journal of Science Education*, *19*(8), 971-980. doi: 10.1080/0950069970190809

European Commision, High Level Group on Science Education. (2007). *Science education now: A renewed pedagogy for the future of europe.* Retrieved from http://ec.europa.eu/research/science-society/

Gagnier, K., Shipley, T., Tikoff, B., Ormand, C., Atit, K., Resnick, I., & Garnier, B. (2016). Training spatial skills in geosciences: A review of tests and tools. In B. Krantz, C. Ormand, & B. Freeman (Eds.), *AAPG Memoirs: Vol. 111. 3-D Structural interpretation: Earth, mind, and machine* (pp. 7–23). doi: 10.1306/M1111307

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1* (4), 586-596.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42* (1-3), 335-346.

Hume, D. (1777). An enquiry concerning human understanding. In P. Millican (Ed.), *Essays and Treatises on Several Subjects: Vol. 2.* Retrieved from http://www.davidhume.org/texts/e

Jacovina, M., Ormand, C., Shipley, T. F., & Weisberg, S. (2014). Topographic map assesment [Test]. Retrieved from http://www.spatiallearning.org/index.php/testainstruments

Jang, S., Vitale, J. M., Jyung, R. W., & Black, J. B. (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Computers & Education*, *106* , 150-165.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, *106* (1), 86-104.

Kant, I. (2004). *Kritiek van de zuivere rede* [Critique of Pure Reason] (J. Veenbaas & W. Visser, Trans.). Amsterdam, The Netherlands: Boom.

Kastens, K. A., Agrawal, S., & Liben, L. S. (2008). Research methodologies in science education: The role of gestures in geoscience teaching and learning. *Journal of Geoscience Education*, *56* (4), 362-368.

King. (2008). Geoscience education: an overview. *Studies in Science Education*, *44*(2), 187-222. doi: 10.1080/03057260802264289

Koffijberg, I. J. P. (2017, October). De AR sandbox als tool voor onderzoek naar embodied cognition [The AR sandbox as tool for research on embodied cognition]. *geografie.nl*. Retrieved from https://geografie.nl

Lakens, D. (2016, January 14). The 20% Statistician [Web log message]. Retrieved from http://daniellakens.blogspot.com/2016/01/power-analysis-for-default-bayesian-t.html

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher*, *42*(8), 445-452.

Marshall, P. J. (2016). Embodiment and human development. *Child Development Perspectives*, *10*(4), 245-250.

Millar, R. (1991). Why is science hard to learn? *Journal of Computer Assisted Learning*, *7*(2), 66-74.

OECD. (2016). *PISA 2015 Results* (Vol. 1: Excellence and Equity in Education). Paris, France: OECD Publishing. doi: 10.1787/9789264266490-en

Osborne, J., & Collins, S. (2000). Pupil's and parent's views of the school science curriculum. *School Science Review*, *82*(298), 23-32.

Osborne, J., & Hennessy, S. (2003). *Literature review in science education and the role of ICT: Promise, problems and future directions* (Research Report No. 6). Retrieved from NESTA Futurelab website: https://telearn.archives-ouvertes.fr/hal-00190441.

Segal, A., Tversky, B., & Black, J. (2014). Conceptually congruent actions can promote thought. *Journal of Applied Research in Memory and Cognition*, *3*(3), 124-130.

Shapiro, L. (2011). Standard cognitive science. In *Embodied cognition* (pp. 7–27). New York, NY: Routledge.

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the p value is not enough. *Journal of Graduate Medical Education*, *4*(3), 279–282.

Thomas, G., & Durant, J. (1987). Why should we promote the public understanding of science. *Scientific Literacy Papers*, *1*, 1-14.

UC Davis W.M. Keck Center for Active Visualization in the Earth Sciences. (n.d.). SARndbox (Version 2.2) [Computer Software]. Retrieved from http://arsandbox.ucdavis.edu/technical-resources/downloads

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science*, *6*(3), 291–298.

Woods, T. L., Reed, S., Hsi, S., Woods, J. A., & Woods, M. R. (2016). Pilot study using the augmented reality sandbox to teach topographic maps and surficial processes in introductory geology labs. *Journal of Geoscience Education*, *64*(3), 199-214.

Woolfolk, A., Hughes, M., & Walkup, V. (2013). Cognitive views of learning. In *Psychology in education* (2nd ed., pp. 228–329). Harlow, England: Pearson Education Limited.

Table 1

*A description of the four codes used to analyze the gestures of students*

| Code | Description |
|---|---|
| Undo | Students would undo their previous manipulation the sand |
| Scoop | Students were either scooping up sand with the sand scoop, or their hands. Either way, the sand would not be in contact anymore with the sand in the sandbox |
| Shove | Students would push or shove the sand in any direction either by means of the sand scoop, or by use of their hands |
| Form | Students were shaping a structure in the sand with either the sand scoop or their hands |

Table 2

*Scores for the pre- and posttest and the learning gain for school A and B*

| | $M_{pre}$ ($SD_{pre}$) | $M_{post}$ ($SD_{post}$) | $M_{\text{learning gain}}$ ($SD_{\text{learning gain}}$) | $p$ | ES ($d$) |
|---|---|---|---|---|---|
| School A + B[a] | | | | | |
| Hands | .33 (.19) | .42 (.25) | .09 (.18) | .16 | 0.36 |
| Scoops | .37 (.16) | .39 (.22) | .02 (.22) | | |
| School A[b] | | | | | |
| Hands | .28 (.17) | .36 (.22) | .079 (.19) | .21 | 0.36 |
| Scoops | .32 (.11) | .33 (.18) | .004 (.22) | | |
| School B[c] | | | | | |
| Hands | .53 (.16) | .70 (.20) | .17 (.11) | .55 | 0.36 |
| Scoops | .56 (.20) | .67 (.14) | .10 (.25) | | |

*Note.* [a]$N = 62$. [b]$N = 50$. [c]$N = 12$.

Table 3

*Results for the one-way repeated measures ANOVA on time (pre- and posttest)*

| | $df_{error}$ | $df$ | $F$ | $p$ | ES ($\eta^2$) | ES ($d$)[a] |
|---|---|---|---|---|---|---|
| Within factors | | | | | | |
| Time | 58 | 1 | 7.40 | .009* | 0.11 | 0.70 |
| Time*Condition | 58 | 1 | 1.21 | .28 | 0.02 | 0.29 |
| Time*School | 58 | 1 | 2.14 | .15 | 0.04 | 0.41 |
| Time*Condition*School | 58 | 1 | 0.00 | .97 | 0.00 | 0.00 |
| Between factors | | | | | | |
| School | 58 | 1 | 41.0 | <.001* | 0.41 | 1.67 |
| Condition | 58 | 1 | 0.02 | .90 | 0.00 | 0.00 |
| Condition*School | 58 | 1 | 0.01 | .94 | 0.00 | 0.00 |

*Note.* [a]Transformation of $\eta^2$ was computed as $d = 2 * \text{f}$, where $f^2 = \eta^2 / (1 - \eta^2)$.

*Significant at $p < .05$.

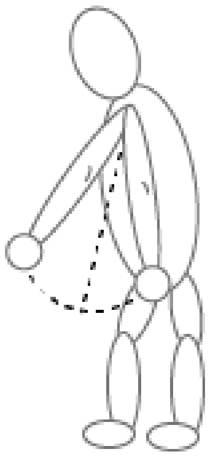


*Figure 1*. Congruent motion of pendulum

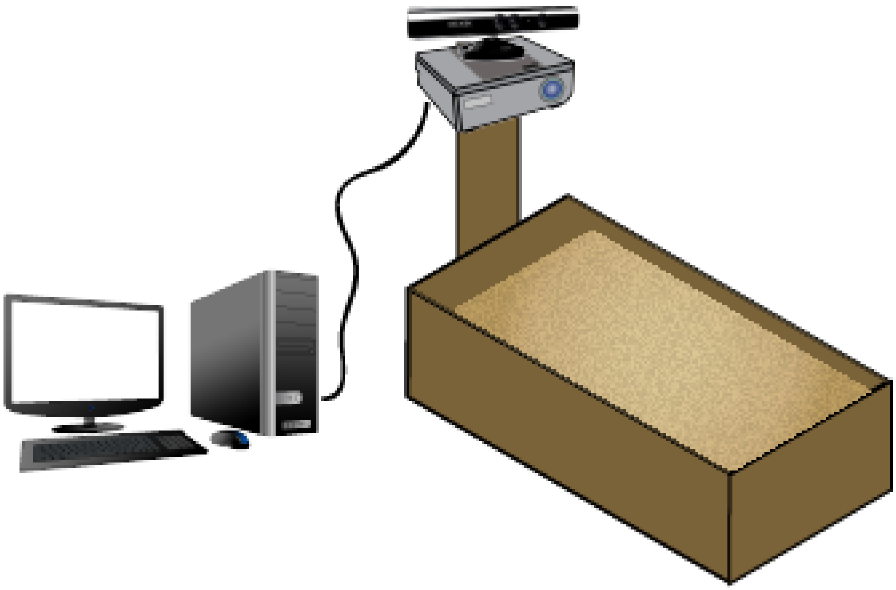


*Figure 2*. A schematic overview of the Augmented Reality Sandbox

*Figure 3*. The sand scoop and glove given to Scoops

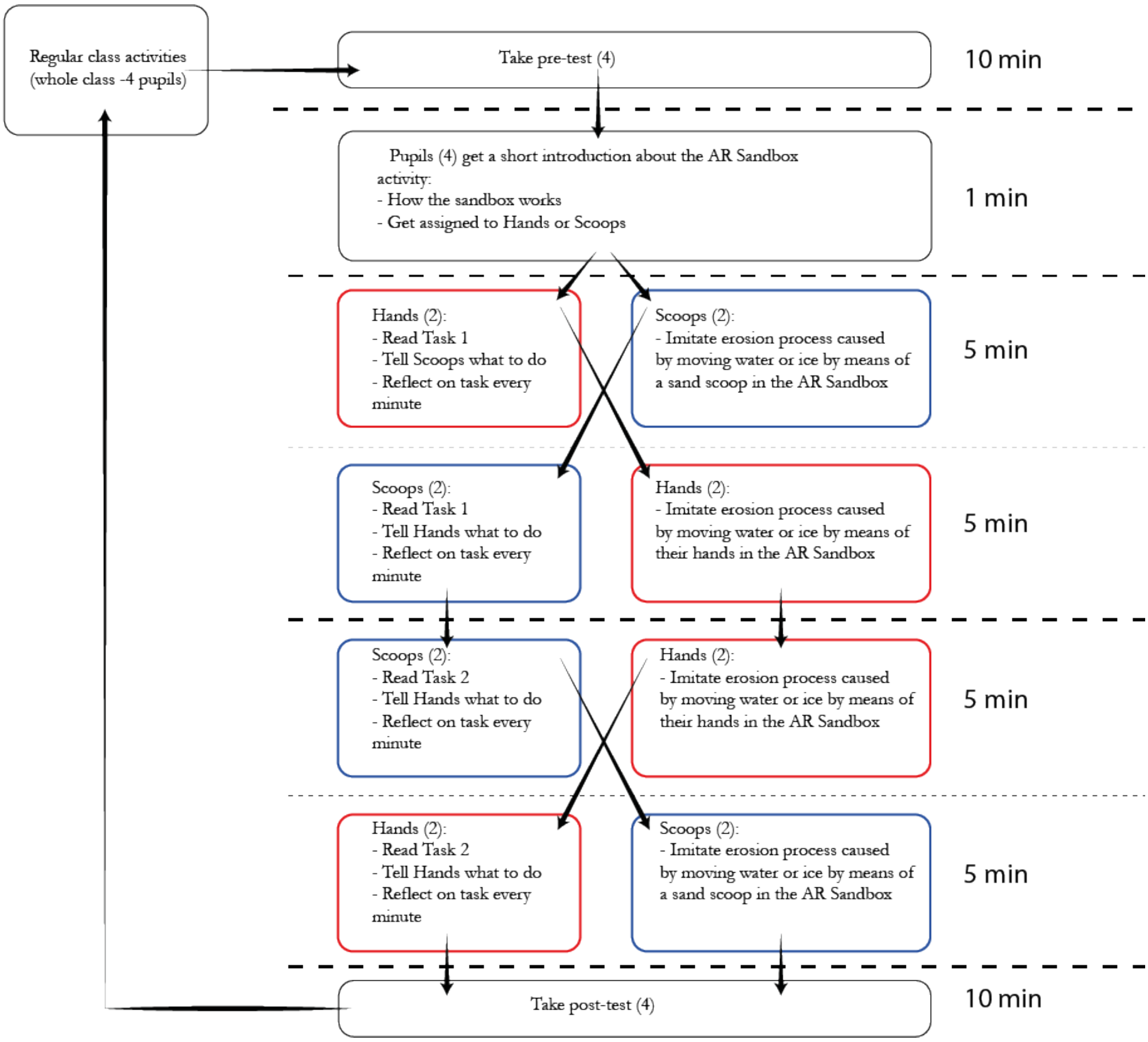
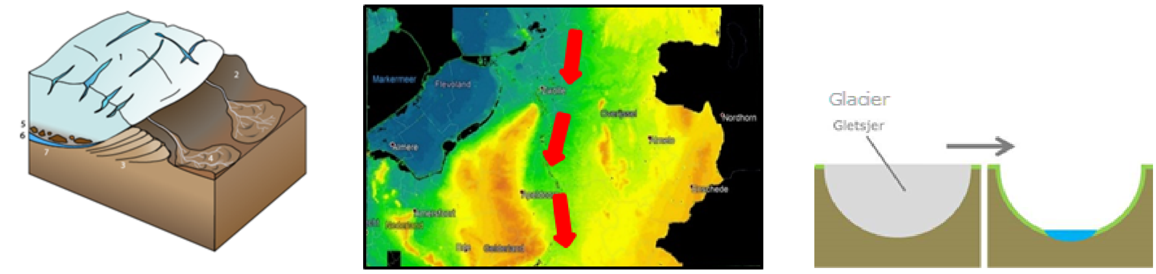
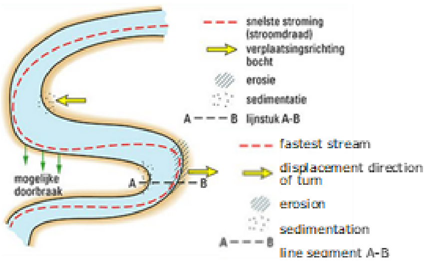



*Figure 4*. Schematic overview of the set up of the intervention



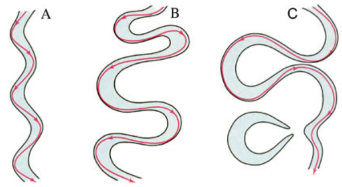

(a) Given at school A (b) Given at school A (c) Given at school B

*Figure 5*. Pictures in the IJssel valley task given to students at school A and B



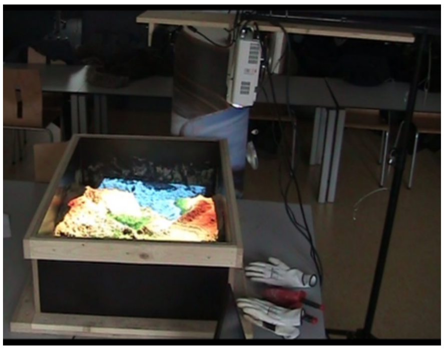

(a) Given at school A

(b) Given at school A

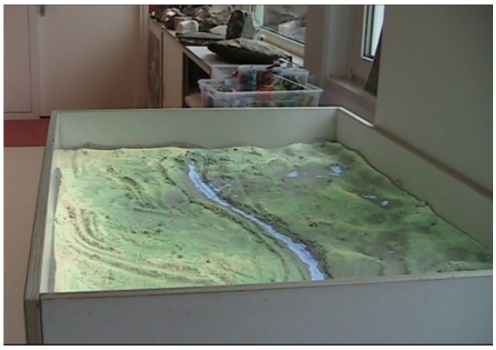



(c) Given at school B

*Figure 6*. Pictures in the meandering river task given to students at school A and B



(a) Used at school A



(b) Used at school B

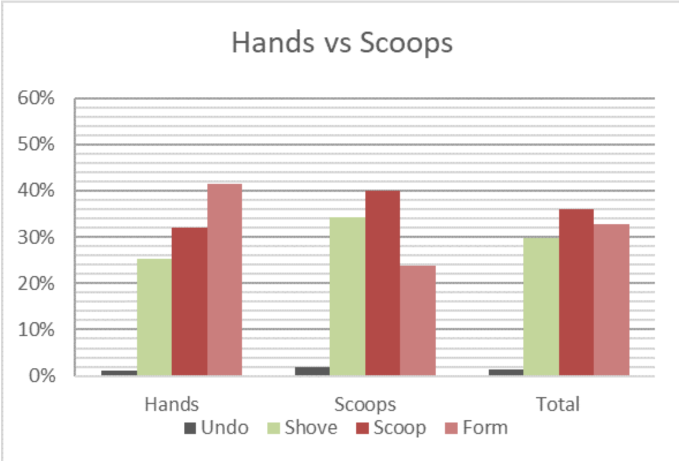
*Figure 7*. Augmented Reality Sandboxes used at school A and B

(a) A student shoving

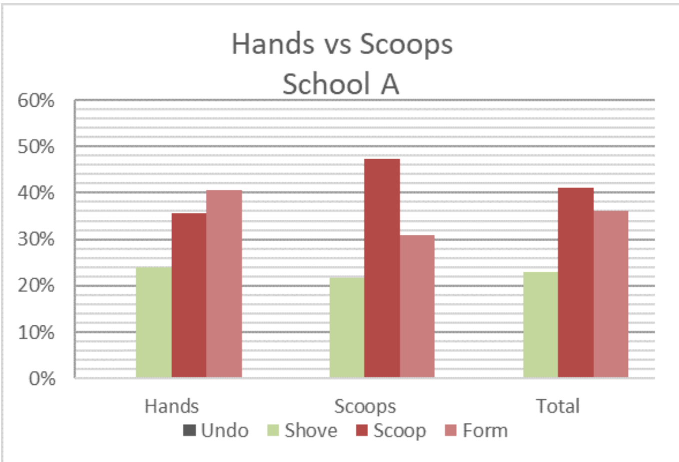
(b) A student scooping

(c) A student forming

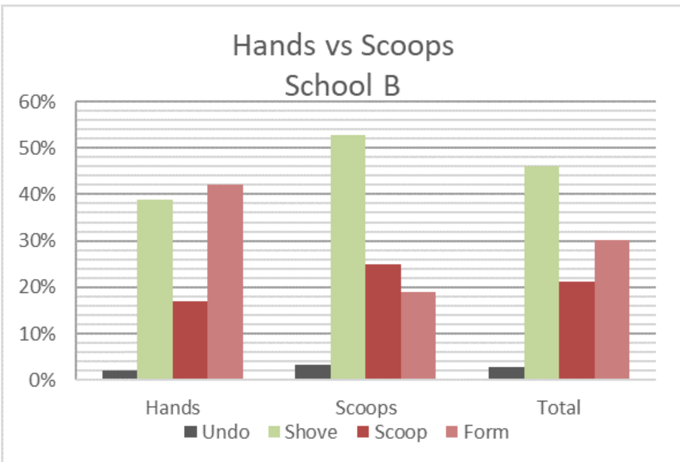
*Figure 8*. Visual representation of the gestures shoving, scooping and forming




(a) All Hands and Scoops




(b) Hands and Scoops at school A




(c) Hands and Scoops at school B

*Figure 9*. Relative time spent on each type of gesture by Hands and Scoops in total and at school A and B respectively

# Appendix A

## Questions in the Pre- and Posttest

Table A1

*Questions used for the pre- and posttest*

| School A | | | | School B | | | |
|---|---|---|---|---|---|---|---|
| Pretest | | Posttest | | Pretest | | Posttest | |
| AAAS | TMA | AAAS | TMA | AAAS | TMA | AAAS | TMA |
| WE015003 | 2 | WE015003[a] | 5 | WE015003 | 2 | WE011002[a] | 5 |
| WE018003 | 9 | WE018003[a] | 8 | WE018003 | 9 | WE027002[a] | 8 |
| WE026004 | 10 | WE039002 | 11 | WE026004 | 10 | WE039002 | 11 |
| WE028003 | 14 | WE042002 | 12 | WE028003 | 14 | WE042002 | 12 |
| WE056001 | 15 | WE048001 | 14[a] | WE056001 | 15 | WE048001 | 13[a] |
| | 16 | | 18 | | 16 | | 18 |

*Note.* The codes match the questions used from AAAS and TMA.

[a]These questions are different at school A and B, however the questions are similar.

# Appendix B

## Simulator for Random Matching

```
#include <stdio.h>
#include <math.h>
#include <stdlib.h>
#include <time.h>

void swap(float *a, float *b) {
        float temp;
        temp = *a;
        *a = *b;
        *b = temp;
}

int main(int argc, const char * argv[]) {
        srand(time(NULL));
        int a;
        int i;
        int N;
        float z[200];
        float d[200];

        puts("Enter n for each condition");
        scanf("%d", &a);

        for (i = 0; i<a; ++i)
        {
                printf("Enter n of experimental condition %d: ", i + 1);
                scanf("%f", &z[i]);
        }

```

```
    for (i = 0; i<a; ++i)
    {
        printf("Enter n of control condition %d: ", i + 1);
        scanf("%f", &d[i]);
    }

    puts("Enter how many simulations you want to run");
    scanf("%d", &N);

    //N step
    FILE *data = fopen("data.dat", "w");
    int step;
    double sum = 0.0;
    for (step = 0; step< N; ++step) {

        for (i = 0; i<a; ++i) {
            int j;
            j = rand() % a;
            // swap two position
            swap(&z[i], &z[j]);
        }

        for (i = 0; i<a; i++)
        {
            printf("%f, ", z[i]);
        }

        // comparision
        int large = 0;
        for (i = 0; i<a; ++i) {
            if (z[i] > d[i]) {
```

```
                    large++;
                }
            }
            sum += large;
            printf("\n%d\n", large);
            fprintf(data, "%d\n", large);
        }


        printf("average \n%d\n", sum / N);
        fprintf(data, "average %f\n", sum / (double)N);
        fclose(data);

        return 0;

}
```