# The Effect of Dynamic Algebra Animation on Learning Outcomes in a Physics Education Video

Rick Oosterwijk

Freudenthal Institute, Utrecht University

Master Thesis (45 ECTS)

Science Education & Communication MSc

Dr. Rogier Bos

Second Reviewer: Prof. Dr. Paul Drijvers

December 12, 2021

Wordcount: 9641

# Abstract

Research remains inconclusive as to what extent dynamic animations contribute to an effective educational video, especially in the context of mathematics and physics. This study investigated whether and why dynamic animation of algebra is beneficial for grasping topics that hold both an algebraic and a conceptual challenge. The question was whether video-design based on conceptual metaphor theory and embodied cognition can improve understanding of both algebraic manipulations and physical concepts for such topics. In a quantitative experiment 38 bachelor students in chemistry or molecular life sciences watched a designed video about time dilation in which the involved algebra was animated in either a static way, a dynamic way, or with animated hands dynamically performing the manipulations. We investigated cognitive load as a possible explanatory factor for the effects of these different conditions. Analyses of the post-test scores showed significantly higher understanding of the physical concepts in the embodied dynamic condition compared to the dynamic condition. For algebra scores as well as cognitive load no significant differences between the conditions were found.

*Keywords:* dynamic animation, embodied cognition, physics education, algebra, conceptual metaphor theory, cognitive load, time dilation.

# Introduction

Audio-visual media such as video are used on an unprecedented scale in education worldwide. More and more, teachers and students turn to online platforms such as *MOOCs* and *YouTube Edu*, as well as domain-specific channels like *Minute Physics* or *3Blue1Brown* for mathematics, to help achieve various kinds of learning goals. But despite their increasingly pertinent role in education, these videos are anything but a solved puzzle. What makes a good educational video? A hundred years after making their debut in the classroom setting, this fundamental debate on video-usage is still nowhere near concluded (Wijnker, Bakker, Van Gog, & Drijvers, 2019).

Over the years, the debate caught the attention of many researchers within the fields of educational sciences and cognitive psychology. Their resulting studies have mainly focused on the aim of maximizing learning gains from educational videos. For this purpose, these studies developed design strategies for factors like efficient processing of audio-visual information and learning in online environments (Wijnker et al., 2019). Methods such as presenting corresponding words and images simultaneously as well as in proximity are examples of such strategies that were found to improve learning (Mayer, 2020; Muller, 2008).

However, past research regarding efficient processing of audio-visual information is inconclusive or insufficient for several video characteristics. This is for example the case when it comes to the effectiveness of static versus dynamic animations, and to the effectiveness of using a body analogy (embodied simulation) in educational videos. For one, research regarding these topics has not yet been conducted elaborately within the full scope of educational fields. In mathematics for example, such research has thus far only been conducted on low-level problems, thereby overlooking vital, more complex topics like algebra. Furthermore, existing research results on the effectiveness of static versus dynamic animations are very mixed. Both conditions are endorsed for instruction over the other by several independent studies (Berney & Bétrancourt, 2016; Castro-Alonso, Ayres, Wong, & Paas, 2018; Höffler & Leutner, 2007). On top of this, for these video characteristics the underlying reasons behind their (in)effectiveness also remain unclear. To further improve learning gains from educational videos, it is important to supplement the existing database of knowledge on these characteristics.

This study investigated the effects of static versus dynamic and embodied dynamic animation of algebra on learning outcomes for topics that are both algebraically as well as conceptually challenging. The aim was to determine whether and why in an educational video these three conditions influence student understanding of the animated algebraic manipulations, as well as the physical concepts involved. For this purpose, a quantitative experiment was performed using designed animation videos about the concept of time dilation. The extensive debate on static versus dynamic animations is important for general video education on all science topics. With their popularity on the rise, the hope is that by shedding more light on this debate, we were able to contribute to the quality of educational videos within many different science fields.

## Theoretical background

In this section we review educational psychology literature on static versus dynamic animation, embodied cognition, and cognitive load. We then review the conceptual metaphor theory in relation to embodied mathematics as well as literature on special relativity theory education. Finally, we discuss the implications of this literature for the present study.

## Educational psychology literature on static versus dynamic animation

In this study we use the definition of the term animation by Bétrancourt and Tversky (2000), defining it as “any application which generates a series of frames, so that each frame appears as an alteration of the previous one, and where the sequence of frames is determined either by the designer or the user” (p. 313). A static animation can then be understood as a depiction of such a series of frames as a traditional slide show, while a dynamic animation makes use of continuous movement and deformation.

Research shows contradictory results concerning the effectiveness of static versus dynamic animations. On the one hand, a meta-analysis by Höffler and Leutner (2007) shows that dynamic animations are more effective in acquiring motor-procedural knowledge and skills, compared to for example conceptual knowledge. On the other hand, Berney and Bétrancourt (2016) conclude in their meta-analysis it is precisely in the acquisition of conceptual and factual knowledge that dynamic animations are most effective. On top of this, although both meta-analyses find dynamic animations to be favourable over static ones, the results show a lot of inconsistencies. In 140 pair-wise comparisons, Berney and Bétrancourt found that in only 30.7% of studies, dynamic animations led to significantly better learning outcomes than static animations. In 59.3%, they found no significant improvement or deterioration.

These mixed results can be explained as there are both advantages and disadvantages to dynamic animations. For one, dynamics can help to facilitate cueing, meaning student attention can be directed as desired. Secondly, the additional temporal information provided in dynamic visualisations allows for seeing actual motion and changes in a continuous process instead of having to infer them, which can decrease cognitive load. A disadvantage of dynamic information is however that due to its transient nature, it can also add cognitive load. With shown elements continually disappearing again from view, it becomes more challenging to remember and integrate all seen information (Ayres & Paas, 2007). Berney and Bétrancourt further explain their inconsistent results by stating that in their analysed studies, many factors and variables that impact the effectiveness of dynamic visualisations have not been appropriately controlled for. For instance, the videos used in these studies differ in the presence of interactive features and accompanying text, both of which have been shown to influence the effectiveness of dynamic animation videos (Berney & Bétrancourt, 2016; Castro-Alonso, Ayres, & Paas, 2016; Höffler & Leutner, 2007). In the current study, this analysis was used to control for such confounding factors in the design of dynamically animated algebra.

## Educational psychology literature on embodied cognition

To counteract the negative effect of transiency in dynamic animations, researchers point to the human movement effect (Castro-Alonso, Ayres, & Paas 2014a; Sweller, Van Merriënboer, & Paas, 2019). This effect implies that in teaching cognitive tasks involving human movement, dynamic animations are more effective than static ones. An explanation for this is argued to be the theory of grounded or embodied cognition. According to this theory, cognitive processes are linked with our bodily or environmental experiences (Garbarini & Adenzato, 2004). Consider for example learning to ride a bike. No one explains exactly how to keep balance and move forward. Instead, you learn these skills in a trial-and-error process through your bodily experiences, thanks to embodied cognition. The theory implies that when watching a dynamic visualisation, the viewer also links all its contents to such experiences (Castro-Alonso et al., 2014a). In the present study, where algebraic manipulations are visualised dynamically, research on embodied cognition showed that a stronger grounding of dynamic algebra in bodily experiences is associated with a stronger human movement effect, leading to more effective dynamic animations (Castro-Alonso et al., 2014a).

The theory of embodied cognition in educational psychology is supported by widespread findings. For starters, research by Sweller et al. (2019) indicates that both observing and making gestures lead to richer encoding of information. Another study by De Koning and Tabbers (2013) shows that the learning effects from an animated video explaining the formation of lightning were greater when the movements in the video were followed by a human hand. However, contradicting results come from research studying the retention of abstract symbols in a small grid. In these studies, observing static hands is found to hinder learning in dynamic videos, while showing conflicting results in the case of static videos (Castro-Alonso, Ayres, & Paas, 2014b; Castro-Alonso et al., 2018). To explain the contradicting results, researchers suggest these static hands might have been experienced as unnecessary, redundant visual information, which can increase cognitive load (Sweller et al., 2019). In past literature in support of embodied cognition theory, moving hands were observed. In the current study, we investigated whether visualising moving hands in an algebraic derivation can be beneficial for mastering both the algebraic and conceptual challenges involved in learning about the topic of time dilation.

## Cognitive load literature

Literature shows that learning is impeded in learning tasks in which the induced cognitive load is too high (Martin, 2014; Sweller et al., 2019). This cognitive load refers to the amount of working memory resources employed in performing a learning task, and is influenced by both the complexity of the task as well as the way in which information is presented. By minimalizing the contribution of this latter factor, learning outcomes are improved (Sweller, 1988; Sweller et al., 2019). In our current study, information is being presented in three different ways over three video conditions. As such, we can check whether for example having a moving hand dynamically move algebraic formulas around in the embodied dynamic condition decreases students' experienced cognitive load, and how changes in this experience correlate to measured learning outcomes. Cognitive load theory however also faces criticism. For example, Martin (2014) describes that a higher cognitive load does not always correspond to a lower task performance. A reason for this is that due to the multidimensional nature of mental performance, the construct is also influenced by factors such as effort, commitment to task and motivation. Furthermore, research has shown learning to also deteriorate in cases where cognitive load is too low (Martin, 2014). As such, in applying cognitive load theory, it is important to both find the optimum amount of load for learning and take individual differences into consideration.

## Conceptual metaphor theory literature

Education based on ideas of embodiment provides a way to shape meaningful education. In mathematics, Wittmann, Flood, and Black (2013) show that students use gestures and speech to represent rearranging terms in mathematical equations as physical objects moving in a landscape. In such embodied mathematics, an important mechanism proposed to explain how abstract mathematical concepts are connected to the world and to each other is the conceptual metaphor mechanism (Lakoff & Núnez, 2000). To explain how this mechanism works, Lakoff and Núnez (2000) make a distinction between 'grounding metaphors' and 'linking metaphors'. In grounding metaphors, fundamental mathematical concepts are directly coupled to physical experiences. For example, through the 'arithmetic is object collection' grounding metaphor, fundamental arithmetical operations like addition and subtraction are understood in terms of for example adding things to a pile or taking things away from a pile. In literature, this is described as a metaphorical ‘mapping’ from the ‘source domain’ of object collection to the ‘target domain’ of arithmetical operations. On the other hand, in linking metaphors, abstract parts of mathematics are understood in terms of each other, linking for example numbers to points. Thus, in contrast to grounding metaphors, linking metaphors are only grounded in physical experiences in an indirect way, resulting in a weaker embodied experience. The idea is that in the end, all of mathematics can be understood as a series of such metaphorical mappings, that are grounded in physical experiences either directly through grounding metaphors, or indirectly through linking metaphors (Winter & Yoshimi, 2020).

Other research explains how conceptual metaphors can be used in practice to achieve a meaningful visual ‘language’ for animating algebraic operations (Bos & Renkema, forthcoming 2022). These researchers showed that dynamic or embodied dynamic algebra animation using this visual language has no significant effect on learning outcomes compared to static animation when the learning goal is to achieve algebraic manipulative skills in new learners. In the present study conceptual metaphors are applied in the same way. Now however, algebraic operations are placed in a physical context, i.e. time dilation, and serve only to support the actual learning goals of understanding both the algebraic derivation as well as the physical concepts and reasoning involved in this time dilation context. As such, apart from the mathematical part, the videos also had to effectively illuminate time dilation from a conceptual point of view.

## Special relativity theory education literature

A review on special relativity theory (SRT) education by Alstein, Krijtenburg-Lewerissa, and Van Joolingen (2020) indicated that the main SRT learning difficulties for students at all educational levels are interpreting frames of reference, relativistic effects, and the SRT postulates. The review showed that these learning difficulties are best dealt with by taking a conceptual approach to teaching SRT centred around the use of thought experiments, as these allow students to place underlying principles, assumptions, and mechanisms in the context of everyday experiences. These results provided further design guidelines for the videos in the current study.

## This study

### *Research questions*

The present study compared the effect of static, dynamic, and embodied dynamic animation of algebra on learning outcomes in designed time dilation videos. The aim was to determine whether and why these three conditions affect student understanding concerning both the algebraic manipulations, and the physical concepts and reasoning involved. To fulfil this aim, three research questions needed to be addressed:

1. *What are the effects of dynamic and embodied dynamic visualisation of algebra on understanding algebraic manipulations and physical concepts and reasoning involved in animation videos about the concept of time dilation, compared to static visualisation?*
2. *How does understanding the algebraic manipulations in the videos relate to understanding their physical concepts and reasoning?*
3. *To what extent can these effects be explained in terms of experienced cognitive load?*

Additionally, regarding this measured student understanding of the videos, we were interested whether there was a ‘wow-effect’ present, meaning we explored the idea that students who consider learning material to be attractive might think they understand the material better than they actually do. We grew interested in this idea through literature done on seductive details in multimedia learning, which refers to the presence of interesting yet redundant pieces of information. A meta-analysis by Sundararajan and Adesope (2020) showed that such details, which might add to the attractiveness of the learning material in which they are embedded, have a detrimental effect on student learning. This got us wondering how the general attractiveness of our designed time dilation videos might be related to student learning outcomes. Somewhat in contrast with the described detrimental effect of seductive details on learning, we intuitively think higher attractiveness of learning material might relate to higher student self-assessment of their understanding of the material. Due to this contrast, we became interested in this self-assessment as well, leading to our fourth research question:

4. *What is the relationship between how attractive the videos are rated, how well students feel they understood the video, and how well they truly understood it?*

### *Hypotheses*

Concerning the effect of the conditions on learning outcome, we saw earlier that past meta-analyses found dynamic animations to be favourable over static ones, as well as that embodied simulation using moving hands enhanced learning effects. Specifically looking at the algebra, the manipulations used in the explanation of time dilation were visualised line by line in the three conditions. Because the old form of the formula most often remained visible on screen, the negative influence of transiency in dynamic animations was weakened. Furthermore, research by De Koning, Tabbers, Rikers, and Paas (2007) shows that strongly drawing attention to one part of an animation through cueing improves understanding and transfer of both the cued information as well as the uncued information in the animation. As earlier described research showed dynamic animation to allow for attention cueing, and attention is most strongly directed in the embodied dynamic condition through the salient moving hands, the first hypothesis corresponding to the first research question was as follows:

1. Embodied dynamic visualisation of algebra is most effective regarding understanding of algebraic manipulations involved in time dilation, followed by dynamic visualisation.

Regarding understanding the physical concepts, the research by De Koning et al. (2007) suggests that understanding of this uncued part is influenced by the conditions in the same way as understanding of the cued algebra video part. Additionally, earlier described research on cognitive load suggests that reducing the cognitive load experienced while watching the algebraic derivation could leave more working memory capacity for conceptual understanding of time dilation. A meta-analysis by Xie, Wang, Hao, Chen, An, Wang, and Liu (2017) found that in multimedia learning, attention cueing reduces the reported cognitive load. Based on the extent to which attention is cued in the different conditions we reached the second, third and fourth hypotheses, respectively corresponding to the first, second, and third research question:

2. Embodied dynamic visualisation of algebra is most effective regarding understanding of physical concepts involved in time dilation, followed by dynamic visualisation.
3. Due to attention cueing, understanding both the algebraic manipulations as well as the physical concepts and reasoning are interwoven in the embodied dynamic condition and the dynamic condition.

4. Cognitive load is an explanatory factor concerning the effects of the conditions on learning outcomes, as embodied dynamic visualisation of algebra leads to the lowest reported cognitive load, followed by dynamic visualisation.

Finally, concerning the wow-effect, literature did not provide any studies that either supported or contradicted our assumption relating rated attractiveness of learning materials to self-assessed student understanding. We tested our assumption to try and add to the existing literature. Our hypothesis is as follows:

5. Students who give a high rating to the attractiveness of the video *think* they understand the content of the video better than they actually do.

## Methods and Materials

### Overview

We designed animated videos about time dilation to perform a quantitative experiment containing two experimental groups and one control group. In the control group, students watched a video in which the involved algebraic manipulations were animated in a static way. In the two experimental groups, these manipulations were animated in a dynamic way, or with animated hands dynamically performing the manipulations (embodied dynamic way).

The students participated in the experiment individually. In all conditions, the participants first completed a pre-test. This test provided a self-assessed and an objective measure of the students' starting level concerning skill in algebraic derivation as well as prior special relativity theory knowledge. The participants then watched their condition's respective video. Finally, the students completed a designed post-test that was identical for participants in all three conditions. This test measured several variables (cf. subsection instruments), but its focus was on measuring student understanding of the algebraic manipulations and physical concepts presented in the video. The entire experiment was conducted online and took about 30 minutes per student.

## Participants

Our target group consisted of students who are experienced with algebraic manipulations but have not mastered them to the extent that such manipulations no longer cost them effort. As such, 38 bachelor students Chemistry and Molecular Life Sciences participated in the experiment. The students applied for participation by voluntarily responding to a post about the study placed on the blackboard page of the course ‘SK-BWSNK1: Wis- en natuurkunde 1’ at Utrecht University, a course that is part of the study program of all students in our target group. We assigned the students to the three conditions based on the order in which they applied to join the study, i.e. the first applicant was placed in the embodied dynamic condition, the second in the dynamic condition, the third in static, the fourth again in embodied dynamic and so on. As a result, the participants per condition were: 13 embodied dynamic, 13 dynamic, 12 control (static).

## Instruments

### *Videos*

**Design and learning goals.** The three videos (see them here: https://tinyurl.com/algani) were identical in every way apart from the way algebraic derivations were presented. In particular, all videos were 7:21 minutes long and had identical audio, algebraic expressions, fonts and font size, colours, timing, outlay and design. We adhered to basic multimedia principles by

Mayer (2020) and Muller (2008) in the design, and tested a first and second version of the videos on test subjects as well as a mathematics educational research expert before reaching the final versions. The learning goals of the videos were as follows:

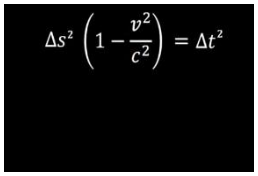
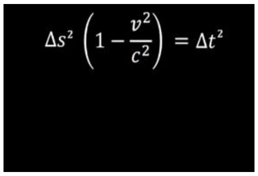
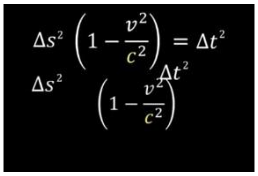
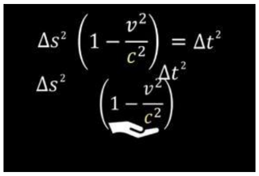
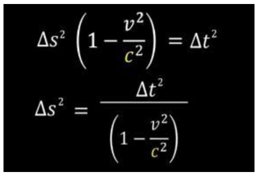
1. After watching the video, the student can reproduce and understand the algebraic derivation of time dilation from the postulates of special relativity theory.
2. After watching the video, the student can reproduce, understand, and to a limited extent apply the physical concepts and reasoning involved in time dilation, including the postulates of special relativity theory.

**Conditions.** In the three conditions, the algebraic manipulations used in the explanation of time dilation were visualised line by line. The conditions differed in the transitions between the different algebraic steps. For the dynamic and embodied dynamic cases, these transitions occurred dynamically, based on conceptual metaphor theory (Lakoff & Númez, 2000). In the latter case, the movement is executed by animated human hands. Table 1 shows the difference between the conditions, using the algebraic manipulation of division as an example. The top row shows the starting point of this manipulation in the videos, which is identical in all three conditions. The second row shows how the actual process of division is visualised in each of the three conditions. Finally, all three videos reach identical expressions again after the division.

**SRT education strategy.** To address the earlier described SRT learning difficulties found in literature, the storyline of the videos was centered around a thought experiment, i.e. cooking an egg on a ‘rocket at rest’ versus on a ‘moving rocket’. The backbone of the video was using the SRT postulates as the foundation based on which we derived the relativistic effect of time dilation in this thought experiment in a logical way. We avoided use of the term frame of reference throughout the video and applied this concept only in the context of Einstein’s frame-dependent time coordinates, i.e. ‘according to A, B is moving closer, while according to B, A is moving closer’. To emphasize, we realize the SRT material would be more effective if, for example, it was explained in several different ways in the video, or if students were allowed to watch the video multiple times. However, the purpose of the video is not to explain time dilation as effectively as possible. Instead, the study is aimed at measuring experimental effects between conditions.

**Table 1**

*Visualisation of Division in Video Conditions*

| Time | Visualisation of Division | | |
|---|---|---|---|
| 4:20 | | *Start all conditions*<br><br>$\Delta s^2\left(1-\frac{v^2}{c^2}\right)=\Delta t^2$ | |
| 4:23 | *Static*<br><br>$\Delta s^2\left(1-\frac{v^2}{c^2}\right)=\Delta t^2$ | *Dynamic*<br><br>$\Delta s^2\left(1-\frac{v^2}{c^2}\right)=\Delta t^2$<br>$\Delta s^2$ $\Delta t^2$<br>$\left(1-\frac{v^2}{c^2}\right)$ | *Embodied dynamic*<br><br>$\Delta s^2\left(1-\frac{v^2}{c^2}\right)=\Delta t^2$<br>$\Delta s^2$ $\Delta t^2$<br>$\left(1-\frac{v^2}{c^2}\right)$ |
| 4:26 | | *End all conditions*<br><br>$\Delta s^2\left(1-\frac{v^2}{c^2}\right)=\Delta t^2$<br>$\Delta s^2=\frac{\Delta t^2}{\left(1-\frac{v^2}{c^2}\right)}$ | |

**Application of conceptual metaphor theory.** Conceptual metaphors were applied similarly as in research by Bos and Renkema (forthcoming, 2022). In the current study however, both the level of applied algebra and the student starting level were higher. This means *arithmetic is object collection* is no longer the main conceptual metaphor used, as the link students made from the algebra to arithmetic, e.g. viewing scalar multiplication as repeated addition, was now likely much weaker. As such, the *source-path-goal* metaphor described in Wittmann et al. (2013) and the *containment* metaphor were dominant in the present study. An example is the substitution property of equality. This rule states that if two quantities are equal, one can be replaced by the other in any equation or expression. In terms of the source-path-goal metaphor, this can be linked to having a physical object (source) that is identical to a second object which is part of a larger system, and interchanging these identical objects such that the source becomes embedded into the larger system instead (goal). Interchanging these object does not change the larger system at hand. This concept is visualised in the video by dynamically placing the quantity $\frac{1}{2}c\Delta t$ over its equal quantity $L$ within an algebraic expression. In terms of the containment metaphor, the quantity $L$ can be seen as a container that is replaced by its contents, being the quantity $\frac{1}{2}c\Delta t$ that $L$ represents. To speak more directly to the concept of replacing a physical object, animated human hands executed the replacement of the quantity in the embodied dynamic condition.

**Script.** To ensure sufficient quality of the video script, preliminary research was done into existing time dilation explanatory videos on *YouTube*. We found for example that in our opinion existing videos spent too little attention on properly explaining the SRT postulates. We thus made sure to give both postulates a central place in our videos. The existing videos also helped us improve the structure of our script, as we ended up dividing the video into several distinct parts instead of discussing time dilation as one elaborate story. After the first version of the script was written, it was discussed with an SRT educational research expert before animation. The final version of the script can be found in Appendix A.

### ***Tests***

The pre-test contained 6 questions and focused on mapping participants' starting level. To this end, students rated both their prior special relativity theory knowledge as well as their prior skill in algebraic manipulations on a 5-point Likert scale. In addition, the test collected objective measures for the starting level by checking whether SRT had been part of the student's high-school program and checking the grade the student had achieved for the mathematics part of the course 'SK-BWSNK1: Wis- en natuurkunde 1' at Utrecht University.

The post-test consisted of 18 questions and was a mixture of open and closed questions, providing both quantitative and qualitative data. Table 2 shows what variables were measured in what questions in this test, using which question type. The score achieved on a variable was then the result of the sum of the scores achieved on the post-test questions that made up that variable, e.g., the achieved algebra score was the sum of the scores achieved on questions 7-11. The table also shows the maximum value achievable for each variable, what research question it addresses, what data-analysis was applied to the variable to link it to which other variable(s), and what figure or table shows the results of these analyses. For example, the post-hoc reported experienced cognitive load was measured in question 1 with the 9-point Likert-scale by Paas (1992; 1 = *very, very low mental effort*, 9 = *very, very high mental effort)*. The variable aided in answering the third research question and was compared over the different conditions through an analysis of covariance (ANCOVA), taking total prior knowledge as the covariate. The result of this analysis is found in Figure 6 in the results section.

Both the pre-test and most of the post-test were designed in this study, as we were dissatisfied with the extent to which existing questionnaires matched our drawn-up learning goals. Besides the used cognitive load question designed by Paas, there were no resources available for the specific variables we wanted to measure in the post-test in our chosen context. To ensure the quality of the designed post-test, it was reviewed by both an SRT and a mathematics educational research expert before use. All post-tests were hand-marked by both researchers independently using a designed answer model. Whenever we disagreed on a score by at least 1 full score point, we talked through our arguments for this score. If neither of us could be convinced, a final score was reached by averaging our two differently appointed scores. The complete post-test can be found in Appendix B.

## Data-analysis

To answer the first research question, we performed four analyses, as can be seen in Table 2. Firstly, we checked whether the mean algebra score differed significantly over the three conditions through an ANCOVA, taking prior skill in algebraic manipulations as a covariate. Through a second ANCOVA we compared the mean score for understanding of physical concepts and reasoning over the conditions, controlling for prior SRT knowledge. We wanted to check whether cueing attention in the embodied dynamic and dynamic conditions influenced student learning outcomes. Here we used the effort students self-reported to have put into following the algebra in the videos as a proxy for attention. Thus thirdly, we checked whether this self-assessed student effort differed over the conditions through an analysis of variance

**Table 2**

*Post-test variables and role in analyses*

| Variable (SA = Self-assessed) | Post-test Questions | Sub Question Type | Max. Score | RQ | Type of Analysis | Linked to | Figure/ Table |
|---|---|---|---|---|---|---|---|
| Experienced cognitive load | 1 | Likert | 9 | 3 | ANCOVA | Conditions using covariate: prior knowledge | Figure 6 |
| Rated video attractiveness | 2 | Likert | 5 | 4 | Kendall correlation | SA understanding video | Figure 7 |
| | | | | | Kendall correlation | Total score | Figure 7 |
| SA understanding video | 3, 5, 6 | Likert | 15 | 4 | Kendall correlation | Rated video attractiveness | Figure 7 |
| SA influence audio-visual traits on comprehensibility | 4a, 4b | Likert & Open | 5 | - | - | - | - |
| Algebra score | 7 - 11 | Open | 10 | 1 | ANCOVA | Conditions using covariate: prior skill in algebraic manipulation | Figure 1 |
| | | | | 2 | Partial Pearson correlation | Conceptual score<br>SA algebra effort | Figure 5 |
| | | | | 1 | Kendall correlation | | Figure 4 |
| Conceptual score | 12 - 15 | Open & Multiple Choice | 8 | 1 | ANCOVA | Conditions using covariate: prior SRT knowledge | Figure 2 |
| | | | | 2 | Partial Pearson correlation | Algebra score | Figure 5 |
| Total score | 7 - 15 | Open & Multiple Choice | 18 | 4 | Kendall correlation | Rated video attractiveness | Figure 7 |
| SA correlation algebraic and conceptual understanding | 16 | Likert | 5 | 2 | Descriptive | - | - |
| SA algebra effort | 18 | Likert | 5 | 1 | ANOVA | Conditions | Figure 3 |
| | | | | 1 | Kendall correlation | Algebra score | Figure 4 |

*Note.* RQ refers to the research question(s) the variable corresponds to.

(ANOVA). Finally, we correlated this effort to the algebra scores through a Kendall correlation. Field (2009) describes that a Kendall correlation ($\tau$) should be used when the dataset is small, a variable is not-normally distributed over the conditions, and many participants achieved the same score on that variable, all of which is the case here. For answering the second research question, we performed a partial Pearson correlation between algebra score and conceptual scores for each condition. Here we controlled for the total prior knowledge, meaning prior skill in algebraic manipulation and prior special relativity theory knowledge combined. The third research question was answered through an ANCOVA, checking whether the post-hoc experienced cognitive load differed significantly over the conditions, also controlling for total prior knowledge. Finally, to answer the last research question, we performed a Kendall correlation between the rated video attractiveness and self-assessed total understanding of the video, as well as a Kendall correlation between this rated video attractiveness and total score, meaning the score on the algebraic and conceptual questions in the post-test combined.

All analyses were performed in SPSS version 26. Out of 38 students, two had not yet completed the course ‘SK-BWSNK1: Wis- en natuurkunde 1’. As such, we had no objective measure for the starting level concerning skill in algebraic manipulation for these students. These two answers were coded as missing, meaning these two students were not included in analyses in which either prior skill in algebraic manipulation or total starting level were involved.

## Results

This section describes the analyses performed for answering the research questions. For AN(C)OVAs, the assumptions and findings are described separately. Firstly however, the role of prior knowledge and video features are described.

### Prior knowledge and video features

A Pearson correlation showed that total prior knowledge correlated moderately with total score, $r = .48$, $p = 0.003$. As such, the expectation that prior knowledge influenced score was confirmed, a result that speaks for the validity of the designed post-test. Descriptive analysis furthermore showed this prior knowledge to be approximately equally distributed over the three conditions, indicating it could be taken as a control variable when desired in further analyses (see Table 3). Additionally, we tested whether the results found regarding our research questions might be attributed to video quality, by checking to what extent video features helped students understand the video, i.e., tempo, structure, use of voice, visualised hand(s) in the algebra, visualised dynamics in the algebraic derivations, other visualised dynamics in the video, and visual design. Corresponding student responses were all positive, with no differences found between conditions.

**Table 3**

*Mean of Normalized Scores on Total Prior Knowledge*

| Condition | *N* | Mean | Standard Deviation |
|---|---|---|---|
| Embodied dynamic | 12 | 0.65 | 0.18 |
| Dynamic | 13 | 0.63 | 0.17 |
| Static | 11 | 0.64 | 0.17 |

### The effect of condition on algebra score

To address the first research question whether dynamic and embodied dynamic algebra animation improve understanding of both the algebraic manipulations, as well as the physical concepts and reasoning in the videos, we performed two ANCOVAs, one ANOVA, and one correlation analysis. The first ANCOVA served to determine whether the means of the algebra scores differed significantly over the three conditions, including prior skill in algebraic manipulation as a covariate. As can be seen in Table 2 (p. 20), a maximum score of 10 could be achieved.

### *Test for assumptions*

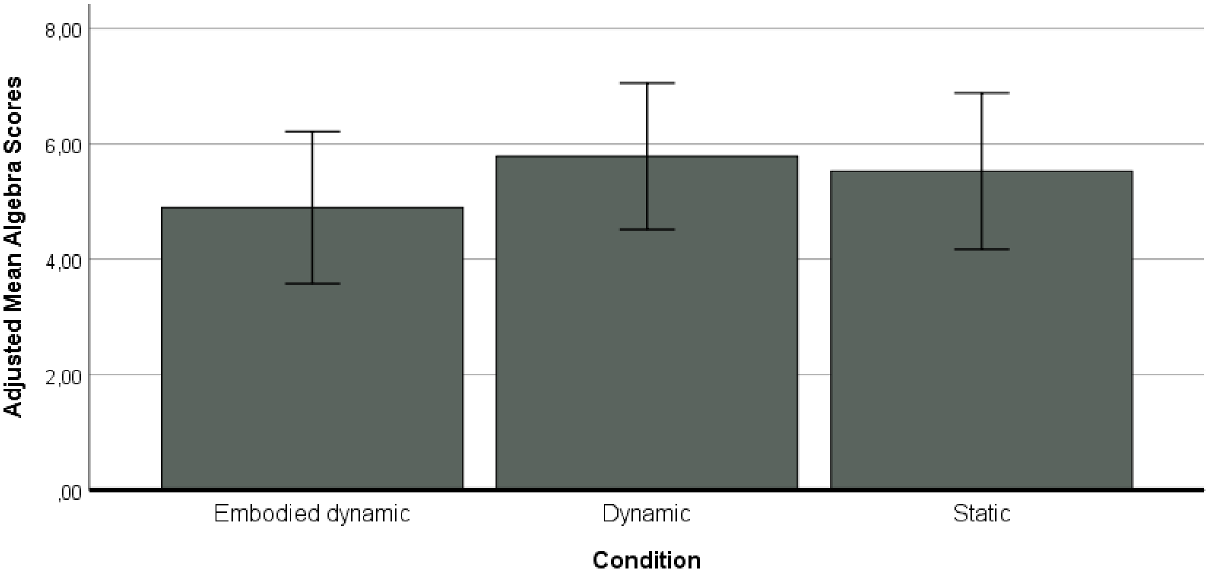
A descriptive analysis revealed no outliers in the algebra scores. Furthermore, an ANOVA showed these scores to be approximately normally distributed for all three conditions, with $D(13) = 0.92$, $p = .26$ for the embodied dynamic condition, $D(13) = 0.92$, $p = .23$ for dynamic, and $D(12) = 0.96$, $p = .83$ for static. This ANOVA also showed both the mean score and regression weights of prior skill in algebraic manipulations as the control variable to not differ significantly over the three conditions, respectively $F(2,33) = 1.10$, $p = .34$ and $F(2,30) = 0.16$, $p = .85$. Levene's test however showed significantly different variances for the three conditions, $F(2,33) = 3.84$, $p = .03$. Nonetheless, the decision was made to still perform the ANCOVA, as evidence suggests AN(C)OVAs to be quite robust to violations of the assumptions of normality and homogeneity of variances (Field, 2009; Stevens, 2012). In particular, this is the case when the group sizes are roughly equal, meaning the ratio between the largest and smallest group size is smaller than 1.5, which is the case in this study (Stevens, 2012).

### *Findings*

In contrast to the expectations, the mean algebra score adjusted for prior skill in algebraic manipulation was lowest in the embodied dynamic condition ($M_{adj} = 4.90$, $SD = 0.65$; dynamic condition: $M_{adj} = 5.79$, $SD = 0.62$; static condition: $M_{adj} = 5.53$, $SD = 0.67$; see Figure 1). There was however no significant effect of condition on this adjusted mean algebra score, $F(2, 32) = 0.50$, $p = .61$, partial $\eta^2 = .03$. Thus, controlled for prior skill in algebraic manipulation, condition explained only 3% of the total variance of the algebra score. We performed additional planned contrasts to check for significance between our two experimental conditions compared to the control condition (static). These planned contrasts revealed no significant effect on algebra score both between the embodied dynamic and static condition, $p = .50$, $\eta^2_{\text{ED vs S}} = 0.01$, and between the dynamic and static condition, $p = .78$, $\eta^2_{\text{D vs S}} = 0.003$.

**Figure 1**

*Differences in Adjusted Mean Algebra Score*



*Note.* The covariate, mean of normalized scores on prior algebraic skill, is evaluated at the value of 0.79. The error bars represent the 95% confidence interval.
*p = .61.*

## The effect of condition on conceptual score

The second ANCOVA performed for answering the first research question determined whether the means of the scores on the questions dealing with understanding of physical concepts and reasoning differed significantly over the three conditions. Now, prior conceptual knowledge of special relativity theory was implemented as a predicting control variable, and the maximum achievable conceptual score was eight (see Table 2, p. 20).

### *Test for assumptions*

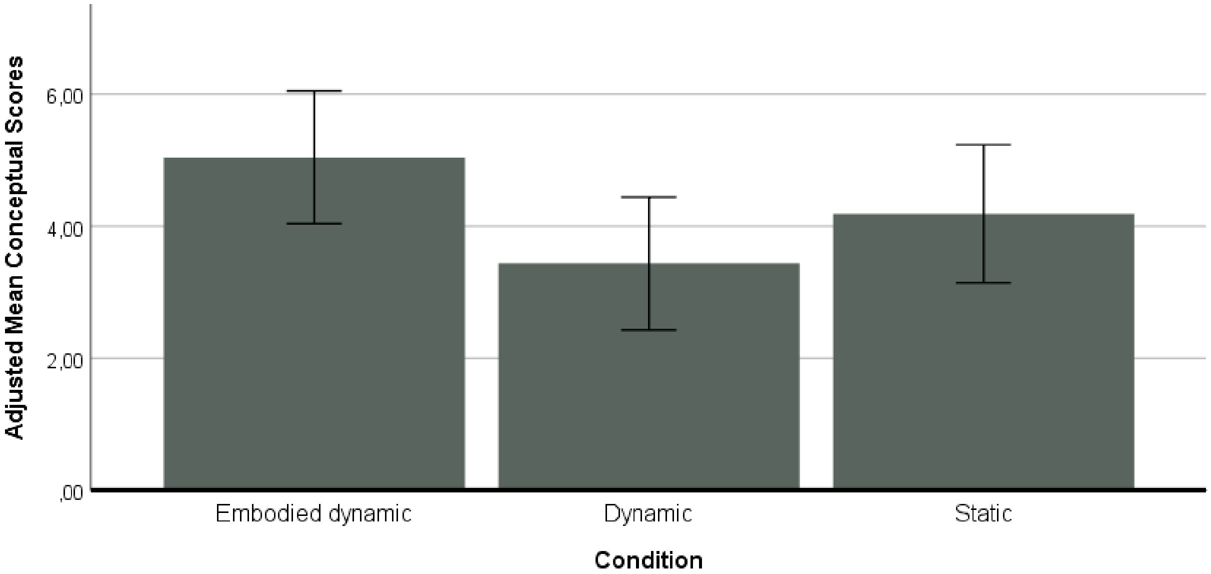
Descriptive analysis of the scores on the conceptual questions revealed one mild outlier in the dynamic condition but no extreme outliers. An ANOVA showed these scores to be approximately normally distributed for all three conditions, with $D(13) = 0.99$, $p = .996$ for embodied dynamic, $D(13) = 0.92$, $p = .28$ for dynamic, and $D(12) = 0.96$, $p = .77$ for static. Furthermore, the variances were approximately equal for the three conditions, $F(2,35) = 3.07$, $p = .06$. Additionally, the mean score and regression weights of prior SRT knowledge as the control variable did not differ significantly over the three conditions, respectively $F(2,35) = 0.03$, $p = .98$ and $F(2,32) = 0.04$, $p = .96$. As such, all ANCOVA preconditions were met.

### *Findings*

As expected, the mean conceptual score adjusted for prior SRT knowledge was highest in the embodied dynamic condition ($M_{adj} = 5.04$, $SD = 0.50$). Unexpectedly however, it was lowest in the dynamic condition ($M_{adj} = 3.44$, $SD = 0.50$; static: $M_{adj} = 4.19$, $SD = 0.52$; see Figure 2). There was no significant effect of condition on this mean adjusted conceptual score, $F(2, 34) = 2.64$, $p = .09$, partial $\eta^2 = 0.13$. Planned contrasts however revealed that the conceptual score was significantly higher in the embodied dynamic compared to the dynamic condition, $p = .03$, $\eta^2_{\text{D vs ED}} = 0.13$. There was no significant difference between the embodied dynamic and the static condition, $p = .24$, $\eta^2_{\text{ED vs S}} = 0.04$, nor between the dynamic and static condition, $p = .30$, $\eta^2_{\text{D vs S}} = 0.03$.

**Figure 2**

*Differences in Adjusted Mean Conceptual Score*



*Note*. The covariate, mean of normalized scores on prior knowledge on special relativity theory, is evaluated at the value of 0.51. The error bars represent the 95% confidence interval.
*p = .09.*

### The effect of condition on algebra effort

The ANOVA corresponding to the first research question determined whether the means of the self-assessed effort students made to follow the algebraic derivations in the video differed significantly over the three conditions. This effort was assessed on a five-point Likert scale.

### *Test for assumptions*

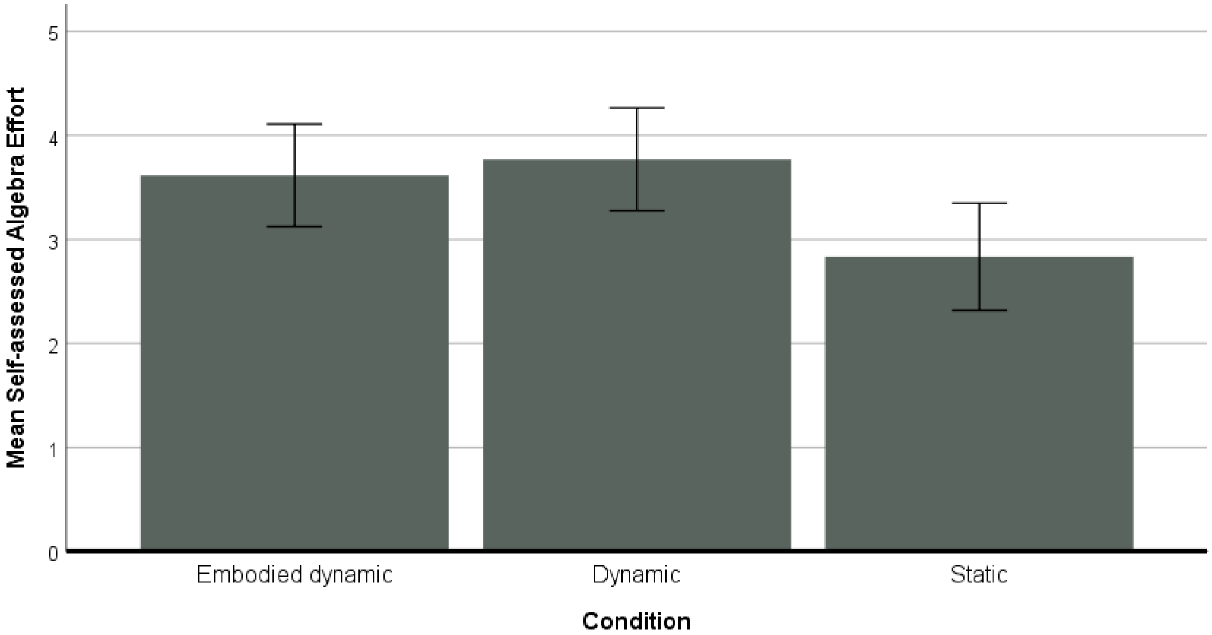
Through descriptive analysis, one mild outlier was found in the self-assessed effort made to follow the algebra in the static condition, but no extreme outliers. An ANOVA however showed the scores to be significantly non-normally distributed in all conditions, with $D(13) = .76$, $p = .002$ for the embodied dynamic condition, $D(13) = 0.84$, $p = .02$ for dynamic, and $D(12) = .75$, $p < 0.01$ for static. Furthermore, the assumption of homogeneity of variances was violated, $F(2,35) = 6.93$, $p < .01$. However, this ANOVA was still performed due to the earlier mentioned robustness of this analysis.

### *Findings*

The ANOVA revealed a significant effect of condition on the mean self-assessed effort made to follow the algebraic derivations in the video, $F(2, 35) = 4.02$, $p = .03$, $\eta^2 = .19$ (see Figure 3). Planned contrasts with the static condition as the reference category revealed that the mean self-assessed algebra effort was significantly higher in the embodied dynamic condition ($M = 3.62$, $SD = 0.77$) than in the static condition ($M = 2.83$, $SD = 0.58$), $p = .03$, $\eta^2_{\mathrm{ED\ vs\ S}} = 0.12$. The algebra effort was also significantly higher in the dynamic condition ($M = 3.77$, $SD = 1.17$) compared to the static condition, $p = .01$, $\eta^2_{\mathrm{D\ vs\ S}} = 0.17$.

**Figure 3**

*Differences in Mean Self-assessed Effort to Follow Algebraic Derivations*



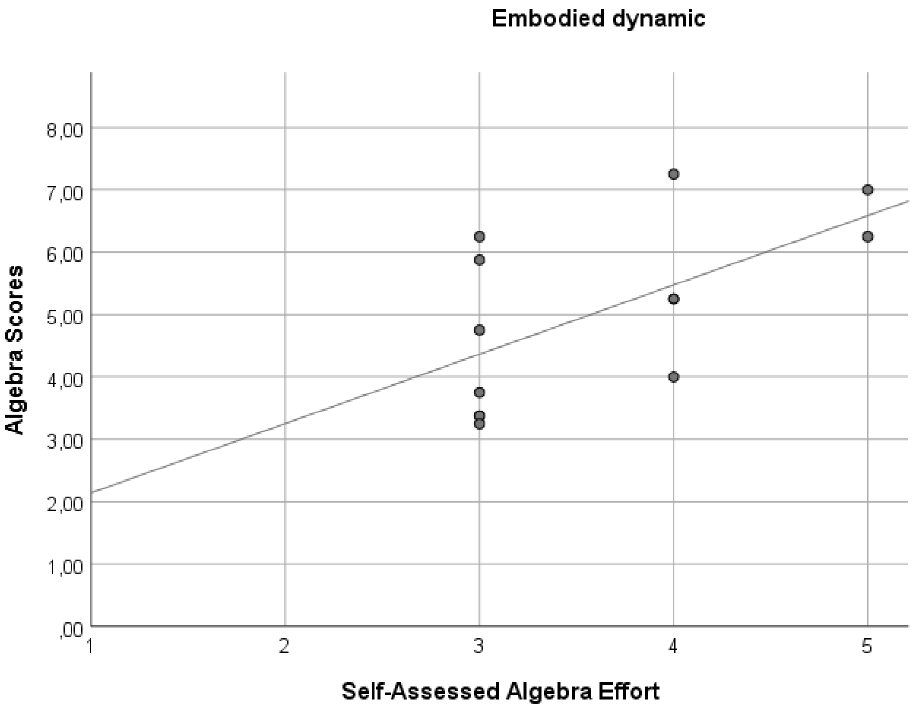
*Note.* The error bars represent the 95% confidence interval.
*$p = .03$.

## The association between self-assessed algebra effort and algebra score

As a final step to answer the first research question we performed a correlation analysis for each condition separately between the self-assessed effort made to follow the algebraic derivations in the video and algebra score. This correlation was moderately positive in the embodied dynamic condition, Kendall's $\tau = 0.51$, $p = .03$, indicating that students in this condition who put more effort into following the algebra in the video scored significantly better on algebra questions in the post-test. There were only weak correlations in the dynamic and static conditions, respectively $\tau = 0.24$, $p = .30$ and $\tau = 0.29$, $p = .24$. A scatterplot of the significant correlation in the embodied dynamic condition can be seen in Figure 4.

**Figure 4**

*Correlation Self-assessed Algebra Effort and Algebra Score Embodied Dynamic Condition*



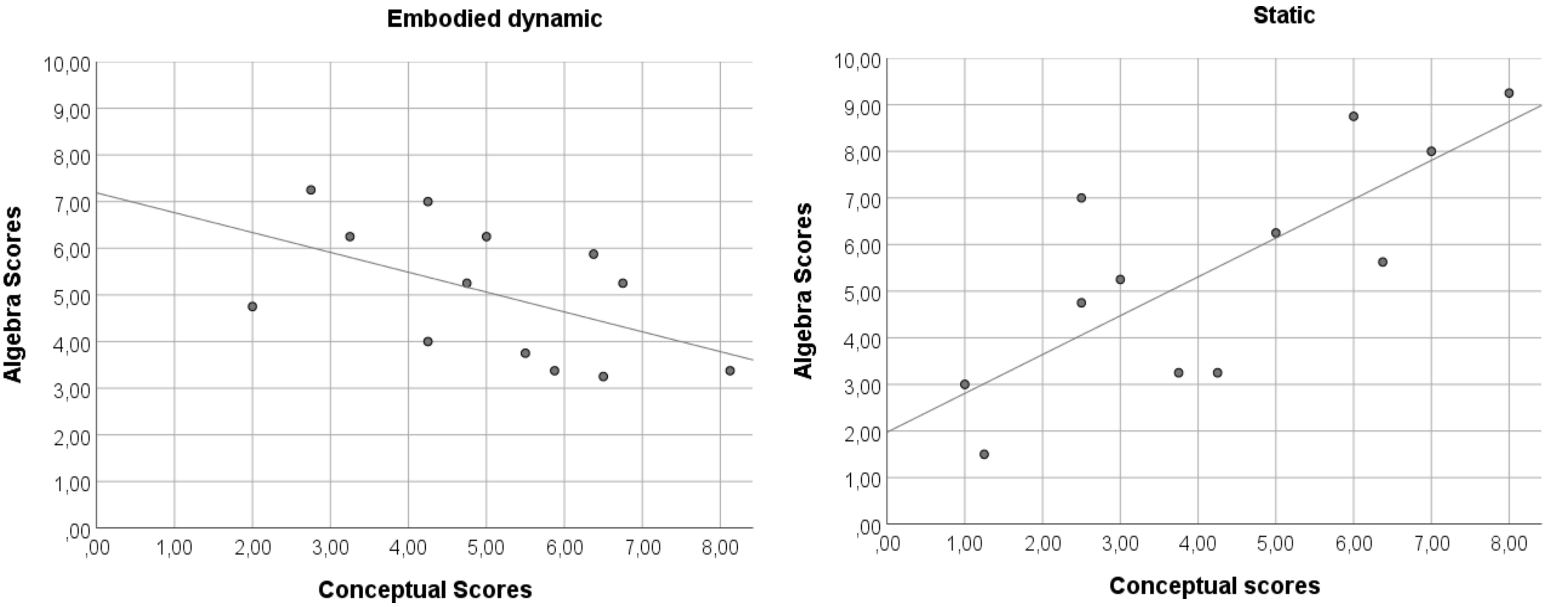
*Note.* **p = .03.*

## The association between algebra score and conceptual score

To answer the second research question whether understanding the algebraic manipulations in the videos relates to understanding their physical concepts and reasoning, we performed a partial correlation for each condition separately between the algebra score and conceptual score, with total prior knowledge included as the control variable. Unexpectedly, we found a strong negative partial correlation in the embodied dynamic condition, Pearson's $r$ = -0.69, $p$ = .02, indicating that a higher algebra score corresponded with a lower conceptual score and vice versa. Moreover, it was the static condition in which we found a strong positive partial correlation, Pearson's $r$ = 0.71, $p$ = .02. Scatterplots of these strong correlations can be found in Figure 5. Note that in these scatterplots, total prior knowledge is not controlled for. These plots merely serve to give an indication of the existing significant correlation. In the dynamic condition, a weak correlation between algebra score and conceptual score was found, *r* = 0.24, *p* = .45.

**Figure 5**

*Correlation Algebra Scores and Conceptual Scores Embodied Dynamic and Static Condition*



*Note.* In these plots prior knowledge is not taken into account.
Respectively *p* = .02, *p* = .02.

We also asked the students themselves to rate on a 5-point Likert scale to what extent they felt an association exists between the understanding of algebraic manipulations and physical concepts. The mean student self-assessment over the three conditions combined was *M* = 2.89 (*SD* = 0.80). This means the students also expected a correlation between understanding of algebraic manipulations and physical concepts and reasoning to exist, as was found in the embodied dynamic and static conditions.

## The effect of condition on experienced cognitive load

To address the third research question to what extent the results concerning the first two research questions can be explained in terms of experienced cognitive load, we performed an ANCOVA to determine whether the post-hoc reported experienced cognitive load differed significantly over the three conditions, including total prior knowledge as a control variable. This cognitive load was measured on a 9-point Likert scale.

### *Test for assumptions*

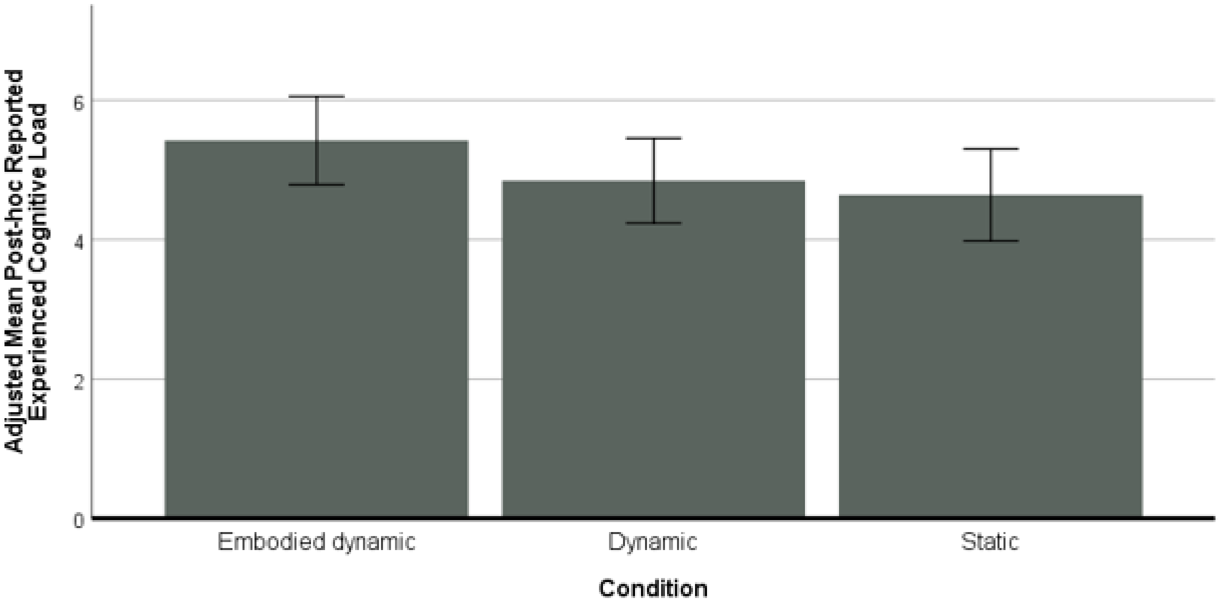
An ANOVA showed that the assumptions of the control variable being independent of the group effect and homogeneity of regression slopes were met, respectively $F(2,33) = 0.09$, $p = .91$ and $F(2,30) = 0.18$, $p = .84$. However, six extreme outliers were found in the dynamic condition. Furthermore, even though the scores were approximately normally distributed in the static condition, $D(12) = 0.86$, $p = .05$, they were significantly non-normal in the other conditions, with $D(13) = 0.71$, $p = .001$ for the embodied dynamic condition, and $D(13) = 0.82$, $p = .01$ for the dynamic condition. Additionally, the variances were significantly different for the three conditions, $F(2,33) = 3.71$, $p = .04$. The earlier discussed ANCOVA robustness aside, the choice to still perform this analysis despite finding six extreme outliers is elaborated in the discussion section.

### *Findings*

Unexpectedly, the mean post-hoc reported experienced cognitive load adjusted for total prior knowledge was highest in the embodied dynamic condition ($M_{adj} = 5.42$, $SD = 0.31$) and lowest in the static condition ($M_{adj} = 4.64$, $SD = 0.33$). There was however no significant effect of condition on this mean adjusted cognitive load, $F(2, 32) = 1.66$, $p = .21$, partial $\eta^2 = .09$ (see Figure 6). Planned contrasts with the static condition as the reference category revealed no significant effect on experienced cognitive load both between the embodied dynamic and static condition, $p = .09$, $\eta^2_{\text{ED vs S}} = 0.09$, and between the dynamic condition ($M_{adj} = 4.84$, $SD = 0.30$) and static condition, $p = .65$, $\eta^2_{\text{D vs S}} = 0.007$.

**Figure 6**

*Differences in Adjusted Mean Post-hoc Reported Experienced Cognitive Load*



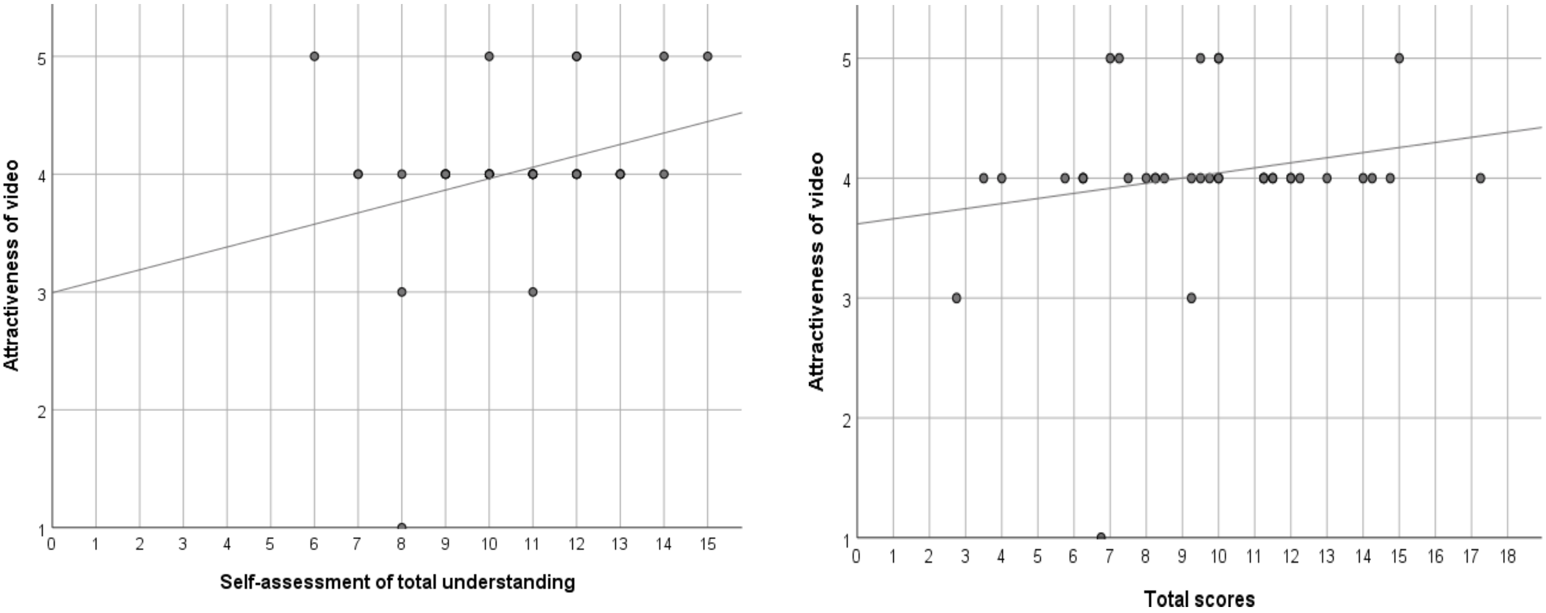
*Note.* The covariate, mean of normalized scores on total prior knowledge, is evaluated at the value of 0.64. The error bars represent the 95% confidence interval.
$p = .21$.

## The associations between rated attractiveness, self-assessed understanding and total score

To answer the final research question what the relationship is between how attractive the videos are rated, how well students feel they understood the video, and how well they truly understood it, we performed two correlation analyses. The first was a Kendall correlation between the video attractiveness, rated on a 5-point Likert scale, and the self-assessed student understanding of the video, rated in three separate 5-point Likert scale questions. As such, the maximum score for this self-assessed understanding was 15 (see Table 2, p. 20). This weak correlation was only marginally significant, Kendall's $\tau = 0.25$, $p = .07$. The second correlation was a Kendall correlation between the rated video attractiveness and the total score on the algebra and conceptual questions combined, on which a maximum score of 18 could be achieved (see Table 2). This very weak correlation was not significant, Kendall’s $\tau = 0.13$, $p = .33$. Thus, rated video attractiveness correlated more strongly to self-assessed understanding than it did to actual performance on the algebraic and conceptual questions. The corresponding scatterplots can be seen in Figure 7.

**Figure 7**

*Correlation Rated Video Attractiveness versus Self-assessed Student Understanding (Left) and Rated Video Attractiveness versus Total Understanding (Right)*



*Note.* In both plots, one datapoint can represent several participants. In the left plot many participants achieved identical scores.
*p = .07 (left), p = .33 (right).*

# Discussion

## *Results*

Through this project, we tried to get one step closer to concluding the fundamental debate on what makes a good educational video. Results regarding the first research question showed that embodied dynamic algebra animation did not improve algebra learning outcomes in comparison with static algebra animation. For dynamic compared to static algebra animation, the algebra learning outcomes did improve, yet not significantly. Embodied dynamic animation did significantly improve learning outcomes related to physical concepts and reasoning compared to dynamic algebra animation. Furthermore, students reported to put significantly more effort into following algebra that is either dynamically or embodied dynamically visualised. For embodied dynamic algebra visualisation, there was a moderate positive correlation between this effort and algebra learning outcomes. With these results, our hypothesis that embodied dynamic algebra animation is most effective regarding understanding of both algebraic manipulations and physical concepts involved in time dilation, followed by dynamic algebra animation, was only partially met. For the second research question, results showed an unexpected strong negative correlation in the embodied dynamic condition, meaning that students who scored higher on algebra, scored lower on conceptual questions and vice versa. The correlation was strongly positive for static algebra animation. Thus, our hypothesis that due to attention cueing better understanding both the algebraic manipulations as well as the physical concepts and reasoning were interwoven in the embodied dynamic and dynamic conditions was not met. As the conditions did not significantly affect our measured post-hoc reported experienced cognitive load, our hypothesis that cognitive load could explain results to the first two research questions was also not met. Finally, rating of the attractiveness of the video correlated more strongly with self-assessed understanding of the video than with objective performance on algebraic and conceptual questions. Hence, this might give an interesting lead to further explore the hypothesis that students who give a high rating to the attractiveness of the video think they understand the content of the video better than they actually do.

Some of the results in the embodied dynamic condition might still be explained by the research by De Koning et al. (2007) underlying our second hypothesis about attention cueing. These researchers state that attention cueing improves understanding and transfer of both the cued information (the algebraic video part) as well as the uncued information (the conceptual video part), meaning understanding of these parts would be interwoven. Considering the strong attention cueing in the embodied dynamic condition due to the animated moving hands, the higher found conceptual scores in the embodied dynamic condition compared to the dynamic condition are in line with the findings of De Koning et al. (2007). Many other results in the embodied dynamic condition are however very different from what these researchers would have us expect. Looking at other research, the influence of attention cueing on learning outcomes points in different directions. Kriz and Hegarty (2007) found that although cueing led to more attention directed to the task, this did not lead to a better understanding of the corresponding learning material. Considering these contradicting results and the extent to which our other found results for the embodied dynamic condition deviate from those expected based on the research by De Koning et al. (2007), we looked for an alternative explanation as to how attention cueing might have influenced the other results in this condition.

One such alternative explanation might be that due to the strong attention cueing, a competition factor may be occurring between the attention spent on the algebraic video part versus on the following conceptual part in this condition, reflecting on the corresponding algebra and conceptual learning outcomes. In the embodied dynamic condition, we now believe we can distinguish between two subgroups based on student attention. The first subgroup consists of students who, due to the strong attention cueing towards the algebra, spent much of their attention span on this algebraic video part. This group then scores higher on this algebra, but, due to the described attention competition factor, may have less attention to spend on the conceptual part, scoring lower there. The second subgroup consists of students who, due to the high level of support provided through the animated human hands, might not be challenged enough in following the algebra, meaning that despite the cueing, they might have spent only little of their attention span on following this algebraic video part. This group scores lower on the algebra, but may score higher on conceptual questions as more attention is left to be spent on this conceptual video part. This existence of these high and low algebra attention subgroups, combined with the attention competition factor, would then explain the negative correlation between algebraic and conceptual learning outcomes, the positive correlation between effort put into the algebra and algebra score, and why these algebra scores are not higher overall in the embodied dynamic opposed to the dynamic and static conditions. Finally, we think the strong attention cueing towards the algebra also explains why students reported putting more effort into following this algebra here than in the static condition.

This last result concerning effort is similar for the dynamic condition. This could also be explained by attention cueing, which in this condition solely stems from the movements animated within the algebra. The other findings in the dynamic condition, such as higher algebra scores compared to the static condition and a positive correlation between algebra scores and conceptual scores (all not significant) are in line with our hypotheses.

In the static condition, attention is not strongly directed to the algebraic manipulations. This might explain reports of lower effort put into the algebra. Furthermore, without the attention cueing, perhaps understanding is the sole main variable underlying learning outcome. Possibly, this allows for the understanding of the algebraic part to support understanding of the conceptual part, explaining their positive correlation.

Concerning post-hoc reported experienced cognitive load, we suspect we did not measure what we wanted to measure. In hindsight, we realised we should have measured this variable not only after students watched the entire video, but also immediately after they watched its algebraic part. Presently, the assumption is made that the cognitive load experienced while watching the conceptual video part was equal over the three conditions. This does not need to be the case, since attention spent on this part possibly differs over the conditions. As such, we cannot conclusively say whether cognitive load can explain the found results.

### ***Data and design choices***

From the analysis of student answers on questions related to algebraic and conceptual learning outcomes it seems that watching the videos enabled students to reason properly within the videos' discourse. Scores were however low when the concept of time dilation had to be transferred onto a new situation, indicating that properly applying the video material after watching is still a challenge. This result might be partly due to difficulties in grasping the different time variables after introduction of Einstein's frame-dependent time coordinates. Considering scores were also low on a conceptual question related to these coordinates and given the absence of practice in properly applying them, perhaps the corresponding explanation provided in the videos went a bit fast.

Considering the data-analysis, both the variables algebra score and conceptual score were not unidimensional. For example, one conceptual question tested understanding of the light postulate, while another tested the ability to transfer the concept of time dilation onto a new situation. As such, no internal consistency of the sub questions that make up these two variables was expected, meaning Cronbach's alpha was not applied. Furthermore, in the ANCOVA for the post-hoc reported experienced cognitive load six extreme outliers were found in the dynamic condition. Thus, these six scores diverged by more than three times the interquartile range from the box of the boxplot (Dawson, 2011). In this dynamic condition, the first and third quartile both corresponded to a reported cognitive load of five, meaning the interquartile range was zero. As such, any reported cognitive load unequal to five on Paas' scale was deemed an extreme outlier by SPSS. Had the interquartile range been even only one, no extreme outliers would have been present. We therefore accepted the presence of these outliers and still performed the ANCOVA.

Reflecting on made design choices, we still stand by most. For example, the decision was made to keep the 'visual language' consistent throughout the video to keep the story clearer, meaning that the same example of a rocket at rest and a moving rocket are repeatedly used in explanations in all parts of the videos. Furthermore, we discussed about the frequency and length of small few-second breaks between different parts of the videos. In hindsight, we think perhaps some of these breaks could have been extended a bit more, to more clearly break-up the time dilation story into smaller bits. Finally, the explanation of Einstein's frame-dependent time coordinates could have been a little more elaborate. Visualising four different time coordinates at the same time on screen may have made this part difficult to follow.

### *Limitations and future research*

A limitation of the study is that the intervention is quite short. In the 7:21 minutes videos, the different conditions are only reflected in the main derivation in the algebraic part, which lasts about 55 seconds. This limits the influence the conditions might have had on the results. Furthermore, regarding the embodied dynamic condition, we made use of embodied simulation instead of enactment. This translates into a relatively weaker bodily experience, further lowering the impact of this condition (Castro-Alonso et al., 2014a). In terms of the participants, students were able to voluntarily apply for participation based on reading a short description of the study on blackboard. As such, participating students might have had a higher-than-average affinity with mathematics, decreasing transferability of the study. Additionally, due to the small group size of

38 participants, our power may be on the low side. It is however comparable to similar research by Castro-Alonso et al. (2014b), who held 14 participants in each condition, compared to 13 and once 12 in the current study. Perhaps some of the multiple found results in our study that bordered on being significant, such as the association between rated attractiveness and self-assessed understanding, might have crossed that line in case of a higher power. Finally, the data was administered online instead of in a controlled environment. Even though we made sure all students participated individually in a quiet room, the fact that the environment was different for different participants might have influenced the data.

Future research might investigate the influence of the embodied dynamic condition on learning outcomes in case enactment is included, for example by having people use their own hands in some way to execute the algebraic manipulations. Perhaps in this case higher conceptual scores in the embodied dynamic compared to the static condition might be found as well. Furthermore, measuring experienced cognitive load both after the video part in which the different conditions are manifested as well as after the end of the video could provide more in-depth insight into whether this cognitive load might explain the found results. Additionally, we investigated the influence of attention cueing by using effort as a proxy for attention and by measuring this variable only for the algebraic video part. Future research might measure attention more directly and elaborately over the video and its different parts, to further elucidate the influence of attention cueing on learning outcomes. Finally, the wow-effect could be explored further. With the association between rated attractiveness and self-assessed understanding being both marginally significant and stronger than the association between this attractiveness and objective performance, we believe there might be an interesting phenomenon there.

# Bibliography

Alstein, P., Krijtenburg-Lewerissa, K., & Van Joolingen, W. R. (2020). Teaching and learning special relativity theory in secondary and lower undergraduate education: A literature review. *arXiv preprint arXiv:2011.08059*.

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 21*(6), 695-700.

Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education, 101*, 150-167.

Bétrancourt, M., & Tversky, B. (2000). Effect of Computer Animation on Users' Performance: A Review. *Le Travail Humain, 63*(4), 311-329.

Bos, R., & Renkema, W. (Forthcoming 2022). Metaphor-based algebra animation. *Accepted for publication in the Proceedings of the 15th International Conference on Technology in Mathematics Teaching*.

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2014a). Dynamic visualisations and motor skills. In W. Huang (Ed.), *Dynamic visualisations and motor skills.* (pp. 551-580). New York: Springer.

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2014b). Learning from observing hands in static and animated versions of non-manipulative tasks. *Learning and Instruction, 34*, 11-21.

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2016). Comparing apples and oranges? A critical look at research on learning from statics versus animations. *Computers & Education, 102*, 234-243.

Castro-Alonso, J. C., Ayres, P., Wong, M., & Paas, F. (2018). Learning symbols from permanent and transient visual presentations: Don't overplay the hand. *Computers & Education, 116*, 1-13.

Dawson, R. (2011). How significant is a boxplot outlier? *Journal of Statistics Education, 19*(2), 1-13.

De Koning, B. B., & Tabbers, H. K. (2013). Gestures in Instructional Animations: A Helping Hand to Understanding Non-human Movements? *Applied Cognitive Psychology, 27*, 683–689.

De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2007). Attention Cueing as a Means to Enhance Learning from an Animation. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 21*(6), 731-746.

Field, A. (2009). *Discovering Statistics Using SPSS.* London: SAGE Publications Ltd.

Garbarini, F., & Adenzato, M. (2004). At the root of embodied cognition: Cognitive science meets neurophysiology. *Brain and cognition, 56*(1), 100-106.

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*, 722-738.

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies, 65*(11), 911-930.

Lakoff, G., & Núnez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being.* New York: Basic Books.

Martin, S. (2014). Measuring Cognitive Load and Cognition: Metrics for Technology-enhanced Learning. *Educational Research and Evaluation*, 592-621.

Mayer, R. (2020). *Multimedia Learning* (third ed.). Cambridge: Cambridge University Press.

Muller, D. A. (2008). *Designing Effective Multimedia for Physics Education.* Sydney: University of Sydney.

Paas, F. (1992). Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach. *Journal of Educational Psychology, 84*(4), 429-434.

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences.* Routledge.

Sundararajan, N., & Adesope, O. (2020). Keeping it Coherent: A Meta-Analysis of the Seductive Details Effect. *Educational Psychology Review, 32*, 707-734.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science, 2*, 25-285.

Sweller, J., Van Merriënboer, J. J., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review, 31*, 261–292.

Wijnker, W., Bakker, A., Van Gog, T., & Drijvers, P. (2019). Educational videos from a film theory perspective: Relating teacher aims to video characteristics. *British Journal of Educational Technology, 50*(6), 3175–3197.

Winter, B., & Yoshimi, J. (2020). Metaphor and the Philosophical Implications of Embodied Mathematics. *Frontiers in Psychology, 11*, 1-13.

Wittmann, M. C., Flood, V. J., & Black, K. E. (2013). Algebraic manipulations as motion within a landscape. *Educational Studies in Mathematics, 82*(2), 169-181.

Xie, H., Wang, F., Hao, Y., Chen, J., An, J., Wang, Y., & Liu, H. (2017). The more cognitive load is reduced by cues, the better retention and transfer of multimedia learning: A meta analysis and two meta-regression analyses. *PloS one, 12*(8).

# Appendix A: Script

| | |
|---|---|
| Introductie (20 sec)<br>Tijddilatatie is het verschijnsel dat een bewegende klok langzamer tikt dan een stilstaande klok. In deze video gaan we deze spectaculaire uitspraak, die onderdeel is van Einstein's beroemde speciale relativiteitstheorie, uitwerken en nuanceren, en drukken we in formules uit hoeveel trager de tijd gaat in een bewegende klok; en daarvoor is niet veel meer wiskunde nodig dan de stelling van Pythagoras. | [woord verschijnt] [plaatje Einstein]<br>[Twee klokken (twee kleuren, zodat aan het einde de rollen nog omgedraaid kunnen worden): één beweegt en tikt langzamer] |
| Lichtpostulaat (60 seconden)<br>Tijddilatatie is een direct gevolg van het zogeheten **lichtpostulaat**: het uitgangspunt dat de lichtsnelheid hetzelfde is, namelijk ongeveer 300.000 km/s, of je nou constant beweegt of stilstaat. | [woord verschijnt]<br>[lichtstraal komt langs en beide waarnemers meten dezelfde snelheid] |
| Laten we dit verduidelijken door het gooien met een bal te vergelijken met het schijnen met een zaklamp. Een bal gegooid vanaf een stilstaande raket vliegt met een lagere snelheid dan een bal gegooid vanaf een raket die met hoge constante snelheid beweegt. De snelheid van de bal hangt dus af van de snelheid van de raket. | [Dit animeren] |
| De lichtstraal van de zaklamp beweegt verrassend genoeg in beide situaties even snel: met de lichtsnelheid van ongeveer 300.000 km/s. Dit is het lichtpostulaat: De snelheid van het licht hangt niet af van de snelheid of richting van de raket. | [Dit animeren en de zaklamp in verschillende richtingen laten schijnen. Bij elke straal in elke richting 300.000 km/s laten verschijnen] |
| Intro klok (60 seconden)<br>Het lichtpostulaat speelt een centrale rol in tijddilatatie, dat we uitleggen met behulp van een gigantische eierwekker. Aan één kant van de | [Verschijnt met A erin] |

| | |
|---|---|
| eierwekker staan een lichtbron en een lichtsensor, aan de andere kant een spiegeltje op 45.000.000 km afstand. De stilstaande lichtbron stuurt een lichtpuls uit. We volgen alleen het deel van deze lichtpuls dat weerkaatst wordt in het spiegeltje en daarna op de lichtsensor valt. De sensor vangt dit licht na precies 5 minuten op. Het ei is klaar, Yum! Sommigen willen hun eitje misschien net wat harder of zachter en daarom variëren we de afstand $L$ tussen de lichtbron en de spiegel. Er geldt dan het verband $\Delta t = \frac{2L}{c}$ voor de tijd in seconden dat het duurt tot de lichtpuls via de spiegel de lichtsensor bereikt, waarbij $c$ staat voor de lichtsnelheid. | [Afstand laser – spiegel verandert en $L$ verschijnt]<br>[tijd = afstand/snelheid<br>= 2 x 45.000.000/300.000<br>= 300 seconden (5 minuten)] |
| Stel eenzelfde soort eierwekker B beweegt met constante snelheid $v$ ten opzichte van A. We volgen weer alleen het deel van het uitgezonden licht dat opgevangen wordt op de lichtsensor. $\Delta s$ is voor déze situatie de tijd tot de eierwekker afgaat. | [Eierwekker B verschijnt, dit animeren] |
| Algebraïsche afleiding (120 sec)<br>Voor inzicht in tijddilatatie is het van belang $\Delta s$ uit te drukken in $\Delta t$. Het pad van de lichtstraal kan worden aangevuld tot twee rechthoekige driehoeken. De belangrijkste stap is de Stelling van Pythagoras voor één van die driehoeken. De ene rechthoekszijde heeft lengte L. De andere rechthoekszijde is de helft van de door de raket afgelegde weg, dat wil zeggen de helft van de snelheid keer tijd: $\frac{1}{2}v\ \Delta s$.<br>De schuine zijde is, op dezelfde manier, de helft van de door het licht afgelegde weg $\frac{1}{2}c\ \Delta s$. Hier is het lichtpostulaat dus cruciaal: de snelheid van het licht is nog steeds c, en niet groter, ondanks de snelheid van de raket. En dan dus de Stelling van Pythagoras toepassen | [Vervang snelheid door c]<br>[Lijnen die de afstand van het licht aangeven uitlichten en daar $c\,\Delta s$ bij plaatsen]<br>[Kleinere driehoek uitlichten, woordformule afstand = snelheid x tijd laten staan, vervang voor $\frac{1}{2}v\ \Delta s$ woorden door symbolen, afstanden verschijnen]<br>[Animeer algebraïsche manipulaties]<br>[In animatie in twee stappen: (1) delen door 2 en keer c. (2) kwadrateren] |

| | |
|---|---|
| $\left(\frac{1}{2} c\,\Delta s\right)^2 = L^2 + \left(\frac{1}{2} v\,\Delta s\right)^2$<br>En die gebruiken we om delta s uit te drukken in delta t.<br>Uit de eerder gevonden uitdrukking voor delta t<br>Keer c<br>Gedeeld door 2<br>Vinden we een uitdrukking voor L. Vervolgens substitutie<br>Haakjes wegwerken<br>En keer 4<br>Omschrijven<br>Delen door $c^2$<br>We proberen delta s in delta t en v en c uit te drukken; daarom halen we hier delta s kwadraat buiten haakjes.<br>Beide zijden delen door de uitdrukking tussen de haakjes | |
| wortel trekken en klaar. | [Dit animeren] |
| Niets kan sneller bewegen dan licht. Dus de term onder de deelstreep zit tussen de 0 en de 1. | v < c verschijnt in beeld |
| Cruciaal punt 1 is: Dit betekent dat eitje B langer kookt dan eitje A. Niets aan de hand, zul je denken, eitje B is dus harder. | Eitjes animeren, 1 hardgekookt 1 zachtgekookt |
| Relativiteitsprincipe (90 sec)<br>En nu cruciaal punt 2. Je voelt geen verschil tussen constant bewegen of niet bewegen met een raket. Preciezer gezegd: volgens het **relativiteitsprincipe** bestaat er geen enkel experiment waarmee A of B kan vaststellen of die met constante snelheid beweegt of niet beweegt. | |
| Bijvoorbeeld het volgende simpele experiment: je stuitert een bal tegen het plafond. De bal gaat recht naar boven en weer naar beneden in je hand, | [Dit animeren] |

| | |
|---|---|
| of je nou constant beweegt of niet beweegt. Het relativiteitsprincipe stelt dat ook met complexere experimenten je dit onderscheid dus niet kunt meten. | |
| Als eitje B echt harder is, dan hebben we daarmee experimenteel vastgesteld dat B beweegt en A niet. En dat kan dus niet volgens het relativiteitprincipe. Beide eitjes moeten daarom even hard zijn! Maar hoe kan dat samengaan met de formule voor $\Delta s$? | Kruis door verkregen ongelijkheid |
| De essentiële stap in Einstein's oplossing in relativiteitstheorie is dat er twee beschrijvingen van de situatie zijn. Vanuit A gezien beweegt B en beschrijven we de tijdsintervallen met $\Delta t$ en $\Delta s$.<br>Vanuit B gezien is A juist in beweging, en beschrijven we de tijdsintervallen met andere variabelen $\Delta t'$ en $\Delta s'$. Vanwege symmetrie geldt een vergelijking van dezelfde vorm… en dus de ongelijkheid $\Delta t' > \Delta s'$. | Wisseling perspectief animeren |
| Dat beide eitjes even hard moeten zijn is nu op te lossen door te stellen dat $\Delta t = \Delta s'$ en $\Delta t' = \Delta s$.<br>Substitueren in de ongelijkheid geeft $\Delta s > \Delta s'$.<br>De conclusie hieruit is dat tijd relatief is: de voor A verstreken tijd $\Delta s$ is langer dan de volgens B verstreken tijd $\Delta s'$.<br>Dit is tijddilatatie! Wil je van de ene beschrijving naar de andere dan moet je een transformatie uitvoeren. En dat is niet geschikt voor zachtgekookte eitjes! | Beide perspectieven tegelijk visualiseren, tijdscoördinaten in de eitjes plaatsen |

## Appendix B: Post-test

**Vraag 1 / 18**

Het kijken van deze video kostte mij

| Hele, hele lage mentale inspanning | Hele lage mentale inspanning | Lage mentale inspanning | Vrij lage mentale inspanning | Noch lage noch hoge mentale inspanning | Vrij hoge mentale inspanning | Hoge mentale inspanning | Hele hoge mentale inspanning | Hele, hele hoge mentale inspanning |
|---|---|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

**Vraag 2 / 18**

Hoe mooi vond je de video eruit zien?

| | Zeer onaantrekkelijk | Onaantrekkelijk | Neutraal | Aantrekkelijk | Zeer aantrekkelijk |
|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ |

**Vraag 3 / 18**

Hoe goed kon je de video volgen, los van de algebraïsche afleidingen?

| | Zeer slecht | Slecht | Enigszins | Goed | Zeer goed |
|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ |

**Vraag 4a / 18**

In welke mate hebben de volgende aspecten bijgedragen aan het kunnen volgen van de video?

| | In zeer beperkte mate | In beperkte mate | Enigszins | In hoge mate | In zeer hoge mate | Niet van toepassing |
|---|---|---|---|---|---|---|
| Tempo | ○ | ○ | ○ | ○ | ○ | ○ |
| Opbouw / structuur | ○ | ○ | ○ | ○ | ○ | ○ |
| Stemgebruik | ○ | ○ | ○ | ○ | ○ | ○ |
| Hand bij algebra | ○ | ○ | ○ | ○ | ○ | ○ |
| Beweging bij de algebraïsche afleidingen | ○ | ○ | ○ | ○ | ○ | ○ |
| Overige bewegingen | ○ | ○ | ○ | ○ | ○ | ○ |
| Visuele vormgeving | ○ | ○ | ○ | ○ | ○ | ○ |

**Vraag 4b / 18**

Leg van één of meer van deze aspecten in één zin (per aspect) uit waarom ze bijdroegen aan het kunnen volgen van de video. Besteed niet meer dan 1-2 minuten aan deze vraag.

**Vraag 5 / 18**

In welke mate kon je de draad vasthouden bij de algebraïsche afleiding van de formule $\Delta s = \frac{\Delta t}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$ ?

| | In zeer beperkte mate | In beperkte mate | Enigszins | In hoge mate | In zeer hoge mate |
|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ |

**Vraag 6 / 18**

In welke mate zou je details van de algebraïsche afleiding kunnen reproduceren?

| | In zeer beperkte mate | In beperkte mate | Enigszins | In hoge mate | In zeer hoge mate |
|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ |

**Vraag 7 / 18**

De algebraïsche afleiding van de formule $\Delta s = \frac{\Delta t}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$ begon met het toepassen van de stelling van Pythagoras op een driehoek.

Schets de rest van deze afleiding in 4 of 5 stappen (zonder ze uit te voeren).

**Vraag 8 / 18**

Wat was de betekenis van de volgende variabelen in de video?

Δs:

Δt’:

c:

L:

**Vraag 9 / 18**

Waarom was de stelling van Pythagoras nodig voor het afleiden van de formule $\Delta s = \frac{\Delta t}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$ ?

**Vraag 10 / 18**

Zeg in één of twee zinnen hoe de formule $\Delta t = \frac{2L}{c}$ gebruikt is om te komen tot de eindformule $\Delta s = \frac{\Delta t}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$

**Vraag 11 / 18**

In de algebraïsche afleiding in de video kwam de volgende stap voorbij: $c^2 \Delta s^2 = c^2 \Delta t^2 + v^2 \Delta s^2$

Wat was de volgende stap in de algebraïsche afleiding?

**Vraag 12 / 18**

Waarom volgde uit het relativiteitspostulaat dat beide eitjes even hard moeten zijn?

**Vraag 13 / 18**

Hoe valt het feit dat de eitjes even hard zijn te rijmen met de ongelijkheid $\Delta s \geq \Delta t$ die volgde uit de vergelijking $\Delta s = \frac{\Delta t}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$ ?

**Vraag 14a / 18**

De Olympische Spelen zijn een twee weken durende sportcompetitie. Geïnteresseerde aliens bekijken de Olympische Spelen vanaf een verre planeet, die met hoge constante snelheid naar de aarde toe beweegt. De voor de aliens verstreken tijd is:

- ○ Langer dan twee weken
- ○ Gelijk aan twee weken
- ○ Korter dan twee weken

**Vraag 14b / 18**

Waarom?

**Vraag 15 / 18**

Wat was de rol van het lichtpostulaat toen de term $\frac{1}{2} c\Delta s$ werd genoemd als de afgelegde weg van het licht?

**Vraag 16 / 18**

In welke mate beïnvloedt het kunnen volgen van de algebra hoe goed je tijddilatatie begrijpt?

| | In zeer beperkte mate | In beperkte mate | Enigszins | In hoge mate | In zeer hoge mate |
|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ |

**Vraag 17 / 18**

Hoe gemotiveerd was je om de video te begrijpen?

| | Zeer ongemotiveerd | Ongemotiveerd | Neutraal | Gemotiveerd | Zeer gemotiveerd |
|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ |

**Vraag 18 / 18**

Wat was je houding tegenover de algebraïsche afleidingen in de video?

| | "Ik geloof het wel" | | "Ik probeerde op zich de algebra wel een beetje te volgen" | | "Ik checkte bij elke stap of ik het ermee eens was" |
|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ |