MASTER'S THESIS

Can additional use of video testimonials in secondary education augment the refutation effect?

't Lam, Marcel

Award date:
2023

:

Can Additional Use of Video Testimonials in Secondary Education Augment the Refutation Effect?

Marcel ‘t Lam

Master Onderwijswetenschappen, Open Universiteit

Mail address: marcarietfam@gmail.com

Course code and course name: OM9906 Masterscriptie

Name thesis supervisor: dr. Leen Catrysse

Word count: 8710

Date: June 15th, 2023

Abstract

Educating students, convincing other people or sharing a message are important interactions in many areas in our society. Especially in case of misconceptions that need to be refuted it is important to find ways to do this effectively. *Refutation texts*, especially texts with negative emotional content, have been proven powerful in realising *conceptual change*. Though quite topical, research on ways to enhance the refutation effect so far mainly focussed on text related enhancements among college undergraduates. This research had K-12 students as participants ($N = 75$) and integrated another factor that potentially could enhance the refutation effect, namely an additional video message with negative emotional content delivered by either a peer or an expert. Learning results were measured using a pre-test, post-test and delayed post-test, after which knowledge scores were compared. Results indicated that the refutation effect can be enhanced by showing a preceding additional video message delivered by a topic expert as compared to a video message by a peer. The effect we found was only applicable to students taking the immediate post-test, which was probably due to single treatment and lack of revision. Results are promising as they show that the already proven power of refutation texts can be augmented thus making them into even stronger tools that can be used to bring about conceptual change.

Keywords: refutation effect, augmentation, video messages, expertise, emotions

Contents

Can Additional Use of Video Testimonials in Secondary Education Augment the Refutation Effect?

1. Introduction

1.1 Problem and Objectives

"Education is the most powerful weapon we can use to change the world" (Mindset Network, 2003). These famous words were spoken by the former South African president Nelson Mandela. They imply that changing the world requires effort and tools and that education is the most effective way to realise a desired change. Educators in a broader sense not only comprise teachers, but also politicians, spokespersons of special interest groups and advertisers (Ode et al., 2021) as well as debaters (Najafi et al., 2016) and rhetoricians (Rapp, 2010). They all share a wish to effectively convey a message, either written or spoken, meant to convince, teach or influence people (King et al., 2012). Even an incitement to action can be part of what they want to achieve (Richards & Curran, 2002).

When aiming at getting a message across, text types that can be used are *argumentative texts* and *persuasive texts* (Diakidoy et al., 2017; Wittenberg et al., 2021). An important difference between the two is that argumentative texts generally try to convince by providing facts based on research while in persuasive texts facts and emotions are blended (*Teaching the Argumentative Standard*, 2012). A special form of text that combines argumentative and persuasive characteristics is the *refutation text* (Thacker et al., 2020). In this text a misconception is mentioned, explicitly refuted and subsequently substituted by the actual scientific viewpoint (Broughton et al., 2010). This process is also known as *conceptual change* (Gill et al., 2022). Research points out that refutation texts can aid learning and are highly efficient in refuting existing misconceptions in many contexts (Schroeder & Kucera, 2022).

Subsequent research found support for enhancing the power of refutation texts aiming at producing conceptual change (Broughton et al., 2010; Danielson et al., 2016; Gill et al., 2022; Thacker et al., 2020). This augmented activation mainly took the form of an additional written text in any form (Ecker et al., 2022; Thacker et al., 2020). A feasible method is adding additional forms of activation that personally implicate the reader and highlight the mismatch between the reader's opinion and the facts as presented in the text (Gill et al., 2022). Trevors and Kendeou (2020) found that adding negative emotional content to refutation texts was such a stimulus capable of personally implicating the reader and bringing about improved knowledge revision. Alternatively, the refutation could be voiced using positive wording. But research into the effect of refutation texts shows that texts with negative emotional content appear to be more effective than texts with positive worded refutation (Trevors & Kendeou, 2020).

Whether augmented activation using a video message would have a similar impact is something research literature does not make clear. Video messages consist of a spoken message being delivered by a sender who is visible to the receiver of the message. They could be even more powerful than written texts in personally affecting somebody and getting a message across (Wittenberg et al., 2021) and deserve more exploration (Ecker et al., 2022; Wittenberg et al., 2021). Some research has been done into the possible impact of video model characteristics on learning results, mainly focussing on the model's gender, the number of models and also skill level (Hoogerheide et al., 2018; Schunk, 1987; Schunk et al., 1987). Following the Model-Observer Similarity (MOS) hypothesis (Bandura, 1994), peer models would hypothetically be more effective than adult or expert models. However, no such conclusion could be drawn so far (Hoogerheide et al., 2018). Research also does not make clear whether knowledge or conceptual change are being impacted as a result of activation

using a video message brought by either a peer or an expert and prior to reading a refutation text.

So far, research has mostly focused on college undergraduates; K-12 students were far less chosen as participants in refutation effect research (Zengilowski et al., 2021). For them, too, gaining knowledge, orientation on and equipping for today's society and the development of an ability to think critically are essential, to mention just a few (Rothstein & Jacobsen, 2006). Following this it seems to make sense to broaden the horizon of research into refutation effect by focussing on secondary school students. Therefore, this study aims to shed more light on potential opportunities to enhance the effect of refutation texts being used in secondary education. It investigates whether different forms of video-based augmented activation have an effect on knowledge revision and transfer of knowledge among final year Havo students reading a refutation text. Whether or not there are differences in effect between video-messages with negative emotional content being delivered by a peer or an expert, will be researched.

1.2 Theoretical Framework

1.2.1 *Knowledge Revision and Conceptual Change*

Misconceptions among students are quite common and they often find their origin in scientifically inaccurate preconceptions and beliefs (Schroeder, 2016). People are often biased when deciding about what is true and they tend to make choices that are based on intuitions and their gut feel (Ecker et al., 2022). Therefore, educators trying to teach students something new that conflicts with already existing ideas should present the new information in an intelligible way that most effectively contributes to revision of false beliefs (Gill et al., 2022). The importance of the foregoing becomes even more urgent when considering that people tend to stick to their original false believes (Ecker et al., 2022; Flynn et al., 2017; Zengilowski et al., 2021), a persistence known as the *continued influence effect* (Ecker et al., 2022).

The reason why people persist in false beliefs and refuse to embrace the refutation may be that they try to reach a desired conclusion when processing information (Flynn et al., 2017) and prefer to reject information that contradicts their standing beliefs (Druckman & McGrath, 2019). According to Ecker et al. (2022) there may be other barriers to belief updating. First, one might easily forget new information, need more causal explanation or be unable to change views due to being in a suboptimal emotional state of feeling angry. Second, one might be prejudiced towards misinformation that feels familiar or not be willing to adopt a new view because that would impact somebody's worldview or identity. Finally, one might reject new information because of fallacies in logic reasoning, because of supposed credibility of the misinformation source or when believing there is consensus about somebody's standing beliefs.

Actual revision of a misconception or false belief that is already in place is also known as *belief revision* or *knowledge revision* both belonging to conceptual change (Chi, 2008). It is not an easy thing to effectuate conceptual change (Gill et al., 2022; Ozdemir & Clark, 2007). Overcoming a mismatch between existing ideas and new, corrective information requires radical changes that, once they do occur, do not take place suddenly, but rather slowly and gradually (Ozdemir & Clark, 2007). Changing previously acquired conceptions about particular topics requires learners to develop a sense of dissatisfaction with their existing ideas (Pintrich et al., 1993). Educators aiming at conceptual change should therefore take pains to present the new content intelligibly and in a way that effectuates knowledge revision best (Gill et al., 2022).

1.2.2 *Refutation Texts and the Knowledge Revision Components (KReC) Framework*

Texts known as *refutation texts* appear to have been powerful in shifting misconceptions and bringing about conceptual change (Lewandowsky et al., 2020; Thacker et al., 2020; Zengilowski et al., 2021). This type of text first identifies a misconception, then

refutes it and finally provides the reader with the correct information (Hynd, 2001; Schroeder, 2016). Refutation texts appear to be more effective than other text types when it comes to knowledge revision because they typically meet with necessary conditions for knowledge revision to take place, as there are *encoding*, *passive activation*, *co-activation*, *integration* and *competing activation* (Trevors & Kendeou, 2020). These five core principles of refutation texts have been proposed by Kendeou and O'Brien (2014) in the Knowledge Revision Components (KReC) framework they developed (Zengilowski et al., 2021). The encoding principle implies that information encoded in long-term memory is there to stay and cannot be easily removed. The passive activation principle implies that already existing misconceptions will be reactivated once related information enters somebody's working memory (Trevors & Kendeou, 2020; Zengilowski et al., 2021). The foregoing two principles lead to the assumption that, unchecked, misconceptions have a power to take their ground and keep somebody from learning and adopting new opinions (Trevors & Kendeou, 2020).

In order to overcome the aforesaid inability to get past previously adopted misconceptions, KReC next introduces three principles that can invalidate a standing misconception (Trevors & Kendeou, 2020). The co-activation principle implies that both the correct information and the misconception it refutes need to be activated simultaneously. Only then the integration principle, implying the encoding of interconnected correct and incorrect information, may become effective. This will result in a clash between the two opposing views, thus making knowledge revision possible (Broughton et al., 2010; Trevors & Kendeou, 2020; Zengilowski et al., 2021). Whether knowledge revision actually does take place depends on which of the two opposing views is activated most; the one that receives most attention will be dominant (Trevors & Kendeou, 2020). The competing activation principle deals with the latter and holds that a boosted activation of one particular view automatically implies less activation of the opposing view, ultimately resulting in a reduction

or elimination of its interference (Kendeou et al., 2014). Critical for knowledge revision to happen is dominance of corrective or refutational information that effectively undermines the misconception and effectuates adopting the new viewpoints (Zengilowski et al., 2021).

1.2.3 *Effect of Emotional Content on Retrieval and Learning*

Emotionally loaded stimuli, for example words like *massacre* or *applause*, positively affect information retrieval (Dolcos et al., 2017, as cited in Trevors & Kendeou, 2020) and have a power to capture the reader's attention (Chang et al., 2019; Larsen et al., 2008; Trevors & Kendeou, 2020). Negative emotionally loaded words, like *massacre* in the previously mentioned pair of stimuli, perform better than words with positive load (Kamp et al., 2015; Yuan et al., 2019). This negativity bias is found to result from *memory narrowing* (Levine & Edelstein, 2009; Mickley Steinmetz et al., 2014). Especially high arousal stimuli, that emotionally affect somebody to a great extent, contribute to *negativity bias* (Yuan et al., 2019). Memory narrowing as a result of exposure to arousing negative emotional stimuli like *anger* or *threat*, makes one focus on and enhances memory for only the most pressing central details (Threadgill & Gable, 2019; Trevors & Kendeou, 2020). And for a reason: when confronted with negative reinforcers, such as loss of money, all available brain resources focus on reducing the risks and increasing the chances of survival (Schechtman et al., 2010), albeit a figurative survival this time.

Through the same mechanism of memory narrowing, negative emotional stimuli have a comparable influence on learning (Trevors & Kendeou, 2020). Here also positive emotional states broaden (Fredrickson, 2013) and negative emotional states narrow focus and attention, the latter resulting in improved accuracy, better recall of main information and increased comprehension (Trevors et al., 2021; Trevors & Kendeou, 2020). Emotions regulating attention to items in somebody's working memory thus also have an effect on storage of information in one's memory (Gokce et al., 2021). Degree of emotion, specifically higher in

arousal and more extreme in valence, predicts the extent in which recall is enhanced (Kamp et al., 2015). This enhancement typically serves memorizing information needed for reaching currently active goals (Levine & Edelstein, 2009).

1.2.4 *Negative Emotional Content and KReC*

The power negative emotional content has to enhance attention and improve memory could be used when looking for a way that effectively "promotes the revision of prior false beliefs", as Gill et al. (2022) put it. Emotions are known to be influential in some of the KReC framework stages that are critical for knowledge revision, as there are *co-activation*, *integration* and *competing activation* and so they do impact knowledge revision success (Trevors et al., 2017). Negative emotions generally narrow attention and make one focus on a reduced number of stimuli, which will affect the KReC framework stages mentioned (Trevors et al., 2017). Narrowing attention, which possibly stems from a bias for negativity (Baumeister et al., 2001), leads to faster reading (Trevors & Kendeou, 2017). This might be due to more activation being directed towards the correct information, because of which reactivation of the misconception is weakened (Trevors & Kendeou, 2017). Although, according to Trevors et al. (2017), "emotions are intertwined with cognitive processes and integral to the success or failure of knowledge revision during reading", they do state that so far little research has examined *emotion* in this particular context of knowledge revision. So future research is needed (Kendeou et al., 2019; Trevors & Kendeou, 2017).

1.2.5 *Adding Video to Augment the Refutation Effect*

Videos could have a power to augment the refutation effect and therefore the combination of the two should be explored (Danielson et al., 2016). Wittenberg et al. (2021) found that people more readily believe something when it is presented to them in a video as compared to text-based presentation. However, they continue that, relatively speaking, the subsequent impact on attitude and behaviour is much smaller. On the subject of *climate*

change, video was found to be powerful in convincing people about current scientific views (Ranney & Clark, 2016) or to be even more powerful than text in that (Goldberg et al., 2019). In a survey it is concluded that, although video is generally thought to be more persuasive than other means of communication, research findings are not unambiguous about this (Wittenberg et al., 2021). According to them video could prove to be rather persuasive in real life environment setting however, only if it would attract quite some attention. That is why future research into this is necessary (Wittenberg et al., 2021). Danielson et al. (2016) and Schroeder (2016) also stress the importance of research into the combination of refutation text and video.

Video is characterised as a rich information modality consisting of graphics, audio and moving images which could easily result in an information overload when being processed by a viewer (Lang, 2000). In an apparent attempt to keep away from such an overload, viewers, when processing information, were found to be biased towards information that requires less cognitive effort and fewer cognitive resources to process (Eagly & Chaiken, 1993, as cited in Sundar et al., 2021). They therefore process video more superficially and because of that users may more readily believe in whatever message video wants to convey (Powell et al., 2018; Sundar et al., 2021). The latter is one of the reasons why video potentially is so influential.

Another reason could be the very nature of video, being an information modality that appeals to our visual senses by offering a concrete and graphical description of whatever message it wants to convey (Sundar et al., 2021). Video makes the described situation more imaginable and therefore the message becomes more believable (Hansen & Wänke, 2010) because people seem to have an innate desire to believe what they are seeing (Sundar et al., 2021). Compared to the persuasiveness of text or audio, the greater realism that video offers when communicating a message results in a greater persuasive appeal potential of video

(Sundar et al., 2021). Adding emotional content to a video that is to augment the refutation effect potentially enhances that effect even more. Anyway, added video selected to target emotion is mentioned as one of the recognised emotion induction techniques (Siedlecka & Denson, 2018).

Another aspect of the extent to which video will be more or less persuasive is whether the producer is perceivably similar to the user or has a more advanced knowledge about the subject matter being dealt with (Paek et al., 2011). In other words: is the producer a peer or an expert? According to the online Cambridge Dictionary, peers can be similar in age, in social position or in abilities. With regard to opinion formation in general, peers are found to be more influential than experts (Sundar et al., 2009). However, with regard to the context of videos meant to inform and convey a message, little is known about any possible difference in persuasiveness between the two sources mentioned (Paek et al., 2011). It was found that viewers *highly committed* to a particular issue most and for all value message quality and argument strength, while *low-involved* viewers are mainly affected by source characteristics (Chaiken, 1980). Peer producers are found to be more influential among viewers that have low issue involvement (Paek et al., 2011). In the same study Peak et al. (2011) indicate that, contrary to the way they manipulated producer type -by providing information in the survey instrument- future research should replicate their finding by highlighting producer type in the video.

1.3 Present Study

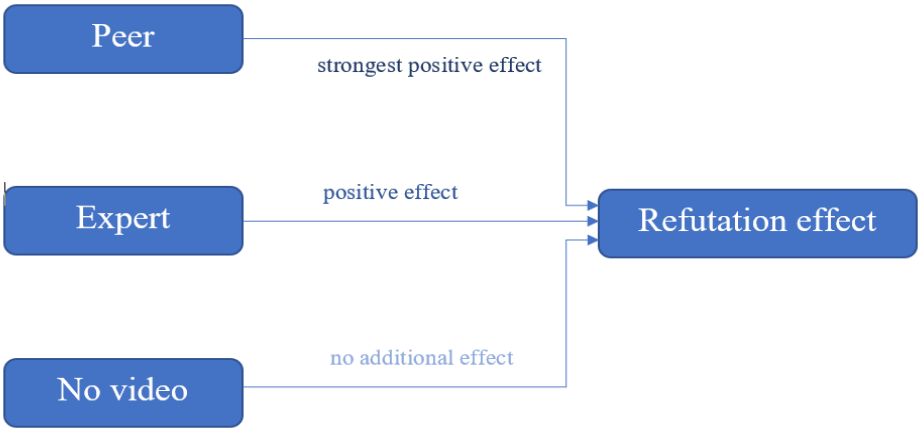
The problem dealt with in this study is how to enhance the effect of refutation texts being used in secondary education. The basics as found to be effective in previous research were left unchanged: a standard refutation text with added negative emotional content. However, contrary to previous research, in which added emotional content was mostly text-based, in this research negative emotional content was included in an additional video

testimonial which was presented to the users before they read the standard refutation text. There were two treatment conditions: a group (expert) that was shown a simple activation video with negative emotional content presented by an expert and a group (peer) that was shown an augmented activation video with negative emotional content presented by a peer. A third group (control) did not receive any treatment before reading the standard refutation text. This group was meant to check whether there was any difference in effect between conditions with and without extra activation, either simple or augmented.

Based on previous research that has been discussed above (Paek et al., 2011; Ranney & Clark, 2016; Sundar et al., 2009) first it is expected that the condition *peer* will positively enhance the refutation effect and second that the condition *expert* will do comparatively less to enhance the refutation effect. Finally, it is expected that the *control* condition will appear to be least influential because it will only show the refutation effect. The extent to which enhancement has been realised will be determined by analysing two factors that play a part in knowledge retention: immediate retention of knowledge and delayed retention of knowledge overtime. In the conceptual model in Figure 1 the expected relationship between the variables has been visualized.

Figure 1

Expected Relationship Between the Variables



The following research questions have been formulated:

(a) Does adding a video testimonial actually augment the refutation effect?

(b) How does the passing of time impact the (augmented) refutation effect?

To find an answer to these research questions, an experiment was conducted in which three groups of final year Havo level secondary school students participated. All groups were asked to read the same standard refutation text. Quantitative experimental research is considered to be a suitable research method to find answers to the aforesaid research questions because it generates numerical insights based on data retrieved from questionnaires (Creswell, 2014).

2. Method

2.1 Participants

Participants in this study are final year Havo students (senior general secondary education) who attend a comprehensive school in the North of the Netherlands. Havo students get prepared for college and their curriculum only comprises a broad range of theoretical

subjects. Not all students follow the same program; on top of the mandatory program they had to choose between a number of optional subjects. Those students who have chosen Geography -and to a much smaller extent Biology- already came across information about the (enhanced) greenhouse effect in the study program of their pre-final year. Halfway the final year this topic will get even more attention and not only in Biology and Geography lessons but also in Physics and Chemistry lessons. This is one of the reasons that we planned data collection for this experiment at the beginning of the final year, right from the start. Interference because of previously acquired knowledge will that way be minimized.

Final year Havo students are between 16-18 years old. At this particular school, about 60% of the final year Havo students are females, so for this study we planned to include 60% females and 40% males. Using G*Power 3.1 (Faul et al., 2007) sample size was calculated, using ANOVA repeated measures, within-between interaction. Here the dependent variable *knowledge* is compared between participants in different conditions and also overtime within the participants themselves: using a post-test and a delayed post-test. Default settings for alpha (α) and *Power* have been left unchanged, first because accepting a 5% chance that the null hypothesis will be wrongly rejected is considered to be reasonable and second because the minimum level of statistical Power is 80%. An average correlation between repeated measures was expected, reason why the correlation coefficient default setting has been left unchanged: $r = .50$. Following G*Power output, design of the present study, which included three conditions (peer, expert and control) and three measurements (pre-test, post-test and delayed post-test) required a minimum sample-size of 54 participants.

Before approaching the prospective participants, the school management was informed and asked for permission for the experiment to take place. Then participants were selected among students from 3 classes by providing information in class, offering students an information leaflet and asking them to join. Initially 72 students consented to participate of

which some appeared to be absent during the first leg of tests (pre-test and post-test) or ultimately when doing the delayed post-test. In the end there were 58 (21 males, 36% and 37 females, 64%) students left that had fully participated. They were between 15 and 17 years old (M=16.10, SD=.48). All participants participated on a voluntary basis and had been informed about the absence of any resulting advantages or disadvantages. Once recruited, participants were randomly assigned and evenly distributed over the three conditions (N = 24). This was realized by using the =ASELECT function in Microsoft Excel, which assigned a random number between 1 and 800 to each entry on the alphabetical list of participants. Next those numbers were used to re-organise the list, starting from the lowest and ending up with the highest number. Finally, the resulting list was split up into three equal parts that each formed one of the three conditions in this experiment.

2.2 Measuring Instruments and Materials

2.2.1 *Refutation Text*

For all three groups the intervention held a refutation text entitled “The enhanced Greenhouse Effect”. This text has been designed by Nussbaum et al. (2017) for research into effectiveness of refutation texts being used to bring about conceptual change, specifically related to the subject of climate change. The target audience of this text consisted of general education students, a category to which also the participants of this research belong. For research purposes, permission has been granted to use this text (Nussbaum et al., 2017). The original text was in English and meant to refute a number of misconceptions, as listed in Table 1 (Nussbaum et al., 2017).

Table 1

Misconceptions Refuted in the Refutation Text "The Enhanced Greenhouse Effect"

1.	There is nothing we can do about global warming
2.	Greenhouse gases are a layer of pollution or dust that trap carbon dioxide inside (like in a greenhouse)
3.	Ozone depletion causes global warming
4.	The greenhouse effect is unnatural and dangerous

Note. Adapted from Nussbaum et all. (2017).

To avoid any possible ambiguity of meaning when using the text in an experiment with Dutch students, the text has been translated into Dutch and, after consulting a Physics teacher, ºF was changed into ºC. It all resulted in a text of 1200 words. A pilot run was done among some third-year students to check whether the text is clear and understandable. The original text is accessible online as supplementary material (Nussbaum et al., 2017) and the translation into Dutch has been enclosed as Appendix A.

2.2.2 *Video Testimonials*

The two experimental groups (*peer* and *expert*) each were shown a video that was meant to augment the refutation effect. Although the two videos have different narrators, the text or message of the video is in both cases exactly the same. It is based on a speech delivered by the well-known British broadcaster and natural historian David Attenborough on the occasion of the UN Climate Action Summit, part of the 74th UN General Assembly (WWF International, 2019). To make it suitable for this current experiment, the text of the speech has been adapted and has been translated into Dutch. Characteristic of the text is its negative emotional content, which could contribute to enhancing the refutation effect (Gill et al., 2022). Words and phrases that may serve as examples of this negative emotional content

are ‘unfortunately’, ‘startling changes’ and ‘there is perhaps no more unsettling thought’. Together with other examples in the text, a general atmosphere of disturbing reality and compelling urgency has been created. Spoken text and background of the videos were the same in both experimental conditions. A pilot run was done among some third-year students to check whether the text is clear and understandable. The earlier mentioned difference in narrators was meant to test whether there is any difference in impact on the refutation effect between a video message presented by a peer and an identical message presented by an expert. In a short intro to each video it was made clear that the narrator is either a peer or an expert. Furthermore, also the narrators’ ages and the clothes they were wearing were to support the idea that they represented either a peer or an expert. The text of the video message has been provided in Appendix B. The videos will be made available on request.

2.2.3 *Knowledge Tests*

Knowledge was tested three times, in a pre-test, a post-test and a delayed post-test. They all have been developed by Nussbaum et al. (2017) and been reviewed for content validity by an expert science educator. The tests have been made available for research purposes. The pre-test consists of the original three open-ended constructed-response questions that test participants’ knowledge and opinions related to the greenhouse effect, the enhanced greenhouse effect and the role that humans play in these. After each open-ended question, students were asked to assess the degree in which they are confident that their answer given is the right answer. This was done by adding to the test a question that, in their own research, has been proposed by Cordova et al. (2014) as a “means to measure participants’ confidence in prior knowledge”: “How confident are you that this answer is correct?” As with them, also here a 7-point Likert scale was used (Cordova et al., 2014) that offers participants answer possibilities that range from 1 (*Not at all confident*) to 7 (*Very confident*). Answers given were used to ascertain whether participants do or do not differ too

much in their confidence in prior knowledge. This was done because “confidence in prior knowledge” might affect the probability of conceptual change to happen, through level of involvement or through commitment to a particular subject (Chaiken, 1980; Cordova et al., 2014).

The three content-related questions of the pre-test aimed to make written production and deeper conceptual understanding visible. The first question was as follows: “Please explain the greenhouse effect”. A scoring rubric has been provided and was also used in this research. It awards 3 or 4 score points per question, each of them linked to ‘ideas’ that together form the elements of the right answer. The maximum score one could get on this test is 11 points. For the sake of clarity, the pre-test has been translated into Dutch. The pre-test, including its scoring rubric, has been enclosed as Appendix C.

Content of the post-test and the delayed post-test were exactly the same. Both started with the same three open-ended constructed-response questions (Part 1) that were used in the pre-test, this time followed by thirteen selected-response multiple-choice questions, each offering four or five answer choices (Part 2). At the end of only the delayed post-test a non-content related question has been added that meant to check on whether participants have or have not been immersing themselves additionally in the climate change subject in the two weeks between post-test and delayed post-test. Post-test and delayed post-test aimed to make knowledge recognition visible (Nussbaum et al., 2017). An example of the questions is: “Which of the following sentences best explains the relationship between the greenhouse effect and ozone depletion?” For each rightly answered multiple-choice question participants got one point. The maximum score for the post-test -and therefore also for the delayed post-test- consisted of two separate scores: 11 (Part 1) and 13 (Part 2). These scores were interpreted separately, because they are two different types of scores that were scaled in different ways and that reflect competencies that are different (Nussbaum et al., 2017). For the

sake of clarity, the post-test has been translated into Dutch. Furthermore, some minor changes have been made in the wording of the instructions. Finally, some changes have been made to a small number of answer choices to prevent participants from being able to exclude some answers just on the basis of logical reasoning and comparing between answer choices. Appendix D contains the (delayed) post-test in which, for this study, the right answers have been bolded.

2.3 Procedure

In three neighbouring classrooms, one for each group, participants were welcomed by a teacher who would normally be teaching one of the participating classes. The participating students had been informed about the classroom they should go to by means of an update of their personal schedule. A copy of the informed consent form of each individual participant been put on a table indicated the student's place in the classroom. Technical details about the current research were shared with the participants. Then they were reminded of the voluntary basis on which they consented to participate in this experiment, the promise of the data being anonymized and the fact that they were free to stop participating at any time. After that all participants completed a demographic survey that inquired after age and sex, followed by a pre-test meant to measure deeper conceptual understanding and also their level of confidence when answering the questions. Next the peer group was first shown a video testimonial presented by a peer, then read the standard refutation text and finally completed a knowledge assignment. Participants in the expert group followed the same procedure with one difference: their video testimonial was presented by an expert. Participants in the control group were not shown a video testimonial; they only read the refutation text and then completed the knowledge assignment. The foregoing design aimed to measure the dependent variable *knowledge* in two experimental conditions and a control condition. Two weeks after the first session, 58 participants took the delayed post-test for which they completed the knowledge

assignment again, this time with changed multiple-choice positions. Retention of the same variable *knowledge* was measured and also a check after interference of information obtained from other sources was included. After completing all parts of this research and analysing the data, the participants were informed about the results.

2.4 Data Analysis

First the collected data were checked on outliers and normality. Checking on outliers was done by first visually inspecting data and applying the interquartile range (IQR) rule. For other checks and analyses SPSS statistical analysis software Version 27 was used. Normality was checked by using SPSS for histograms, QQ plots and measuring kurtosis and skewness. Also z-scores for skewness and kurtosis were used to check whether these scores were between the threshold values -1.96 and 1.96. To decide on whether or not the distribution of the dependant variable scores on the different tests was normal, a Shapiro-Wilk test was performed for each test.

Furthermore, the pre-test (t0) results were analysed using an ANOVA to find out whether there were any already existing differences between the three conditions, First homogeneity was checked by interpreting Levene's test. Where this test indicated that the assumption of homogeneity had been violated, Welch's *F* was used because Field (2018) strongly recommends this test for it being more robust than Kruskal-Wallis test and applicable in all situations. If there were differences in *knowledge* or *confidence in prior knowledge*, including a covariate when testing the hypotheses would be considered.

Regarding our first research question, ANOVAs were used to analyse whether the condition has an effect on learning outcome. There is one independent variable and one dependent variable. The independent variable (or factor) *augmentation* describes the particular type of video that was used to augment the refutation effect brought about by reading the refutation text. It has 3 levels: peer, expert or control (no video). *Augmentation* is

a between-subjects factor. The dependent variable *knowledge* was measured on interval level. For the first hypothesis it was checked whether there is a significant effect of augmentation on knowledge, comparing the control group with the video groups. For the second hypothesis it was checked whether there is any difference of effect between the video conditions peer and expert. For this comparison, Bonferroni adjusted multiple comparison is reported. The aforesaid checks were performed for the post-test (t1) and delayed post-test (t2) using the results on MC and open questions separately.

Retention of knowledge was tested by running a repeated measures ANOVA. The dependent variable is knowledge and the independent variables are condition and time (pre-test, post-test and delayed post-test). Because the dataset consisted of three conditions, sphericity could be an issue. Following Field's suggestion (2018) to ignore Mauchly's *W* when testing for sphericity and routinely apply one of the corrections provided in the same output, either Greenhouse-Geisser's estimate or Huynh-Feldt's estimate was used, depending on the value of Epsilon $\hat{\varepsilon}$ (Huynh-Feldt was only used when $\hat{\varepsilon} > 0.75$). Post-hoc tests with Bonferroni adjustments were used to compare differences between time points.

3. Results

3.1 Descriptive statistics

Table 2 presents an overview of the descriptive statistics of the knowledge measures. Appendix E presents an overview of the results. The distribution was not normally distributed for the "Delayed Post-test Open" variable ($W = .949$, $p = .016$). For the distributions of all other variables Shapiro-Wilk did not yield any proof of non-normality.

Table 2

Descriptive Statistics of the Pre-test and Post-test Measures per Condition

	CONTROL ($N = 20$)		EXPERT ($N = 18$)		PEER ($N = 20$)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Pre-test Confidence (max = 7)	3.70	1.42	3.69	1.51	3.89	1.15
Pre-test Open (max = 11)	2.75	1.42	3.00	1.14	2.93	1.89
Post-test Open (max = 11)	3.85	1.29	4.42	1.55	3.45	1.51
Post-test MC (max = 13)	9.35	2.23	9.78	.94	8.55	1.61
Delayed Post-test Open (11)	3.60	1.17	3.47	1.55	2.75	1.48
Delayed Post-test MC (13)	8.65	2.23	8.50	2.33	8.15	1.90

Following Field (2018) a significant test should not be much of a problem if after visual inspection the plotted data would appear close to normal. Appendix F provides an overview of the distribution of each variable. For the delayed post-test open, scores are only slightly different from normal scores. Additionally, scores for skew and kurtosis were converted to z-scores: the z-score for skewness of the delayed post-test open score was found to be .80, indicating that the distribution was just slightly right-skewed. For the same test the z-score for kurtosis was found to be .15, meaning that the distribution of the data was pretty much close to normal. These scores are between the threshold values of -1.96 and 1.96. Indeed, none of z-scores for skewness and kurtosis of the other tests were below or above the threshold values of -1.96 and 1.96 respectively. When zooming in and calculating z-scores for the delayed post-test open again, but this time per condition, it appears that the control condition scores 2.18 on skewness) and 2.73 on kurtosis. Consulting the associated Stem-and-Leaf plot made clear there was one extreme score that, however, was not out of range when compared to the highest scores of other conditions or scores scored in an earlier stage (on

either pre-test or post-test open). So it was concluded that there was no need to manipulate the dataset: no respondents were removed from the data set.

Furthermore, the three experimental conditions did not differ significantly in *prior knowledge*: $F(2, 55) = 0.14$, $p = .873$, nor was there any significant difference in *confidence in prior knowledge*: $F(2, 55) = 0.12$, $p = .884$. Therefore, the option of including *prior knowledge* or *confidence in prior knowledge* as a covariate in our analysis was not effectuated.

3.2 Does Adding a Video Testimonial Actually Augment the Refutation Effect?

To investigate whether there is an effect of augmentation on knowledge an ANOVA was conducted, first evaluating post-test scores on the open-ended constructed-response questions. When checking homogeneity of variances, Levene's statistic for post-test open scores indicated that variances for the three conditions did not differ significantly ($p = .503$).

No significant effect of augmentation was found on the overall post-test open scores, $F(2, 55) = 2.12$, $p = .130$, $\eta^2 = 0.07$. However, planned contrasts revealed that the mean score on open questions of the expert group was significantly higher than the score of the peer group: $t(55) = -2.05$, $p = .045$, $d = -0.67$. Other contrasts did not yield any significant result.

A second ANOVA was conducted, evaluating the effect of augmentation on the post-test scores of the multiple-choice questions. When checking homogeneity of variances, Levene's statistic for post-test MC indicated that variances for the three conditions were significantly different, $F(2, 55) = 7.32$, $p = .002$, and therefore it was decided to report Welch's statistic.

Welch's test revealed that there was a significant effect of augmentation on MC questions, $F(2, 34.00) = 4.16$, $p = .024$, $\eta^2 = 0.09$, indicating that means are different. As Levene's statistic indicated homogeneity of variances was violated ($p < .05$) the "equal variances not assumed" section of planned contrasts was consulted, resulting in only one

significant statistic that indicates means for MC scores of the peer group and the expert group differ: $t(55) = -2.91$, $p = .007$, $d = -0.72$, more specifically indicating *expert* is higher than *peer*.

Quite similarly, also for the delayed post-tests ANOVAs were conducted, first for open-ended questions. Levene's statistic did not yield any significant result ($p = .32$). Results of an ANOVA did not reveal a significant delayed effect of augmentation on the open questions, $F(2, 55) = 2.11$, $p = .131$, $\eta^2 = 0.07$.

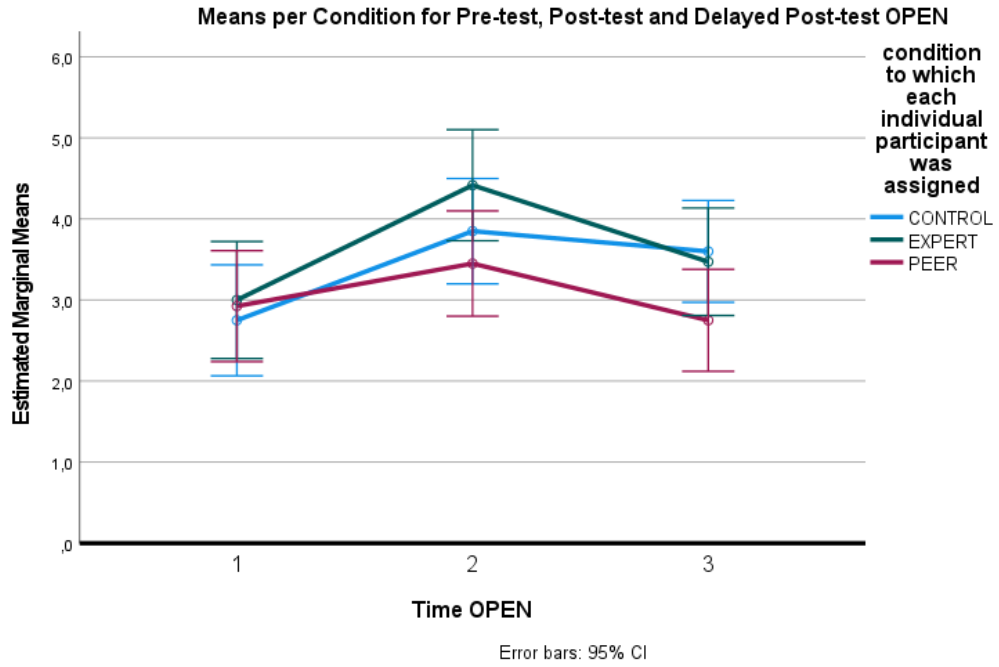
Also here a second ANOVA was conducted, evaluating the effect of augmentation on the delayed post-test scores of multiple-choice questions. Levene's test for equality of variances did not yield any significant result ($p = .76$). The results of an ANOVA did not reveal a significant delayed effect of augmentation on MC questions, $F(2, 55) = .282$, $p = .755$, $\eta^2 = 0.01$.

3.3 Retention of Knowledge Over Time

To investigate whether and to which extent knowledge was retained, a repeated measures mixed ANOVA was performed to compare the effect of passing of time on knowledge retention. Mauchly's test indicated that the Huynh-Feldt estimate of the departure from sphericity was $\varepsilon = 1.00$, so there is no violation of sphericity. Knowledge retention was significantly affected by the passing of time, as followed from evaluating first pre-test, post-test and delayed post-test scores on open questions, $F(2, 110) = 11.20$, $p = <.001$, $\eta_p^2 = 0.17$. Bonferroni adjusted pairwise comparisons between three levels first show that the significant main effect reflects a significant difference ($p < .001$) between Levels 1 and 2, post-test open being higher than pre-test open. Second, they also show a significant difference ($p = .007$) between Levels 2 and 3, post-test open being higher than delayed post-test open. Figure 2 provides a graphic overview of the means per condition for the open-ended tests.

Figure 2

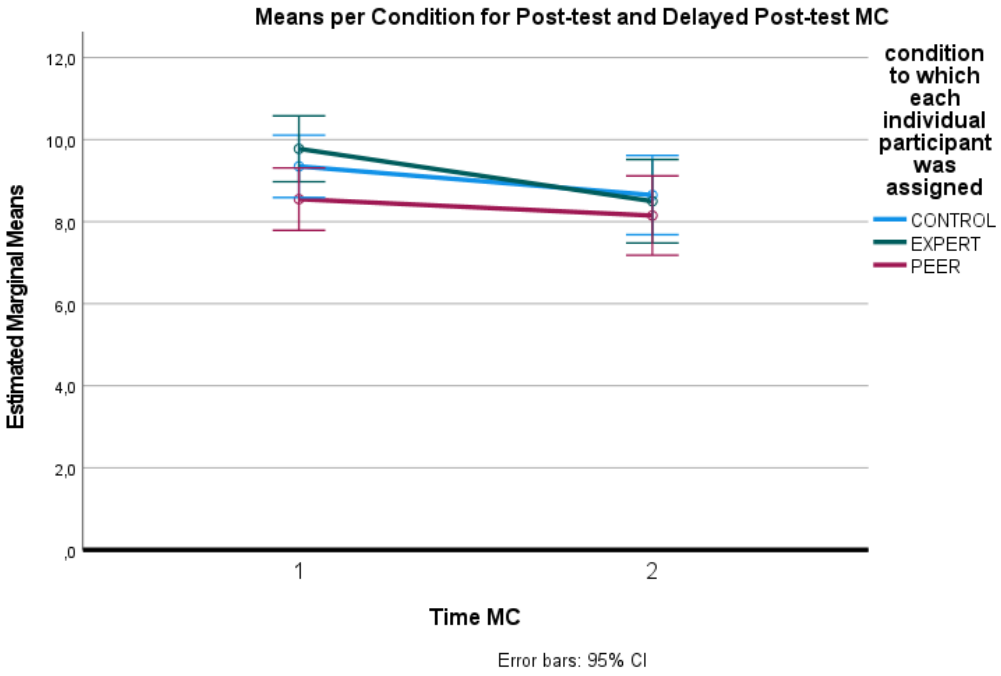
Means for Three Tests with Open-ended Questions



Next knowledge retention over time between post-test MC and delayed post-test MC was evaluated by running a second repeated measures ANOVA. Mauchly's test indicated that the Huynh-Feldt estimate of the departure from sphericity was $\varepsilon = 1.00$, so there is no violation of sphericity. Knowledge retention was significantly affected by the passing of time, as followed from evaluating overall post-test and delayed post-test scores on MC questions, $F(1, 55) = 9.06$, $p = .004$, $\eta_p^2 = 0.14$, post-test MC scores being higher than delayed post-test MC scores. As appears from Figure 3, for all three conditions scores for the MC tests follow the same pattern.

Figure 3

Means for Two Tests with MC Questions



4. Discussion

This study aimed at investigating whether additional use of video testimonials would augment the refutation effect in secondary education. What has been subject of earlier studies was mainly augmented activation on the basis of additional text (Ecker et al., 2022; Thacker et al., 2020). Studies into the effect of non-textual forms of augmented activation, specifically the use of additional video messages either brought by a peer or an expert, were scarce and deserve more attention (Ecker et al., 2022).

4.1 Influence of Adding a Video Testimonial on the Post Test

Regarding our first research question "Does adding a video testimonial actually augment the refutation effect?" results demonstrate that the expert group scores significantly higher than the peer group on the open and multiple-choice questions in the immediate post-test. So here the effects on the MC questions and open questions coincide, which seems to

enhance support for the idea that in this research expert group members' attention has somehow been raised to a higher level by the video message that was shared with them. This could indicate that participants' test results can be influenced by showing them a video message before having them read the refutation text. The aforesaid is in line with previous research in which it was suggested that video messages could have a power to enhance the refutation effect (Goldberg et al., 2019; Ranney & Clark, 2016; Wittenberg et al., 2021).

In showing that participants who saw the expert video augmentation did better than those who saw the peer video, this research yielded a result that is opposite to what we expected based on previous research. There it was found that related to opinion formation in general, a perceivably similar peer in a talking head video is more influential than somebody who is being presented as an issue expert (Paek & Gunther, 2007; Paek et al., 2011). Our seemingly contradictory finding could be accounted for by possible differences in which the video messages either brought by 'the peer' or 'the expert' were perceived by their respective audiences. The expert could have been more convincing, resulting in a high-involved audience, which according to previous research is more focussed on communicator's arguments and persuasive message quality (Chaiken, 1980). In contrast, a lower involved audience primarily focusses on a communicator's personal characteristics (Chaiken, 1980; Paek et al., 2011) consequently picking up less of the persuasive message.

4.2 Influence of Adding a Video Testimonial on the Delayed Post Test

Next to evaluating the results on the post-test questions also the results on the delayed post-test questions were evaluated. Both results on MC questions and results on open questions did not reveal any significant delayed effect of augmentation. Where the immediate post-test results showed a significant initial effect of augmentation among those participants who saw the expert-video, two weeks later when taking the delayed post-test this difference was not present anymore. This could be because the participants had forgotten the new

information, because people tend to easily forget new information (Ecker et al., 2022). After all, effectuating conceptual change is not an easy thing (Gill et al., 2022; Ozdemir & Clark, 2007) and it requires radical changes that occur slowly and gradually (Ozdemir & Clark, 2007). Learners should first become dissatisfied with their existing ideas before being able to adopt new information (Pintrich et al., 1993).

Secondly, we did not investigate whether the participants always tried to provide answers that were in line with what they had been reading in the text. Especially when taking the delayed post-test the participants could have fallen back on their personal, potentially dissenting opinion to make up for forgetting that which was being said in the text. This would not be surprising at all because people tend to stick to their original opinions (Ecker et al., 2022; Flynn et al., 2017; Zengilowski et al., 2021). Following the *competing activation principle* of the KReC framework (Kendeou et al., 2014) *knowledge revision* requires the provided corrective information to become dominant thus eliminating interference of any opposing view. Just once being exposed to information that refutes already existing misconceptions may not have been enough for encoding the information learned or for substituting misconceptions. After all, following the coding principle, information stored in the long-term memory is there to stay and cannot be easily removed (Kendeou et al., 2014).

Alternatively, initial acceptance of corrective information provided in the refutation text may have been fuelled by emotions evoked by the compelling video message that preceded reading the refutation text. Previous research found that added video can be a suitable means to target emotion and affect the audience (Siedlecka & Denson, 2018). Emotions in general impact the knowledge revision process either favourable or unfavourable (Trevors et al., 2017). Negative emotional content is held responsible for narrowing the attention and directing the reader's attention towards the correct information (Trevors & Kendeou, 2017). Two weeks after the post-test the participants took the delayed post-test

which came without preceding video message. So any resulting emotion could have been lost long ago and was not being reinforced, which may have resulted in a lower score on the delayed post-test.

4.3 Retention of Knowledge Over Time

Concerning the second research question "How does the passing of time impact the (augmented) refutation effect?" results demonstrate that the immediate post-test scores were significantly higher than the pre-test and delayed post-test scores. For as long as newly acquired information is not being encoded in the long-term memory, initial gains may result into losses later on because the new information simply did not receive enough attention to become dominant (Trevors & Kendeou, 2020).

Repeated measures results are highly prone to order effects as a result of exposing participants to multiple treatments (Creswell, 2014; Danielson et al., 2016). In such a case test scores may change, either becoming better or showing a decline. In this research delayed post-test scores had dropped which could have resulted from what we proposed above, but which alternatively could also have been a result of fatigue. Anyway, evaluation of the answers given made clear that answers on the delayed post-test questions were in general somewhat shorter and showed reduced quality.

4.4 Limitations and Suggestions for Future Research

This research has a number of limitations. First, no difference was made between participants who were less or alternatively more motivated about reading or about the topic of climate change. Motivation is known to positively affect reading comprehension (Ahmadi, 2017) and fascination for a particular subject positively contributes to an advanced long-term knowledge gain (Schneiderhan-Opel & Bogner, 2020). It also does matter whether a participant is either high- or low-involved because of a difference in focus: either on content or on a communicator's personal characteristics respectively (Chaiken, 1980). A focus

primarily on content may lead to higher levels of alertness, especially when feeling the topic being important. If someone is affected when listening to a compelling message that contains negative emotional stimuli, it will narrow that person's focus and attention (Trevors & Kendeou, 2020). This implies that accuracy will improve and recall of main information will increase (Trevors et al., 2021; Trevors & Kendeou, 2020). The degree in which somebody has been touched matters, because that defines the extent in which enhancement of recall will be established (Kamp et al., 2015). Future research should take into account reader motivation thus trying to find out whether the findings of this study hold for both motivated and less motivated readers.

Second, regarding the videos that were used in this study, differences between the expert and peer boiled down on personal and appearance characteristics (age, looks, tone of voice, clothes) and attributed knowledge level or expertise (peer or expert). All of these belong to nonverbal communication (McDermott, 2023). The actual meanings that are communicated through nonverbal communication depend on the listeners' subjective interpretations (Todorović et al., 2017). Whether and in what way these nonverbal communication cues have been impacting the power with which the message was brought by peer or expert and how this possibly impacted transfer of knowledge and knowledge retention is something that future research should focus on. This could be done by including questions that tap into whether or not participants sympathize with and/or relate to the communicator in the video.

Additionally, future research could focus on participants' opinions about the persuasiveness of the speaker thus checking on whether they consider the speaker to be either an expert or a peer. This because it is not clear whether the participants actually regarded the peer as 'a peer' and the expert as 'an expert' only based on the fact that we branded them as such. According to the online Cambridge Dictionary, peers can be similar in various ways,

among which are age and abilities. We have not been researching which aspect (if any at all) was most important to the participants. The expert in our research might have been somebody the participants identified with more as compared to identification with the peer we had in mind for them. After all, persuasiveness depends on whether the speaker is perceivably similar to the listener (Paek et al., 2011) as subjectively determined by each individual listener.

Also, with regard to the materials that were used, it should be noted that, when taking the tests, the participants were asked to answer open questions without any clues or hints being provided. They may not have realised that the scoring rubric identified multiple aspects for the model answer to a question, each of them linked to a partial score and all of them necessary for a full score. They could have been easily satisfied after having answered the question without having been trying to give a full, multi-aspect answer. Future research could provide participants with some cues or with a pre-structured worksheet which would make clear to them how many different aspects the model answer should consist of.

Finally, this study focused on determining whether video testimonials can be used to effectively augment the refutation effect. For this, participants first watched a video of a peer or an expert and then read the refutation text or only read the refutation text in the control condition within a set time without offering them any possibility to take notes or revise. This implies that they were using a learning strategy, namely attentive reading and limited re-reading, that is considered to be less effective (Dunlosky et al., 2013). This is because for new information to be stored in the long-term memory, revising it -or *spaced learning*- is crucial (Smolen et al., 2016). The delayed post-test being done two weeks after the post-test, a decline in main scores was to be expected, due to not revising newly-acquired information (Murre & Dros, 2015; Underwood, 1957). Future research could replicate this research while

alternatively asking participants to revise, thus assessing whether this would affect the outcomes.

4.5 Conclusion

This study provides information about the potential of video testimonials being used to augment the refutation effect in secondary education. Though future research is needed, we found that the refutation effect can be enhanced by additionally showing a video message delivered by a topic expert. Knowing about this contributes to learning more about the effectiveness of refutation texts and bridging the gap between sender and receiver when aiming at conceptual change. The effect we found was only applicable to students taking the immediate post-test, which was probably due to single treatment and lack of revision. Future research should try and find ways to make it a lasting effect and thus contribute to getting to know more about lasting conceptual change and knowledge retention.

References

Ahmadi, M. R. (2017). The impact of motivation on reading comprehension. *International Journal of Research in English Education*, *2*(1), 1-7. https://doi.org/10.18869/acadpub.ijree.2.1.1

Bandura, A. (1994). Self-efficacy. In V.S. Ramachaudran (Ed.), *Encyclopedia of Human Behavior*, (Vol. 4, pp. 71-81). Academic Press.

Baumeister, R., Bratslavsky, E., Finkenauer, C., & Vohs, K. (2001). Bad is stronger than good. *Review of General Psychology*, *5*. https://doi.org/10.1037/1089-2680.5.4.323

Broughton, S., Sinatra, G., & Reynolds, R. (2010). The nature of the refutation text effect: An investigation of attention allocation. *Journal of Educational Research*, *103*(6), 407-423. https://doi.org/10.1080/00220670903383101

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*(5), 752-766. https://doi.org/10.1037/0022-3514.39.5.752

Chang, E. P., Ecker, U. K. H., & Page, A. C. (2019). Not wallowing in misery – Retractions of negative misinformation are effective in depressive rumination. *Cognition & Emotion*, *33*(5), 991-1005. https://doi.org/10.1080/02699931.2018.1533808

Chi, M. T. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 61-82). https://www.public.asu.edu/~mtchi/papers/Chi_concpetualchangechapter.pdf

Cordova, J. R., Sinatra, G. M., Jones, S. H., Taasoobshirazi, G., & Lombardi, D. (2014). Confidence in prior knowledge, self-efficacy, interest and prior knowledge: Influences on conceptual change. *Contemporary Educational Psychology*, *39*(2), 164-174. https://doi.org/10.1016/j.cedpsych.2014.03.006

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Danielson, R. W., Sinatra, G. M., & Kendeou, P. (2016). Augmenting the refutation text effect with analogies and graphics. *Discourse Processes*, *53*(5-6), 392-414. https://doi.org/10.1080/0163853X.2016.1166334

Diakidoy, I.-A. N., Ioannou, M. C., & Christodoulou, S. A. (2017). Reading argumentative texts: Comprehension and evaluation goals and outcomes. *Reading and Writing*, *30*(9), 1869-1890. https://doi.org/10.1007/s11145-017-9757-x

Druckman, J. N., & McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change*, *9*(2), 111-119. https://doi.org/10.1038/s41558-018-0360-1

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4-58. https://doi.org/10.1177/1529100612453266

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E.K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, *1*(1), 13-29. https://doi.org/10.1038/s44159-021-00006-y

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Field, A. P. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Flynn, D. J., Nyhan, B., & Reifler, J. (2017). The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Political Psychology*, *38*, 127-150. https://doi.org/10.1111/pops.12394

Fredrickson, B. L. (2013). Chapter one - Positive emotions broaden and build. In P. Devine & A. Plant (Eds.), *Advances in Experimental Social Psychology* (Vol. 47, pp. 1-53). Academic Press. https://doi.org/10.1016/B978-0-12-407236-7.00001-2

Gill, M. G., Trevors, G., Greene, J. A., & Algina, J. (2022). Don't take it personally? The role of personal relevance in conceptual change. *The Journal of Experimental Education*, *90*(1), 1-22. https://doi.org/10.1080/00220973.2020.1754152

Gokce, A., Zinchenko, A., Annac, E., Conci, M., & Geyer, T. (2021). Affective modulation of working memory maintenance: The role of positive and negative emotions. *Advances in Cognitive Psychology*, *17*(2), 107-116. https://doi.org/10.5709/acp-0321-7

Goldberg, M. H., van der Linden, S., Ballew, M. T., Rosenthal, S. A., Gustafson, A., & Leiserowitz, A. (2019). The experience of consensus: Video as an effective medium to communicate scientific agreement on climate change. *Science Communication*, *41*(5), 659-673. https://doi.org/10.1177/1075547019874361

Hansen, J., & Wänke, M. (2010). Truth from language and truth from fit: The impact of linguistic concreteness and level of construal on subjective truth. *Personality and Social Psychology Bulletin*, *36*(11), 1576-1588. https://doi.org/10.1177/0146167210386238

Hoogerheide, V., van Wermeskerken, M., van Nassau, H., & van Gog, T. (2018). Model-observer similarity and task-appropriateness in learning from video modeling examples: Do model and student gender affect test performance, self-efficacy, and perceived competence? *Computers in Human Behavior*, *89*, 457-464. https://doi.org/10.1016/j.chb.2017.11.012

Hynd, C. R. (2001). Refutational texts and the change process. *International Journal of Educational Research*, *35*(7), 699-714. https://doi.org/https://doi.org/10.1016/S0883-0355(02)00010-1

Jarodzka, H. M. (n.d.). *Online Leren en Instructie*. Open Universiteit. Retrieved June 11, 2022, from https://www.ou.nl/onderzoek-onderwijswetenschappen-online-leren-en-instructie

Kamp, S. M., Potts, G. F., & Donchin, E. (2015). On the roles of distinctiveness and semantic expectancies in episodic encoding of emotional words. *Psychophysiology*, *52*(12), 1599-1609. https://login.ezproxy.elib11.ub.unimaas.nl/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eoah&AN=37740228&site=ehost-live&scope=site

Kendeou, P., Kendeou, P., Robinson, D. H., & McCrudden, M. T. (2019). *Misinformation and fake news in education*. Information Age Publishing, Inc.

Kendeou, P., Walsh, E. K., Smith, E. R., & O'Brien, E. J. (2014). Knowledge revision processes in refutation texts. *Discourse Processes*, *51*(5-6), 374-397. https://doi.org/10.1080/0163853X.2014.913961

King, R., Erickson, C., & Sebranek, J. (2012). *Inquire: A student handbook for 21st Century learning*. Thoughtful Learning.

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, *50*(1), 46-70. https://doi.org/10.1111/j.1460-2466.2000.tb02833.x

Larsen, R. J., Mercer, K. A., Balota, D. A., & Strube, M. J. (2008). Not all negative words slow down lexical decision and naming speed: Importance of word arousal. *Emotion*, *8*, 445-52. https://doi.org/10.1037/1528-3542.8.4.445

Levine, L. J., & Edelstein, R. S. (2009). Emotion and memory narrowing: A review and goal-relevance approach. *Cognition and Emotion*, *23*(5), 833-875. https://doi.org/10.1080/02699930902738863

Lewandowsky, S., Cook, J., & Lombardi, D. (2020). *Debunking Handbook 2020.* Databrary. https://doi.org/10.17910/B7.1182

McDermott, N. (2023, January 19). 6 Types of nonverbal communication, what they mean and how to spot them. *Forbes Health*. https://www.forbes.com/health/mind/nonverbal-communication-types

Mickley Steinmetz, K. R., Waring, J. D., & Kensinger, E. A. (2014). The effect of divided attention on emotion-induced memory narrowing. *Cognition & Emotion*, *28*(5), 881-892. https://doi.org/10.1080/02699931.2013.858616

Mindset Network. (2003, July 16). *Launch of Mindset Network - Nelson Mandela Speech* [Video]. YouTube. https://www.youtube.com/watch?v=e6X5rCyQqn0

Minitab Blog Editor. (n.d.). Repeated Measures Designs: Benefits, Challenges, and an ANOVA Example. https://blog.minitab.com/en/adventures-in-statistics-2/repeated-measures-designs-benefits-challenges-and-an-anova-example

Murre, J. M. J., & Dros, J. (2015). Replication and Analysis of Ebbinghaus' Forgetting Curve. *PLoS ONE*, *10*(7), e0120644. https://doi.org/10.1371/journal.pone.0120644

Najafi, M., Motaghi, Z., Nasrabadi, H. B., & Heshi, K. N. (2016). "Debate" learning method and its implications for the formal education system. *Educational Research and Reviews*, *11*(6), 211-218. https://login.ezproxy.elib11.ub.unimaas.nl/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1094430&site=ehost-live&scope=site

Nussbaum, E. M., Cordova, J. R., & Rehmat, A. P. (2017). Refutation texts for effective climate change education. *Journal of Geoscience Education*, *65*(1), 23-34. https://doi.org/10.5408/15-109.1

Ode, D. A., Chauhan, J. D., & Sharma, A. R. (2021). *Concepts of education*. Red'Shine Publication Private Limited. https://books.google.nl/books?id=zgQVEAAAQBAJ

Ozdemir, G., & Clark, D. (2007). An overview of conceptual change theories. *Eurasia Journal of Mathematics, Science and Technology Education*, *3*, 351-361. https://doi.org/10.12973/ejmste/75414

Paek, H.-J., & Gunther, A. C. (2007). How peer proximity moderates indirect media influence on adolescent smoking. *Communication Research*, *34*(4), 407-432. https://doi.org/10.1177/0093650207302785

Paek, H.-J., Hove, T., Ju Jeong, H., & Kim, M. (2011). Peer or expert? *International Journal of Advertising*, *30*(1), 161-188. https://doi.org/10.2501/IJA-30-1-161-188

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, *63*(2), 167-199. https://doi.org/10.3102/00346543063002167

Powell, T. E., Boomgaarden, H. G., De Swert, K., & de Vreese, C. H. (2018). Video killed the news article? Comparing multimodal framing effects in news videos and articles. *Journal of Broadcasting & Electronic Media*, *62*(4), 578-596. https://doi.org/10.1080/08838151.2018.1483935

Ranney, M. A., & Clark, D. (2016). Climate change conceptual change: Scientific information can transform attitudes. *Topics in Cognitive Science*, *8*(1), 49-75. https://doi.org/10.1111/tops.12187

Rapp, C. (2010, February 1). *Aristotle's rhetoric*. Retrieved February 23, 2022, from https://plato.stanford.edu/archives/spr2010/entries/aristotle-rhetoric/#Bib.

Richards, J., & Curran, C. (2002). Oracles on "advertising": Searching for a definition. *Journal of Advertising*, *31*(2), 63-77. https://login.ezproxy.elib11.ub.unimaas.nl/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eoah&AN=30484465&site=ehost-live&scope=site

Rothstein, R., & Jacobsen, R. (2006). The goals of education. *Phi Delta Kappan*, *88*, 264-272. https://doi.org/10.1177/003172170608800405

Schechtman, E., Laufer, O., & Paz, R. (2010). Negative valence widens generalization of learning. *Journal of Neuroscience*, *30*(31), 10460-10464. https://doi.org/10.1523/jneurosci.2377-10.2010

Schneiderhan-Opel, J., & Bogner, F. X. (2020). How fascination for biology is associated with students' learning in a biodiversity citizen science project. *Studies in Educational Evaluation*, *66*, 100892. https://doi.org/10.1016/j.stueduc.2020.100892

Schroeder, N. L. (2016). A preliminary investigation of the influences of refutation text and instructional design. *Technology, Knowledge and Learning*, *21*, 325-340. https://doi.org/10.1007/s10758-016-9278-8

Schroeder, N. L., & Kucera, A. C. (2022). Refutation text facilitates learning: A meta-analysis of between-subjects experiments. *Educational Psychology Review*, *34*, 957-987. https://doi.org/10.1007/s10648-021-09656-z

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research*, *57*(2), 149-174. https://doi.org/10.3102/00346543057002149

Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer-model attributes and children's achievement behaviors. *Journal of Educational Psychology*, *79*(1), 54-61. https://doi.org/10.1037/0022-0663.79.1.54

Siedlecka, E., & Denson, T. F. (2018). Experimental methods for inducing basic emotions: A qualitative review. *Emotion Review*, *11*(1), 87-97. https://doi.org/10.1177/1754073917749016

Smolen, P., Zhang, Y., & Byrne, J. H. (2016). The right time to learn: Mechanisms and optimization of spaced learning. *Nature Reviews Neuroscience*, *17*(2), 77-88. https://doi.org/10.1038/nrn.2015.18

Sundar, S. S., Molina, M. D., & Cho, E. (2021). Seeing is believing: Is video modality more powerful in spreading fake news via online messaging apps? *Journal of Computer-Mediated Communication*, *26*(6), 301-319. https://doi.org/10.1093/jcmc/zmab010

Sundar, S. S., Xu, Q., & Oeldorf-Hirsch, A. (2009). Authority vs. peer: How interface cues influence users. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems* (pp. 4231–4236). https://doi.org/10.1145/1520340.1520645

Teaching the Argumentative Standard. (2012). Retrieved February 21, 2022, from https://www.smekenseducation.com/wp-content/uploads/bsk-pdf-manager/2020/06/Op_v_Pers_v_Arg.pdf

Thacker, I., Sinatra, G. M., Muis, K. R., Danielson, R. W., Pekrun, R., Winne, P. H., & Chevrier, M. (2020). Using persuasive refutation texts to prompt attitudinal and conceptual change. *Journal of Educational Psychology*, *112*(6), 1085-1099. http://dx.doi.org/10.1037/edu0000434

Threadgill, A. H., & Gable, P. A. (2019). Negative affect varying in motivational intensity influences scope of memory. *Cognition & Emotion*, *33*(2), 332-345. https://doi.org/10.1080/02699931.2018.1451306

Tippett, C. D. (2010). Refutation text in science education: A review of two decades of research. *International Journal of Science and Mathematics Education*, *8*(6), 951-970. https://doi.org/10.1007/s10763-010-9203-x

Todorović, T., Pavko-Čuden, A., Košak, K., & Toporišič, T. (2017). Language of dressing as communication system and its functions – Roman Jakobson's linguistic method. *Fibres and Textiles in Eastern Europe*, *25*, 127. https://doi.org/10.5604/01.3001.0010.4639

Trevors, G., Bohn-Gettler, C., & Kendeou, P. (2021). The effects of experimentally induced emotions on revising common vaccine misconceptions. *Quarterly Journal of Experimental Psychology*, *74*(11), 1966-1980. https://doi.org/10.1177/17470218211017840

Trevors, G., & Kendeou, P. (2020). The effects of positive and negative emotional text content on knowledge revision. *Quarterly Journal of Experimental Psychology*, *73*(9), 1326-1339. https://doi.org/10.1177/1747021820913816

Trevors, G. J., Kendeou, P., & Butterfuss, R. (2017). Emotion processes in knowledge revision. *Discourse Processes*, *54*(5-6), 406-426. https://doi.org/10.1080/0163853X.2017.1312201

Trevors, G., & Kendeou, P. (2017). *The role of affect in reducing misinformation about vaccinations*.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review*, *64*, 49-60. https://doi.org/10.1037/h0044616

Wittenberg, C., Tappin, B. M., Berinsky, A. J., & Rand, D. G. (2021). The (minimal) persuasive advantage of political video over text. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *118*(47). https://doi.org/10.1073/pnas.2114388118

WWF International (2019, September 24). *Sir David Attenborough | A message to world leaders* [Video]. Youtube. https://www.youtube.com/watch?v=fyYpExl8AJU&t=3s

Yuan, J., Tian, Y., Huang, X., Fan, H., & Wei, X. (2019). Emotional bias varies with stimulus type, arousal and task setting: Meta-analytic evidences. *Neuroscience & Biobehavioral Reviews*, *107*, 461-472. https://doi.org/https://doi.org/10.1016/j.neubiorev.2019.09.035

Zengilowski, A., Schuetze, B. A., Nash, B. L., & Schallert, D. L. (2021). A critical review of the refutation text literature: Methodological confounds, theoretical problems, and possible solutions. *Educational Psychologist*, *56*(3), 175-195. https://doi.org/10.1080/00461520.2020.1861948

Appendix A

Refutation Text - Dutch

Het versterkt broeikas Effect

Mensen verwarren het broeikas effect nogal eens met de opwarming van de aarde. Veel mensen denken dat het broeikas effect gevaarlijk is en dat het wordt veroorzaakt door menselijk handelen. Misschien denk jij dit ook. Echter, het is niet de waarheid. Het broeikas effect op aarde is NIET gevaarlijk. Het wordt NIET veroorzaakt door mensen. Het broeikas effect op aarde heeft een natuurlijke oorsprong. Het helpt de gemiddelde temperatuur op aarde comfortabel te houden voor menselijke bewoning. Zonder broeikas effect zou de gemiddelde temperatuur op aarde ongeveer -18°C zijn. Dat is ongeveer 33°C kouder dan de normale gemiddelde temperatuur, die ongeveer 15°C is. Zonder het broeikas effect zou er misschien helemaal geen leven op aarde voorkomen.

Wat is het broeikas effect? Energie in de vorm van zichtbaar licht afkomstig van de zon komt de atmosfeer van de aarde binnen. Een deel van deze zonne-energie wordt weer teruggekaatst de ruimte in. Een ander deel wordt geabsorbeerd door wolken, gassen en kleine deeltjes in de atmosfeer. De atmosfeer van de aarde maakt het mogelijk dat veel van de inkomende zonne-energie het aardoppervlak bereikt. Ongeveer de helft van de energie die van de zon komt wordt geabsorbeerd door landoppervlak en oceanen.

De aarde raakt deze energie weer kwijt door onzichtbaar infrarood licht uit te stralen. Echter, sommige gassen in de atmosfeer absorberen infrarood licht. Zij worden ‘broeikas gassen’ genoemd. Hieronder vallen onder andere waterdamp, kooldioxide en methaan. Broeikas gassen zenden dit infrarood licht vervolgens opnieuw uit. Een deel van deze energie komt weer terug op het aardoppervlak of in de oceanen. Dit doet de oppervlakte temperatuur van de aarde stijgen.

Wetenschappers hebben vastgesteld dat de gemiddelde temperatuur op aarde stijgt. In feite komen de twintig warmste jaren ooit gemeten alle na 1981. De top 10 aan warmste jaren bevindt zich volledig binnen de periode van de afgelopen twaalf jaren. Sommige mensen geloven dat de opwarming van de aarde alleen maar is toe te schrijven aan natuurlijke oorzaken. Dit is echter onjuist. De opwarming van de aarde is NIET het gevolg van alleen

maar natuurlijke oorzaken. Het is waar dat klimaatverandering een natuurlijk proces is. Echter, de snelle opwarming van de aarde die we op dit moment waarnemen kan niet worden verklaard door alleen maar natuurlijke factoren. De meeste klimaat wetenschappers zijn het er over eens dat menselijke activiteiten er de oorzaak van zijn dat broeikas gassen zich ophopen in de atmosfeer. Dit veroorzaakt een toename in de hoeveelheid energie die wordt geabsorbeerd en weer teruggekaatst naar het aardoppervlak. Het gevolg is dat het aardoppervlak opwarmt. Deze opwarming van de aarde, die bovenop de natuurlijk opwarming komt, wordt 'het versterkte broeikas effect' genoemd. Dit versterkt broeikas effect veroorzaakt de opwarming van de aarde.

Sommige mensen denken aan een laag met vervuiling, stof of gassen rond de aarde. Zij denken dat deze laag de energie in onze atmosfeer blokkeert en vasthoudt, wat opwarming van de aarde veroorzaakt. Echter, dit is niet waar. Er is geen laag met vervuiling, stof of gassen rond de aarde die de zonne-energie vasthoudt. De hoeveelheid energie die de atmosfeer van de aarde binnenkomt is ongeveer even groot als de energie die vanuit de atmosfeer de ruimte in gaat.

Sommige mensen geloven dat het dunner worden van de ozonlaag verantwoordelijk is voor de opwarming van de aarde. Echter, dit is onjuist. Niet het dunner worden van de ozonlaag veroorzaakt opwarming van de aarde; het door mensen veroorzaakt broeikas effect veroorzaakt opwarming van de aarde.

Het dunner worden van de ozonlaag is grotendeels een heel andersoortig probleem. Ozon moleculen hoog in de atmosfeer van de aarde absorberen schadelijke ultraviolette straling van de zon, met name UV-B en UV-C straling. Menselijke uitstoot van chemische stoffen die chloor bevatten, zoals chloorfluorkoolstoffen (CFCs), hebben het proces van ozon vernietiging versneld. Dit heeft een afname van ozon veroorzaakt, wat ertoe kan leiden dat meer schadelijke ultraviolette straling van de zon het aardoppervlak bereikt. Zelfs hoewel het dunner worden van de ozonlaag betekent dat meer ultraviolette straling het aardoppervlak bereikt, heeft dit geen gevolgen voor de absorptie van energie in de atmosfeer. Daarbij komt dat het grootste deel van de energie die wij van de zon ontvangen, de vorm heeft van zichtbare energie (licht). Daarom beïnvloedt een toename van ultraviolette straling als gevolg van het dunner worden van de ozonlaag niet echt de totale hoeveelheid energie die de aarde

van de zon ontvangt. Een toename van ultraviolette straling kan wel schadelijke effecten, zoals huidkanker, met zich meebrengen. Maar de gemiddelde temperatuur op aarde wordt er niet door beïnvloed. Er zou nog steeds sprake zijn van opwarming van de aarde, zelfs als het probleem van het dunner worden van de ozonlaag niet meer bestond.

Het belangrijkste broeikas gas dat in de atmosfeer terecht komt door menselijke activiteiten is koolstofdioxide. De grootste koolstofdioxide uitstoot is het gevolg van het verbranden van fossiele brandstoffen (kolen, olie en aardgas) ten behoeve van transport en het opwekken van elektriciteit.

Koolstofdioxide maakt voor minder dan 1% deel uit van de atmosfeer. Toch kunnen kleine veranderingen in koolstofdioxine concentratie heel *grote* gevolgen hebben. De gemiddelde temperatuur op aarde is gedurende de gehele 20^{e} eeuw met net iets meer dan 0.5°C toegenomen. Klimaat wetenschappers voorspellen dat in de huidige 21^{e} eeuw de gemiddelde temperatuur op aarde waarschijnlijk zal toenemen met tussen 1°C en 6°C. Dat klinkt alsof het niet veel zou zijn. Maar een toename van slechts een aantal graden zal grote gevolgen hebben voor alle facetten van onze beschaving. Denk hierbij onder andere aan voedsel productie, beschikbaarheid van drinkwater en energie voorziening.

Veel mensen denken dat de gevolgen van de opwarming van de aarde zich *niet* in onze tijd zullen voordoen. Jij hebt dit misschien ook wel eens gedacht. Echter, dit is niet juist. We zien de gevolgen *nu* al. De gevolgen bestaan onder andere uit smeltende gletsjers en heviger en vaker voorkomende stormen. Een ander gevolg wordt gevormd door groter neerslaghoeveelheden. Ook stijgen de zeewater temperatuur en de zeespiegel. Hele ecosystemen worden getroffen. Klimaat verandering veroorzaakt ook wijzigingen in waar sommige plantensoorten en dieren leven. Ook beïnvloedt het de omvang van plant- en dierpopulaties.

Sommige mensen geloven dat er niets is dat mensen kunnen doen om het probleem van klimaat verandering aan te pakken. Dit is niet juist. Mensen kunnen de uitstoot van broeikas gassen in de atmosfeer *verminderen*. Dit kan bereikt worden door het gebruik van fossiele brandstoffen (zoals olie en gas) voor elektriciteit, verwarming en transport te verminderen. Het gebruik van meer duurzame energie in plaats van fossiele brandstoffen zou helpen. Het

vergroten van bossen en het verminderen van de hoeveelheid door ons geproduceerde afval kan *ook* helpen.

Welke eenvoudige dingen kan jij doen? Je kunt kleine stapjes nemen, zoals het halen van de stekker uit het stopcontact als elektrische apparaten niet in gebruik zijn. Wat ook kan helpen is de verwarming een graadje lager zetten en de airco een paar graden hoger. Overweeg het vervangen van gloeilampen door spaarlampen. Kies ervoor te lopen, te fietsen, te carpoolen of gebruik het openbaar vervoer indien mogelijk. Vermijd het kopen van producten die je niet nodig hebt. Zoek naar andere manieren om de uitstoot van broeikas gassen te verminderen. Tenslotte, informeer anderen over alles wat je zojuist hebt geleerd.

Appendix B

Text Video

De natuur is fascinerend mooi en vol leven, maar ... voor hoe lang nog? Er vinden veranderingen plaats, veranderingen die momenteel het voortbestaan van onze beschavingen bedreigen.
Beschavingen hebben zich in het verleden kunnen ontwikkelen, grotendeels door een lange periode van klimatologische stabiliteit. Hierdoor konden mensen zich ergens vestigen en voedsel verbouwen. Dit resulteerde uiteindelijk in een internationale uitwisseling van ideeën en handel in goederen. De wereld werd een dorp.

Helaas gingen we er van uit dat deze stabiliteit een vast gegeven was. Onze economieën en politieke systemen zijn gegrond op de aanname dat de natuur zal blijven bijdragen aan de vereiste omstandigheden voor onze voorspoed. We dachten dat het altijd zo zou blijven: regelmatige seizoenen, voldoende betrouwbaar drinkwater, zeeën vol vis, bestuivers voor de bloemen, mineralen, vruchtbare grond.

Inmiddels is duidelijk dat de periode van klimatologische stabiliteit is geëindigd en onze planeet niet meer bestaat in de betrouwbare en stabiele vorm van destijds.

De gevolgen van deze destabilisatie zullen het leven in alle landen op deze planeet ernstig beïnvloeden.

Als je dit goed tot je door laat dringen is er wellicht *geen* gedachte die *meer* verontrustend is: de enige leefomstandigheden die wij als moderne mens kennen, veranderen! En ze veranderen heel snel.

Er is geen enkele tijd meer te verliezen. Er *moet* actie ondernomen worden. En ook jij, als mens, als burger, als weldenkend wezen kan bijdragen aan de oplossing. Hoe bewust leven we en schaffen we spullen aan? Wat doen we met ons afval? Op welke wijze verkleinen we de negatieve gevolgen van menselijk handelen voor natuur en klimaat? Hoe voorkomen we verdere opwarming van de aarde? Hoe draag *ik* bij aan schonere energie, duurzamer leven en bewuster omgaan met grondstoffen?

Er is niets dat op dit moment belangrijker is en verstrekkender gevolgen heeft. Ons aller handelen in de komende paar kalenderjaren zal cruciaal bepalend zijn voor de navolgende

honderden of duizenden jaren. Wees je er van bewust je verantwoordelijkheid te nemen en ook anderen op te wekken zich in te zetten voor deze zaak van leven of dood.

Appendix C

Pre-test + Scoring Rubric "Versterkt broeikas Effect"

Deel 1

In deel 1 geef je wat relevante informatie over jezelf. Kruis aan:

Wat is je leeftijd? 0 15 jaar 0 16 jaar 0 17 jaar 0 Anders, namelijk:

Wat is je geslacht? 0 Man 0 Vrouw 0 Ander/ Verkies ik niet te zeggen

Deel 2

Beantwoord elke *open vraag* zo compleet mogelijk.

1. a. Leg het broeikas effect uit. ……

   ……

   ……

   b. Hoe zeker ben je er van dat je antwoord goed is? Kleur het rondje van je keuze.

Helemaal NIET zeker						**Helemaal zeker**
1	2	3	4	5	6	7
0	0	0	0	0	0	0

2. a. Leg het versterkt broeikaseffect uit. ……..……

   ……

   ……

   b. Hoe zeker ben je er van dat je antwoord goed is? Kleur het rondje van je keuze.

Helemaal NIET zeker						**Helemaal zeker**
1	2	3	4	5	6	7
0	0	0	0	0	0	0

3. a. Leg de invloed van menselijk handelen op klimaatverandering uit. …….

……

……

……

b. Hoe zeker ben je er van dat je antwoord goed is? Kleur het rondje van je keuze.

Helemaal NIET zeker						**Helemaal zeker**
1	2	3	4	5	6	7
0	0	0	0	0	0	0

Scoring Rubric for the EGHE Refutation Text Constructed-Response Questions

(Versterkt broeikas Effect Pre-test)

Question 1. Please explain the greenhouse effect.

Scored for **3** main ideas (1 total point possible for each idea with the option to give half of a point or zero). Max possible score = 3. Min possible score = 0.

- Idea 1. Natural: The participant demonstrates the understanding that the greenhouse effect is a natural phenomenon.
- Idea 2. Warming of Earth to Habitable State: The participant demonstrates the understanding that the greenhouse effect warms the Earth, thus making it habitable. It is not dangerous. (Without it, the Earth's average temp would be about -1 degrees F. therefore, not habitable.)
- Idea 3. Mechanisms of the greenhouse effect: The participant demonstrates some understanding of the mechanisms involved. (The sun's energy comes into the Earth's atmosphere and some is absorbed by clouds, gases, etc. Some reflected back to space. Some reaches the Earth's surface (land and oceans). The Earth's surface loses this energy as infrared light, some of which is absorbed by greenhouse gases, which re-emit the infrared light. Some of this is transferred back to the Earth's surface resulting in an increase in the Earth's surface temperature.)

Question 2. Please explain the enhanced greenhouse effect.

Scored for **4** main ideas (1 total point possible for each idea with the option to give half of a point or zero). Max possible score = 4. Min possible score = 0.

- Idea 1. Human impact: The participant demonstrates the understanding that the enhanced greenhouse effect is due to human activity (for example, Burning of fossil fuels).

- Idea 2. Warming of Earth beyond that which is natural: The participant demonstrates the understanding that the enhanced greenhouse effect is raising the Earth's average temperature beyond that which is natural (for example, may mention global warming, or an increase in the Earth's temperature).
- Idea 3. Greenhouse gas emissions or carbon dioxide (CO_2): The participant demonstrates some understanding that greenhouse gases (particularly CO_2) are increasing in the atmosphere as a result of the enhanced greenhouse effect.
- Idea 4. Mechanisms of the enhanced greenhouse effect: The participant demonstrates some understanding of the mechanisms involved such as the understanding that there is an increase in the amount of energy that is being absorbed and re-radiated back to the Earth's surface.

Question 3. Explain the role of human activity in climate change.

Scored for **4** main ideas (1 total point possible for each idea with the option to give half of a point or zero). Max possible score = 4. Min possible score = 0.

- Idea 1. One point if mentions cars/transportation, deforestation, or electricity.
- Idea 2. One point if mentions burning fossil fuels, CO_2, greenhouse gases (0.5 point just for gases or pollution).
- Idea 3. One point (only if have already earned a point) if mentions one of the following impacts: extreme weather, melting of glaciers/poles, shifting ecosystems.

Idea 4. One point if mention a mitigating action (0.5 point for just reuse or recycling).

Appendix D

Post-test + Answer Model

Versterkt broeikas Effect – vragenlijst *(correct answers in bold)*

De vragenlijst hieronder bestaat uit twee delen. Deel 1 bevat drie open vragen en deel 2 bevat dertien meerkeuze vragen.

Deel 1

Beantwoord elke vraag zo compleet mogelijk.

1. Leg het broeikas effect uit.
2. Leg het versterkt broeikaseffect uit.
3. Leg de invloed van menselijk handelen op klimaatverandering uit.

Deel 2

Omcirkel bij iedere vraag het beste antwoord.

4. Het broeikas effect is:

   A. het gevolg van menselijk handelen

   B. een natuurlijk verschijnsel

   C. het gevolg van een deken van stof of gassen die de aarde omringt

   D. geen van bovengenoemde antwoorden

5. Het broeikas effect:

   A. is gevaarlijk voor het leven op aarde

   B. maakt de aarde geschikt voor bewoning door mensen

   C. is niet gerelateerd aan het leven op aarde

   D. geen van bovengenoemde antwoorden

6. Hoeveel van de zonne-energie die de atmosfeer van de aarde binnenkomt wordt opgenomen door het aardoppervlak?

   A. niets

   B. 26%

   C. 51%

   D. alles

7. Welke soort energie geeft het aardoppervlak af?

   A. zichtbaar licht

   B. ultraviolet licht

   C. infrarood licht

   D. geen van bovengenoemde antwoorden

8. De zonne-energie die de atmosfeer van de aarde binnenkomt:

   A. is minder dan de energie die van de atmosfeer van de aarde de ruimte in gaat

   B. is ongeveer gelijk aan de energie die van de atmosfeer van de aarde de ruimte in gaat

   C. is meer dan de energie die van de atmosfeer van de aarde de ruimte in gaat

   D. is niet gerelateerd aan de energie die van de atmosfeer van de aarde de ruimte in gaat

9. Welke van de volgende is geen broeikas gas?

   A. waterdamp

   B. koolstofdioxide

   C. methaan

   D. koolstofmonoxide

10. Het versterkt broeikas effect heeft geleid tot:

    A. het dunner worden van de ozonlaag

` **B. een stijging van de gemiddelde temperatuur op aarde**

    C. meer zichtbaar licht dat het aardoppervlak bereikt

    D. hogere aantallen huidkanker patiënten

11. Opwarming van de aarde:

    A. vindt in werkelijkheid niet plaats

    B. vindt wel plaats maar stelt niet zo veel voor

    C. vindt wel plaats maar de effecten zullen pas na onze tijd zichtbaar worden

    D. vindt wel plaats en de gevolgen ervan zijn al zichtbaar

12. Wat van het hieronder genoemde draagt NIET bij aan de opwarming van de aarde?

    A. het verbranden van fossiele brandstoffen om elektriciteit op te wekken

    B. het kappen van bossen

    C. de uitstoot van auto's, vrachtauto's, vliegtuigen en treinen

    D. het dunner worden van de ozonlaag

    E. Opwarming van de aarde vindt in werkelijkheid niet plaats

13. Wat kunnen mensen zelf doen om de uitstoot van broeikas gassen in de atmosfeer te verminderen?

    A. de kachel een paar graden lager zetten

    B. minder vaak de auto pakken

    C. afzien van het kopen van niet-noodzakelijke producten

    D. al het bovenstaande

    E. mensen kunnen hier zelf niet echt wat aan doen

14. Welke van de onderstaande zinnen legt het best het verband uit tussen het broeikas effect en de opwarming van de aarde?

    A. er is geen verband tussen het broeikas effect en de opwarming van de aarde

    B. een versterking van het broeikas effect leidt tot de opwarming van de aarde

    C. een toename van de opwarming van de aarde veroorzaakt het broeikas effect

    D. het broeikas effect en/of opwarming van de aarde bestaat/bestaan niet

15. Welke van de onderstaande zinnen legt het best het verband uit tussen het broeikas effect en het dunner worden van de ozonlaag?

    A. er is geen verband tussen het broeikas effect en het dunner worden van de ozonlaag

    B. een versterking van het broeikas effect leidt tot het dunner worden van de ozonlaag

    C. versnelling van het dunner worden van de ozonlaag leidt tot het broeikas effect

    D. het broeikas effect en/of het dunner worden van de ozonlaag bestaat/bestaan niet

16. Welke van de onderstaande zinnen legt het best het verband uit tussen het dunner worden van de ozonlaag en de opwarming van de aarde?

    A. er is geen verband tussen het dunner worden van de ozonlaag en opwarming van de aarde

    B. versnelling van het dunner worden van de ozonlaag leidt tot opwarming van de aarde

    C. een toename van de opwarming van de aarde veroorzaakt het dunner worden van de ozonlaag

    D. het dunner worden van de ozonlaag en/of de opwarming van de aarde bestaat/bestaan niet

17. Heb je jezelf tussen de vorige toets afname en die van vandaag nog verdiept (bijv. middels een documentaire of beschouwend artikel) in het onderwerp of de onderwerpen van dit onderzoek?

    Ja Nee *(omcirkel het juiste antwoord)*

    Indien *Ja*: wanneer en waarmee/hoe ?
    ……..……..

Einde.

Dank voor het beantwoorden van de vragen ! Dank voor het deelnemen aan het onderzoek !

Appendix E

Results Normality Tests

Results Shapiro-Wilk Test

Tests of Normality

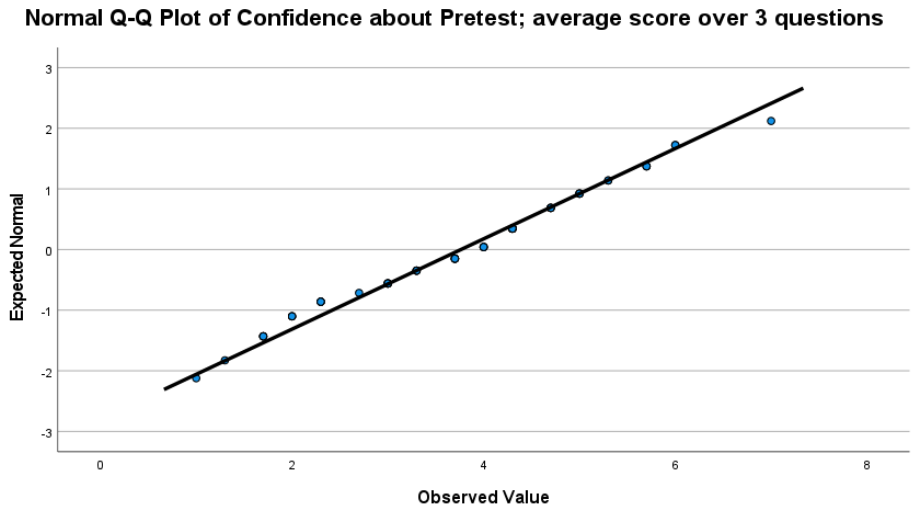
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pre-test OPEN	.125	58	.024	.971	58	.175
Post-test OPEN	.123	58	.030	.966	58	.103
Post-test MC	.141	58	.006	.960	58	.056
Delayed Post-test OPEN	.116	58	.052	.949	58	**.016**
Delayed Post-test MC	.192	58	.000	.962	58	.066
Confidence (Pre-test) [b]	.104	58	.186	.979	58	.419

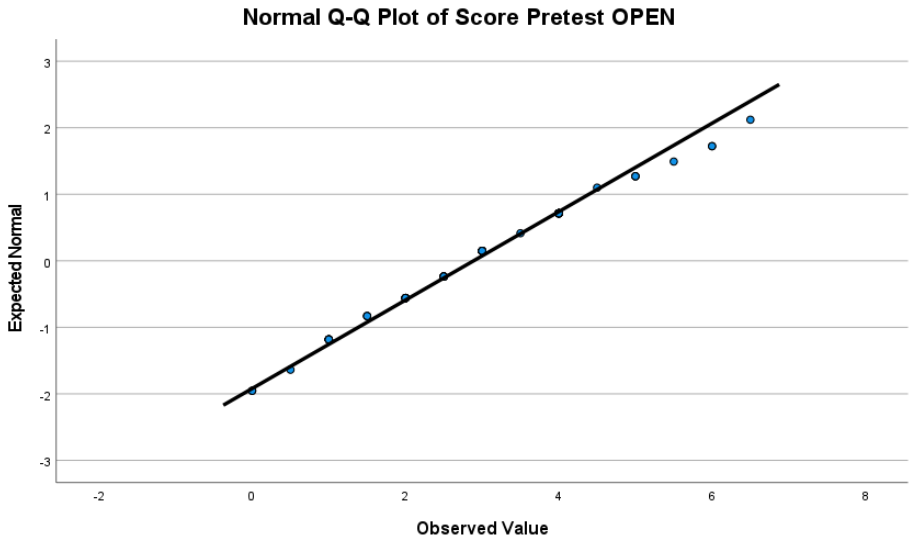
a. Lilliefors Significance Correction

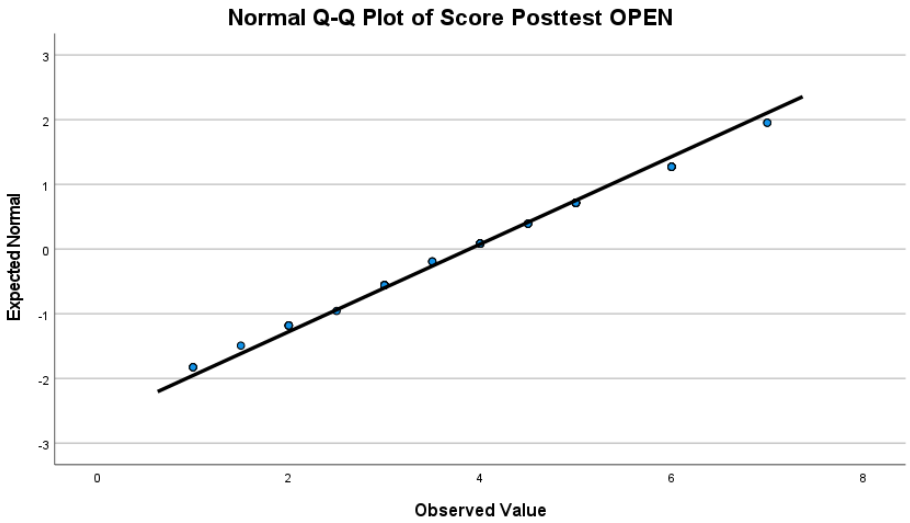
b. average score over 3 questions on a 7-points Likert scale

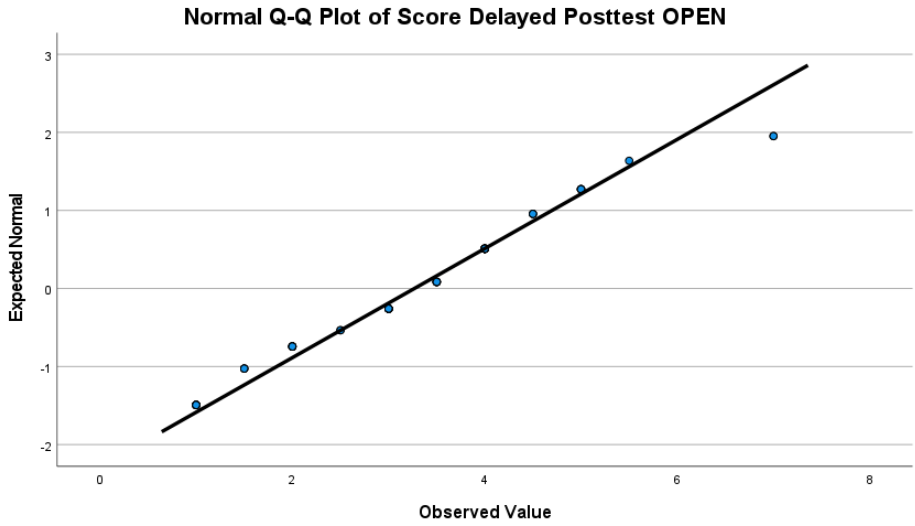
Appendix F

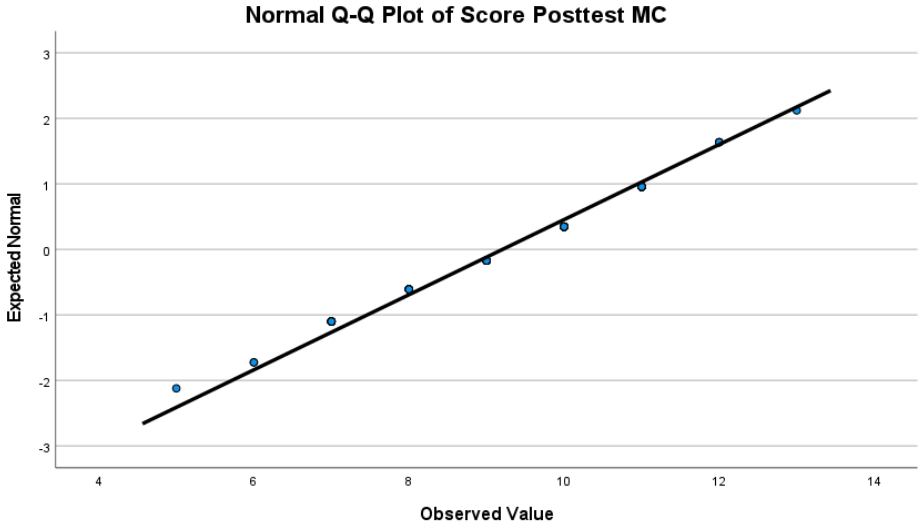
Distributions of Scores per Test

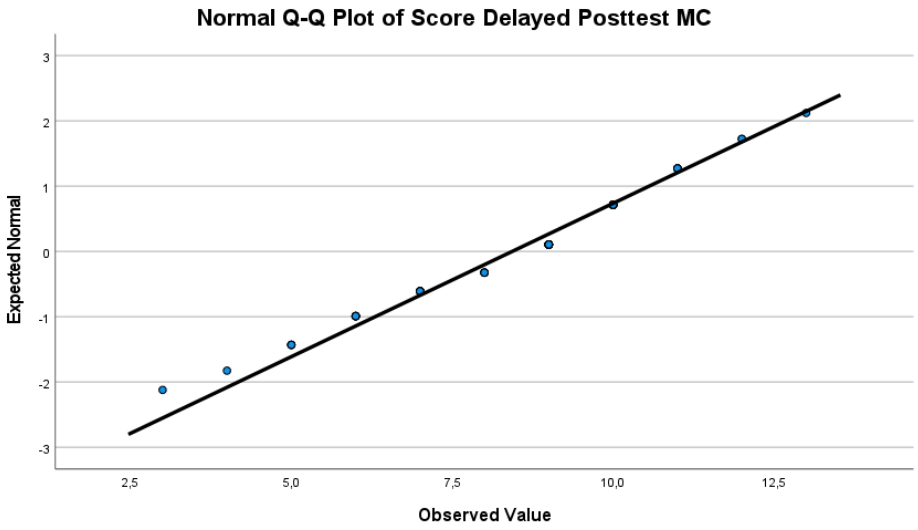





Normal Q-Q Plot of Score Posttest OPEN
Expected Normal
Observed Value
3
2
1
0
-1
-2
-3
0
2
4
6
8


Normal Q-Q Plot of Score Delayed Posttest OPEN
Expected Normal
Observed Value
3
2
1
0
-1
-2
0
2
4
6
8


Normal Q-Q Plot of Score Posttest MC
Expected Normal
Observed Value
3
2
1
0
-1
-2
-3
4
6
8
10
12
14


Normal Q-Q Plot of Score Delayed Posttest MC
Expected Normal
3
2
1
0
-1
-2
-3
2,5
5,0
7,5
10,0
12,5
Observed Value