Radboud University


# The Effect of Foreign Language and Subtitles on the Effectiveness of Instructional Videos

Sèfie Bekkers
MA Communication and Persuasion
Radboud University Nijmegen
5 October 2021

**Abstract**

More and more manufacturers of assembly required products are starting to make use of online instructional videos, in addition to or instead of print manuals. As businesses are increasingly starting to operate on a global scale, these companies must cater to consumers from various countries. The current study has therefore investigated whether adaptation (using spoken L1 and/or L1 subtitles) or standardisation (using the lingua franca English) are preferred to improve the effectiveness of instructional videos. The effectivity of language use in instructional videos in this study was measured by using an instructional video on a LEGO house, which represented an assembly required product such as an IKEA wardrobe. Dutch participants were investigated as the Dutch population has been presumed to speak English as a foreign language. The research has drawn on the dual-coding hypothesis and the cognitive load theory and has tried to add to this body of literature by varying the cognitive load by manipulating the instructional video using three different types of information (visual, verbal and audiovisual). The findings imply that, in general, instructional videos neither become more effective by adapting language choice (L1 versus L2), nor by changing the presence or absence of subtitles. This suggests that the extra costs of adaptation do not outweigh its potential benefits, resulting in an argument in favour of standardisation.

*Keywords:* instructional videos, foreign language use, L1 subtitles, same language subtitles, adaptation versus standardisation

## Introduction

### *Background*

Globalisation has caused more and more businesses and organisations to start operating on an international scale, which has resulted in companies having to cater to consumers from different countries. Businesses translate their advertisement slogans, create distinct Instagram-accounts for various countries and they write instruction manuals in the target language of the countries that consume their products and services. Companies, especially retailers and manufacturers of assembly required products, are also increasingly starting to make use of online instructional videos, next to or instead of print manuals. This has been facilitated by technological developments that have led to the rise and popularity of video sharing platforms like YouTube. Furthermore, earlier research has proven that dynamic instructions outperform static instructions, which also include written instructions such as manuals (Castro-Alonso, Ayres & Paas, 2016), leading to more efficient use of the instruction by consumers. Businesses and organisations would thus benefit from understanding what makes an effective instructional video and the way they might affect their consumers, by figuring out which factors play a role in the overall effectivity of instructional videos. By gaining more insight in the exact workings of instructional videos, and in this study more specifically insights in the effect of using a foreign language or the mother tongue, companies allow themselves to make a balanced choice between adaptation (using the mother tongue or L1 subtitles) and standardisation (which often comes down to the use of English as a lingua franca) on the current global market.

*Adaptation* takes place when instructional videos are translated to the target language of the consumer country and/or by means of subtitles in the domestic language of a country. Instructional videos can also be *standardised*, by using one commonly spoken language or no language at all. The English language would be one of the obvious choices for standardisation of instructional videos, as English is often considered a lingua franca and roughly three out of four speakers of English in the world are not even native speakers (Crystal, 2003). According to Education First, most European countries have a very high to moderate English proficiency (EF, 2020). The Netherlands has the highest English proficiency of all countries that do not have English as their native language (EF, 2020). In countries like the Netherlands, it may therefore seem effective and the most cost-efficient for companies to simply standardise their instructional videos using the English language or by adapting them to the country by adding L1 subtitles. However, effective and well-understood instructional videos are not only important for a good evaluation of the videos themselves. Gök, Ersov and Börühan (2019) found that user manual quality has a positive significant relation with perceived product quality. These findings can be combined with earlier research that has shown that product performance quality has a positive effect on overall brand reputation and brand satisfaction, which in

turn increase brand loyalty (Selnes, 1993). Thus, instructional videos can also affect customer loyalty and the way consumers view the quality of the company and their products in general. It would therefore be worthwhile for companies to investigate how the use of either L1 or L2 spoken language, subtitles and even same language subtitles might affect the effectivity and evaluation of instructional videos to gain insight in the effects of adaptation and standardisation.

The current research will therefore try to provide insight into the workings of L1 or L2 spoken language and L1 subtitles accompanying L2 or L1 spoken language in commercial instructional videos. The study will do this by looking at the effectivity of an instructional video on building a LEGO house. The LEGO house will represent an assembly required product, such as an IKEA wardrobe. This research will compare the effectiveness of using either English or Dutch spoken language in itself, as well as in combination with Dutch subtitles. Participants will be Dutch, as 90% of the Dutch population has been found to speak English as a foreign language (Eurobarometer, 2012; Edwards, 2014). Dutch participants can thus be presumed to have at least a basic understanding of the English language. Subtitles are usually used to enhance understanding for consumers trying to comprehend a foreign language. However, the aim of this study is to first establish the effects of standardisation and adaptation on relatively proficient L2 speakers in general. Dutch participants are thus appropriate subjects for investigating the effects of spoken L1 and L2 language and L1 subtitles in instructional videos. Once more evidence has been found on whether standardisation or adaptation is preferred by relatively proficient L2 speakers, the differences between less proficient and more proficient speakers can be considered. The current study will therefore only attempt to make an initial tentative prediction towards the differences between levels of language proficiency.

The remainder of this chapter will first discuss the dual-coding hypothesis, showing the positive effects of including language in instructional videos as opposed to using no language at all. It will then continue to explain the effectivity of either standardisation by using a spoken L2, or adaptation by using a spoken L1. After this, the cognitive load theory, the split-attention effect and the redundancy effect will be discussed. These theories will then be combined with previous research on the effectiveness of L1 subtitles accompanying L2 spoken language. The chapter will then discuss previous literature on the use of same language subtitles. Eventually the literature review will lead to the research question as well as the hypotheses this study proposes.

*Literature Review*

Ready-to-assemble furniture retailer IKEA has started to produce instructional videos showing the assembly of their products. The company has decided to standardise their instructional videos by almost completely omitting the verbal component, leaving only a limited amount of written language. The absence of the verbal component which instructional videos might have, may seem like a low-

cost and effective way to make the videos suitable for a larger audience. The dual-coding hypothesis, however, proposes that the understanding of a video is the most effective when both visual *and* verbal elements are used (Paivio, 1990). This theory suggests that two distinct mental representations can occur through both a visual and a verbal mechanism. When the visual and the verbal mechanisms are working at the same time, the given information is easier to understand as this allows the learner to connect the verbal and the visual mental representation together. This suggests that instructional videos are most effective when they are combined with simultaneous speech.

Mayer and Anderson (1991) have tested the dual-coding hypothesis in two separate experiments. The first experiment was designed to test the hypothesis that the simultaneous creation of the visual and the verbal mental image allowed the participants to connect them together, such that understanding would be facilitated. During the first experiment students were presented with an animation showing how a bicycle tire pump functions. The experiment consisted of two conditions. One group viewed the animation including a simultaneous voice-over, whereas the other group first heard the voice-over and then watched the animation. The former turned out to be significantly more knowledgeable about the bicycle tire pump. The result demonstrates that verbal and visual mental representations indeed can be connected more easily when they are presented simultaneously. During the second experiment, participants were divided into four groups. The first group only had access to the verbal component of the animation, the second only saw the visual animation and the third group was presented with the animation simultaneously with the voice-over. The fourth group served as a control group, which is why they were neither presented with the visual animation nor the voice-over. As expected, the third group outperformed the other groups regarding their knowledge on the bicycle tire pump. Both experiments add to the body of evidence for the dual-coding hypothesis, as they show that both visual and verbal elements facilitate the understanding of instructional videos and neither of them should thus be left out.

The effectivity of the addition of verbal elements has so far predominantly been considered in the L1 of the target group. The question remains whether the verbal component could be as effective, or even more effective, in an L2. Research concerning the verbal use of an L1 versus an L2 in the domain of commercial instructional videos has not been conducted yet. A potentially useful field to consider in this regard, is the field of second language acquisition and L2 education. Earlier research seems to suggest that the use of instructional videos within the foreign language classroom, as opposed to classical classroom instruction, is effective in improving listening (Herron, Morris, Secules & Curtis, 1995), speaking (Namaziandost, Esfahani & Hashemifarnia, 2018) or even language proficiency as a whole (Mekheimer, 2011). None of these studies, however, have investigated whether the L2 or the L1 would be the most effective to use for such instructional videos. The potential differences in effectiveness of the use of the L1 versus the L2 in instructional videos

therefore remain unclear. Previous research, however, did look into the use of the L1 versus the L2 within the field of L2 education outside of instructional videos, but within the domain of classroom instructions. Cook (2001, as qtd. in Bruhlmann 2012) and Anton and DiCamilla (1998) have emphasised the importance of *providing instructions* in the classroom in the L1 to facilitate the acquisition of the L2. Learners cannot be completely separated from their L1, as they will still think in their native language and probably use the L1 whilst (de)constructing the L2. The use of the L1 during instructions has been shown to facilitate task and classroom management procedures (Cook, 2001). These arguments, however, specifically pertain to classroom instructions regarding the acquisition of a language, which requires long-term memory and understanding. The current study investigates instructional videos for more procedural tasks, like product assembly or maintenance. These videos only require short-term understanding, setting them apart from the domain of second-language acquisition. Even though the abovementioned arguments pertain to educational instructions within the classroom, the L1 seems to play a specific role concerning the clarity, understanding and effectiveness of instructions. Besides these arguments, consumers generally should have a better understanding of their mother tongue than their L2. The current research therefore proposes that instructional videos in the L1 might be more effective than their L2 counterparts.

Another option concerning adaptation, besides translation to the L1, is the use of subtitles. Following the dual-coding hypothesis, the addition of subtitles would force a person to include written text into their visual mechanism, in addition to the other visual information already present, namely the video image. This may lead to cognitive overload, as the visual channel now receives too much information to process. The cognitive load theory, which is based on the dual-coding hypothesis, could provide more insight into the effectiveness of adding subtitles to the verbal component of instructional videos. Cognitive load (CL) is a theoretical construct that comprises the internal processing of tasks, which cannot be readily observed (Sweller, 2011; Kruger, Hefer & Matthew, 2013). CL can be divided into three segments: intrinsic, extraneous and germane CL. Intrinsic CL is caused by external factors of the material, including the difficulty of understanding, and can thus not be manipulated. Extraneous CL originates from the way information is presented, for instance by in- or excluding subtitles. Germane CL consists of the remaining cognitive resources and is used to process and understand the subject matter. The emergence of more intrinsic and extraneous CL takes up space for germane CL to properly work, which in turn might lead to cognitive overload. Within the educational context, subtitles are generally considered to be detrimental to learning as they are demanding for the extraneous CL, resulting in a reduction of available germane CL to understand the material (Paas & Van Merriënboer, 1993, as qtd. in Kruger, Hefer & Matthew, 2013). CL theory therefore seems to suggest that the addition of subtitles would not be beneficial for instructional videos because of the increased extraneous CL.

To fully understand the predicted ineffectiveness of subtitles by CL theory, two effects that follow from CL theory and the dual-coding hypothesis are relevant: the split-attention effect and the redundancy effect (Sweller, Ayres & Kalyuga, 2011). The split-attention effect implies that simultaneous animation and written text lead to comprehension difficulties, as dividing attention between and combining both visual aspects is very demanding for the extraneous CL. This leaves relatively less cognitive room for germane CL. In addition, the redundancy effect occurs when information is displayed concurrently in written and spoken form, which is detrimental to understanding the given material. According to the redundancy effect, processing the same information in two different ways is demanding for the cognitive load. Simultaneous written and spoken language for instance, require constant revision on whether the two presentations of language state the same thing. This takes up a lot of cognitive space leaving little germane CL to actually understand the material. Evidence has been found for both the split-attention effect and the redundancy effect (Sweller, 2002; Moreno & Mayer, 2002). Both the split-attention effect and the redundancy effect therefore seem to lead to the suggestion that subtitles will increase cognitive load and in turn lower the effectiveness of the instructional video, regardless of the language of the video.

Several empirical studies have looked at the effects of subtitles in an educational context. However, conflicting results have been found. A study by Kruger, Hefer and Matthew (2013) found that the presence of L1 subtitles leads to higher self-reported frustration levels and mental effort, compared to when subtitles are absent. Participants' native language was Sesotho and they were all enrolled in a study program taught in the English language. Participants' eyes were tracked during the experiment, in which they were presented with a video of an English psychology lecture, either with English subtitles, Sesotho subtitles or no subtitles at all. After watching the video, participants completed a comprehension test and they reported how much cognitive effort understanding the lecture took through a questionnaire. The results showed that comprehension did not significantly differ between the three groups. The eye-tracking results, however, showed that the L1 subtitles group neglected the subtitles significantly more than the L2 subtitles group. On average the Sesotho subtitles were only utilised half of the number of times the English subtitles were utilised. The L1 subtitles group also reported higher levels of frustration and cognitive effort than both the L2 subtitles group and the no subtitles group. Even though there was no significant difference in comprehension, the outcome of this experiment suggests that the presence of L1 subtitles may indeed lead to an increase in CL. This in turn might have led to the neglection of the L1 subtitles and the higher self-reported frustration. These findings seem to be in line with the CL theory, the redundancy effect and the split-attention effect. Looking at same language subtitles, however, the L2 subtitles group reported significantly less frustration and mental effort and they utilised the subtitles significantly more than

the L1 subtitles group. These results seem to be contradictory to the aforementioned theories and effects.

Vulchanova, Aurstad, Kvitnes and Esthuis (2015) found a similar result regarding L2 subtitles but have suggested an opposing view on the effectiveness of L1 subtitles. Norwegian students were divided into three groups: no subtitles, L1 subtitles (Norwegian) and L2 subtitles (English). They watched an episode of an English spoken cartoon and then filled in a comprehension questionnaire. Both L1 and L2 subtitles turned out to significantly improve comprehension over no subtitles. In contrast to the findings of Kruger, Hefer and Matthew (2013), this research seems to imply that L1 subtitles do *not* lead to an increase in CL but instead seem to help the viewer to process and comprehend the spoken L2.

The effectivity of native language subtitles supporting a spoken L2 thus remains to be a bit of a grey area. On the one hand, the redundancy effect may be weaker or non-existent because the written and the spoken language differ from each other, meaning the information that has to be processed twice is not fully identical. Subtitles may therefore not lead to cognitive overload and instead prove useful for the comprehension of an L2. On the other hand, L2 speakers might cognitively process the L1 subtitles as identical, or somewhat identical, to the spoken L2 language since it conveys the same meaning. In this case the redundancy effect should still apply. The split-attention effect may apply to L1 subtitling regardless, indicating that CL theory might still affect the effectiveness of native language subtitles accompanying a spoken L2. Limited empirical research concerning instructional videos aiming to specifically investigate the use of native language subtitles as opposed to no subtitles has been done and within the educational domain opposing views have been proposed. The question remains which of these variables also play a role in the context of instructional videos for procedural tasks within a commercial context. This study aims to investigate this.

The aforementioned studies have already provided some evidence towards the effectiveness of same language subtitles (SLS) and more empirical research has found additional evidence in the same direction (Diao, Chandler & Sweller, 2007; Hirose & Kamei, 1993; Kruger, Hefer & Matthew, 2013; Neuman & Koskinen, 1992). Hirose and Kamei (1993) and Neuman and Koskinen (1992) found that for bilingual speakers with English as their L2, English subtitles accompanying spoken English language led to a higher comprehension than when subtitles were absent. Diao, Chandler and Sweller (2007) found a similar result conducting an experiment to provide insight in the effects of a written L2 accompanying spoken L2 on comprehension. Native Chinese participants were divided into three groups: English audio only, English audio including an English written script and English audio including simultaneous English subtitles. They performed several different tasks concerning comprehension and vocabulary. In all comprehension tasks the simultaneous English subtitles group significantly outperformed the other two groups, implying that the split-attention effect and the

redundancy effect did not apply to SLS in this experiment and SLS did not lead to an increase in extraneous CL. These findings contradict the predictions made by the CL theory, the split-attention effect and the redundancy effect since SLS are expected to cause cognitive overload, a redundancy effect and split-attention problems. The assumption would be that the redundancy effect might even be more present with SLS than with L1 subtitles accompanying spoken L2, as SLS offer a literal transcription of the spoken information. However, research as has been discussed so far seems to disprove these predictions. One distinct difference between all these studies and the current research is the fact that all these researchers investigated SLS in the L2, whereas the current study will focus on SLS in the L1. The cognitive load resulting from L1 SLS may differ from L2 SLS but, since this has not been looked at yet, the current research will base its hypothesis on previous empirical evidence.

*Research Question and Hypotheses*

All the aforementioned information has led to the current research focusing on the effects of the choice between spoken L1 and spoken L2 and the presence or absence of L1 subtitles. The goal of the study is to gain more insight into how language use can contribute to the effectiveness of instructional videos. This leads to the following research question: *How do the choice of language (L1 vs L2) and the presence of L1 subtitles influence the effectiveness of instructional videos?* In response to this question, the current research proposes the following hypotheses based on previous literature:

> **H1**: The use of a spoken L1 in instructional videos is more effective than the use of a spoken L2.
>
> **H2**: Instructional videos with same language subtitles (spoken L1 accompanied by L1 subtitles) are more effective than instructional videos without subtitles.

Different views on whether L1 subtitles accompanying spoken L2 either help the viewer to comprehend and process the L2 or lead to cognitive overload, have been proposed. L1 subtitles do seem to have an effect within the field of education but following the results of previous research, as well as the fact that limited empirical research regarding L1 subtitles within the field of instructional videos has been done, the direction of this effect remains unclear. H3 therefore does predict an effect without expecting a specific direction.

> **H3**: The presence of L1 subtitles supporting a spoken L2, in comparison to no subtitles, affects the effectiveness of instructional videos.

To gain more specific insight into the exact workings of cognitive load and the interaction of the presence of subtitles, the instructional video used in this study contains another manipulation. The LEGO house consists of ten layers. For some layers, information is only given visually. This,

according to the cognitive load hypothesis, should not lead to cognitive overload, regardless of the presence or absence of subtitles. Since only visuals are used, neither spoken language nor subtitles are present, leaving enough cognitive capacity to process the given information. The expectation according to the dual-coding theory, however, is that instructional videos are more effective when both the visual and the verbal mechanism are used concurrently to process the instructions (Paivio, 1990; Mayer & Anderson, 1991).

Some of the layers are only accompanied by verbal information. According to the cognitive load hypothesis, this may or may not lead to cognitive overload when subtitles are present. This depends on whether the split-attention effect or the redundancy effect gets the most weight. When L1 subtitles are present, the split-attention effect does not apply as the viewer does not have to focus on two visual elements at the same time. Considering the split-attention effect only, enough cognitive space for germane CL should be left. However, the redundancy effect may apply since the same verbal information has to be processed twice, leaving less cognitive room to process the given instructions (Sweller, Ayres & Kalyuga, 2011). This manipulation can thus be used to gain a better understanding on whether L1 (Dutch) subtitles are processed visually or verbally when watching an instructional video in either an L1 (Dutch) or an L2 (English).

Lastly, some layers contain simultaneous visual and verbal information. Looking at the dual-coding theory, visual or verbal information only should be less effective than when both visual and verbal information are included. Dual-coding theory proposes that instructions are most effective when they can be processed through both the visual, as well as the verbal mechanism simultaneously (Paivio, 1990; Mayer & Anderson, 1991). The question remains however, whether L1 subtitles are indeed processed visually or verbally, leading to cognitive overload when both visuals as well as a spoken L2 are present, or whether L1 subtitles are processed in a different way, instigating a better understanding of the spoken language (H3). When visuals and spoken L1 language are used at the same time, same language subtitles accompanying spoken language are expected to cause cognitive overload according to the split-attention or the redundancy effect. Previous research, however, has found results pointing into a different direction (H2). The manipulation can therefore also be used to check whether L1 subtitles work the way cognitive load theory proposes, or whether they improve understanding of the spoken language through a different mechanism and whether this differs between spoken L1 and spoken L2.

## Method

*Materials*

Participants were presented with an instructional video on building a LEGO house, in which the LEGO represents an assembly required product such as an IKEA wardrobe. The video was filmed using a tripod, meaning only one static viewpoint was used. The video consisted of two hands building the LEGO house and a voice-over giving instructions. The whole video took four minutes and 22 seconds to watch. The complete LEGO house consisted of 48 bricks distributed over ten layers, each with a different composition of either 2x2 square bricks or 4x2 rectangular bricks which could either be red, blue, white, green or yellow. The LEGO house was constructed such that frequent LEGO builders would still have to pay attention instead of building it on autopilot. Figures 1 and 2 show a screenshot of the instructional video as well as the composition of the ten layers.

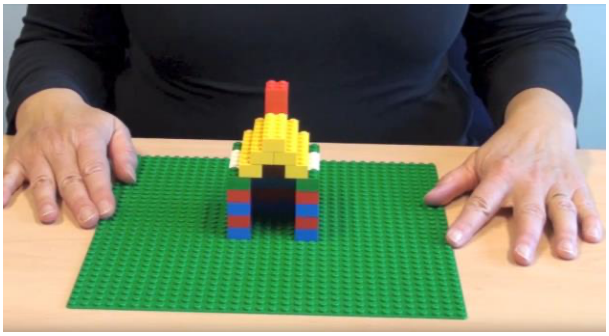

*Figure 1. Screenshot from the instructional video showing the finished LEGO house*

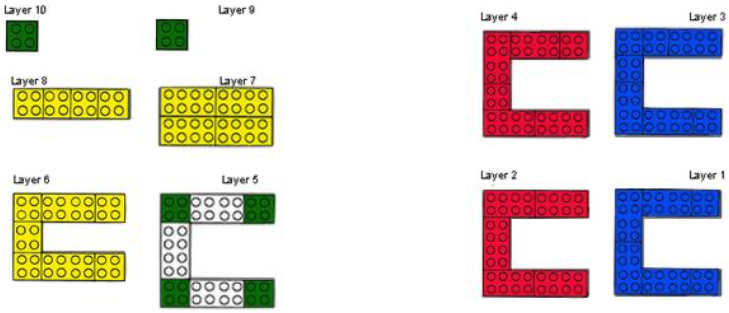


*Figure 2. The ten layers of the LEGO house*

There were two independent variables (Language and Subtitles) with each two levels leading to four conditions, because the spoken language in the instructional video could be either Dutch (L1) or English (L2) and Dutch (L1) subtitles were either present or absent. Participants were randomly assigned to one of the four conditions: spoken L1 without subtitles ($Dutch_{AudioOnly}$), spoken L2 without subtitles ($English_{AudioOnly}$), spoken L1 with L1 subtitles ($Dutch_{DutchSubs}$) and spoken L2 with L1 subtitles ($English_{DutchSubs}$). All language use in the instructional video was designed to be of a basic level and was relatively easy to understand.

The experiment contained a third independent variable, namely Type of Information (Type of Info) provided in the video. This independent variable was operationalised by using three levels that were constructed based on the information given per layer. Some layers were only accompanied by visual information (Visual), some layers only contained verbal information (Audio) and some layers consisted of both verbal and visual information (Audio-visual). The Visual condition consisted of layers 2, 3 and 7, for which the verbal component was omitted and thus no verbal information regarding the size and the order of the bricks was given. The Audio condition was designed by

visually building layer 9 and 10 in a certain colour, while verbally communicating that this layer should be built in another colour. The Audio-visual condition comprised layers 1, 4, 5, 6 and 8 for which both verbal and visual information were given regarding the size, colour and position of the bricks.

*Subjects*

Participants were native Dutch speakers, who are generally considered to have at least a basic understanding of the English language. As mentioned earlier the Dutch population has the highest English proficiency of all countries that do not have English as their native language (EF, 2020) and 90% of the Dutch population speaks English as a foreign language (Eurobarometer, 2012; Edwards, 2014). Randomly selected Dutch participants above 18 years old took part in the experiment ($M$ = 39.26, $SD$ = 9.57, [18-61]). Participants were divided over the four conditions as follows: $\text{Dutch}_{\text{AudioOnly}}$ – 39 participants (collected during previous research) $\text{English}_{\text{AudioOnly}}$ - 34 participants (collected during previous research), $\text{English}_{\text{DutchSubs}}$ – 45 participants (collected during previous research) and $\text{Dutch}_{\text{DutchSubs}}$ - 30 participants (collected specifically for the current research). This comes down to 148 participants in total. A total of 70 men and 78 women participated in the experiment. The most frequent educational level was HBO and participants educational level ranged from MBO to WO. The native language of all participants was Dutch.

Gender was distributed evenly over the four conditions ($\chi^2$ (3) = 7.51, $p$ = .057), as was English Proficiency ($\chi^2$ (9) = 10.01, $p$ = .349). The distribution of Dutch Proficiency could not be checked because data was missing. A separate one-way analysis of variance also showed that Manual Preference did not significantly differ across conditions ($F$ (3, 144) < 1).

However, educational level was not evenly distributed across the four conditions ($\chi^2$ (6) = 14.55, $p$ = .024), nor was LEGO Dexterity ($\chi 2$ (12) = 46.12, $p$ < .001). The crosstab for educational level showed that this uneven distribution was caused only by the $\text{Dutch}_{\text{DutchSubs}}$ condition, which consisted of a higher amount of WO educated participants (37.8%) as compared to the other three conditions. The crosstab for LEGO Dexterity showed that this uneven distribution was caused only by the $\text{Dutch}_{\text{DutchSubs}}$ condition again, which contained a lower number of participants that had used LEGO during the last week (0.0%) and during the last month (2.4%) as compared to the other three conditions. The $\text{Dutch}_{\text{DutchSubs}}$ condition also contained a higher number of participants whose last time using LEGO was longer than ten years ago (54.5%) as compared to the other three conditions (see also table 1 and 2 for the crosstabs). The scope of this thesis, however, is too limited to take these differences into account. The results should nonetheless be interpreted with care.

A one-way analysis of variance on age for the four conditions showed a significant main result ($F$ (3, 141) = 12.18, $p$ < .001, $\eta^2$ = .206). The $\text{Dutch}_{\text{DutchSubs}}$ condition ($M$ = 30.83, $SD$ = 13.66)

contained significantly younger participants on average than $English_{DutchSubs}$, ($M = 41.00$, $SD = 6.46$; $p < .001$, Bonferroni-correction), $English_{AudioOnly}$ ($M = 41.21$, $SD = 7.89$; $p < .001$, Bonferroni-correction) and $Dutch_{AudioOnly}$ ($M = 42.22$, $SD = 5.83$; $p < .001$, Bonferroni-correction).

Table 1. Division of educational level over the four experimental conditions

| | Education | | | |
|---|---|---|---|---|
| Condition | MBO | HBO | WO | Total |
| $Dutch_{AudioOnly}$ | 12 (36.4%) | 14 (25.5%) | 9 (20.0%) | 35 (26.3%) |
| $English_{AudioOnly}$ | 4 (12.1%) | 15 (27.3%) | 4 (8.9%) | 23 (17.3%) |
| $English_{DutchSubs}$ | 12 (36.4%) | 18 (32.7%) | 15 (33.3%) | 45 (33.8%) |
| $Dutch_{DutchSubs}$ | 5 (15.2%) | 8 (14.5%) | 17 (37.8%) | 30 (22.6%) |
| Total | 33 (100%) | 55 (100%) | 45 (100%) | 133 (100%) |

Note. $\chi^2 = 14.55$, $p = .024$, $n = 133$, df = 6. Row percentages are shown next to observed cell counts.

Table 2. Division of LEGO dexterity over the four experimental conditions

| | Last used LEGO | | | | | |
|---|---|---|---|---|---|---|
| Condition | 1 | 2 | 3 | 4 | 5 | Total |
| $Dutch_{AudioOnly}$ | 9 (45.0%) | 10 (24.4%) | 14 (31.8%) | 1 (10.0%) | 5 (15.2%) | 39 (26.4%) |
| $English_{AudioOnly}$ | 4 (20.0%) | 13 (31.7%) | 8 (18.2%) | 3 (30.0%) | 6 (18.2%) | 34 (23.0%) |
| $English_{DutchSubs}$ | 7 (35.0%) | 17 (41.5%) | 15 (34.1%) | 2 (20.0%) | 4 (12.1%) | 45 (30.4%) |
| $Dutch_{DutchSubs}$ | 0 (0.0%) | 1 (2.4%) | 7 (15.9%) | 4 (40.0%) | 18 (54.5%) | 30 (20.3%) |
| Total | 20 (100%) | 41 (100%) | 44 (100%) | 10 (100%) | 33 (100%) | 148 (100%) |

Note. $\chi^2 = 46.12$, $p < .001$, $n = 148$, df = 12. Row percentages are shown next to observed cell counts. Columns denote time since last used LEGO: 1 = last week, 2 = last month, 3 = over a year ago, 4 = over 5 years ago, 5 = over 10 years ago.

*Design*

A real-life experiment was conducted with three independent variables, two with two levels and one with three levels. The first independent variable was a between-subject variable named Language, which could either be Dutch or English. The second between-subject variable was Subtitles, which were either present or absent. The third and last independent variable was a within-subject variable called Type of Info, which could either be Visual, Audio or Audio-visual.

*Instruments*

All dependent variables were measured based on a paper by Nederstigt and Hilberink-Schulpen (2019). The participants' performance was calculated with an accuracy score (0-100%) per layer. The layers were then further coded into three categories, namely Colour, Brick Size and Position. Each layer was first rewarded with a score of 0-100% based on the colour of the layer. Participants were rated 100% if they used the correct colour(s), 50% if they added an incorrect colour and 0% if they used the wrong colour(s). No participants used more than one incorrect colour in a layer. Brick Size was calculated by counting the number of bricks used in a layer and comparing this to the correct number. If a participant used the correct number of bricks, the participant was given a score of 100%. Each extra or missing brick lowered the score by 100% divided by the correct number of bricks. Position was examined by counting the number of bricks placed in the correct position in a layer and was scored in a similar manner as Brick Size. Coding these layers individually allowed for the calculation of a total score on Colour, Brick Size, Position and Total Performance as well as a total score on the Audio layers, the Visual layers and the Audio-visual layers separately. The different types of information (Audio, Visual, Audio-visual) could then also be separated into a total score for Colour, Brick Size and Position per Type of Info. The dependent variables that comprise the performance thus are Total Performance, Colour Total, Brick Size Total, Position Total, AudioVisual Total, Visual Total, Audio Total, Colour AudioVisual, Colour Visual, Colour Audio, Brick Size AudioVisual, Brick Size Visual, Brick Size Audio, Position AudioVisual, Position Visual and Position Audio.

Attitude was operationalised into five separate variables: attitude towards the video (Attitude Video), attitude towards the audio (Attitude Audio), attitude towards the task (Attitude Task), attitude towards the subtitles (Attitude Subtitles) and attitude towards the difficulty of the task (Task Difficulty). All variables were measured on a 5-point Likert scale that ranged from 1 (completely agree) to 5 (completely disagree).

Attitude Video was measured using six items regarding the participants' attitude towards the video. Participants rated the video on the items *structured well*, *clear*, *not interesting*, *easy to remember*, *of good quality* and *portrayed well*. The item *not interesting* was recoded positively since this was the only negatively worded statement. A Cronbach's Alpha was calculated separately for all four components of Attitude to check the reliability of the scales. The Cronbach's Alpha in the current study has first been calculated by using the data that was generated by the $\text{Dutch}_{\text{DutchSubs}}$ group, as the Cronbach's Alpha was already calculated earlier for the other three conditions. The decisions following from these analyses were made based on the previously mentioned research conducted by Nederstigt and Hilberink-Schulpen (2019). The reliability of Attitude Video within the $\text{Dutch}_{\text{DutchSubs}}$ condition, consisting of six items, was acceptable: $\alpha = .77$ and could be turned into good by deleting

the item *not interesting*: $\alpha$ = .80. Nederstigt and Hilberink-Schulpen (2019) found that deleting the item *not interesting* made the scale more reliable for the other three conditions as well. The reliability of Attitude Video within the other three conditions, consisting of six items, was indeed acceptable: $\alpha$ = .77 and could be turned into good by deleting the item *not interesting*: $\alpha$ = .84. The current research will therefore delete this item. The mean of the remaining five items was used to calculate the compound variable Attitude Video, which was used in the remaining analyses.

Attitude Audio was measured with six items pertaining to participants' attitude towards the spoken language. The items were *easy to understand*, *difficult to follow*, *distracting from the task*, *aiding in the task*, *too fast* and *too informative*. The items *easy to understand* and *aiding in the task* were recoded negatively to make them consistent with the other items. The reliability of Attitude Audio comprising six items turned out to be acceptable: $\alpha$ = .74. The reliability of Attitude Audio within the other three conditions, consisting of five items (excluding *too informative*) turned out to be marginally acceptable: $\alpha$ = .67. Earlier research for the other three conditions found that Attitude Audio was most reliable when the item *too informative* was deleted (Nederstigt & Hilberink-Schulpen, 2019). Even though the reliability of Attitude Audio (including all six items) turned out to be acceptable within the $\text{Dutch}_{\text{DutchSubs}}$ group, the compound variable Attitude Audio was calculated without the item *too informative* since this resulted in the most reliable scale for all conditions in general. It was then used in all further analyses.

Attitude Task was measured using four items: *fun to do*, *easy to do*, *boring to do* and *harder than anticipated.* Again, the items *fun to do* and *easy to do* were negatively recoded to make them consistent with the other two items. The reliability of Attitude Task within the $\text{Dutch}_{\text{DutchSubs}}$ condition comprising four items was acceptable: $\alpha$ = .71. Attitude Task turned out to be unreliable: $\alpha$ = .56 in all other conditions in previous research (Nederstigt & Hilberink-Schulpen, 2019), however, which is why the four items will be analysed separately in further analyses.

Attitude Subtitles was measured with six items: *difficult to understand*, *easy to follow*, *too slow*, *aiding in the task*, *distracting from narration* and *distracting from the animation*. *Easy to follow* and *aiding in the task* were recoded negatively, as the other items were worded negatively as well. Attitude Subtitles was only measured in two of the four conditions, where subtitles were present. The reliability of Attitude Subtitles within the $\text{Dutch}_{\text{DutchSubs}}$ condition, which consisted of six items, was found to be acceptable: $\alpha$ = .78. Nederstigt & Hilberink-Schulpen (2019), however, found that the reliability of Attitude Subtitles, comprising six items, within the $\text{English}_{\text{DutchSubts}}$ condition was inadequate: $\alpha$ = .59. These researchers therefore established two separate concepts regarding the attitude towards subtitles. Some questions pertained to the concept of understanding the subtitles, whereas others related more to the distraction resulting from the subtitles. *Too slow* related to neither of these two concepts and was therefore deleted (Nederstigt & Hilberink-Schulpen, 2019). The

current research therefore calculated these two concepts separately as well for the condition $Dutch_{DutchSubs}$ and comprised them into two separate compound variables, which will be used in all further analyses. The reliability of Subtitles Distraction, comprised of three items, was marginally acceptable within the $Dutch_{DutchSubs}$ condition: $\alpha = .61$, as well as within the $English_{DutchSubs}$ condition: $\alpha = .66$. A significant positive correlation was found between the two questions regarding Subtitles Understandability (*difficult to understand* and *easy to follow*, which was recoded negatively) within the $Dutch_{DutchSubs}$ condition ($r$ (30) = .64, $p < .001$) and within the $English_{DutchSubts}$ condition ($r$ (45) = .40, $p = .007$). The higher the difficulty to understand the subtitles, the harder to follow.

Task Difficulty was measured using two items: *I understood the task* and *I executed the task correctly*. A significant positive correlation was found between the two questions regarding Task Difficulty (*I understood the task* and *I executed the task correctly*) within the $Dutch_{DutchSubs}$ condition ($r$ (30) = .42, $p = .023$) and within the other three conditions ($r$ (118) = .55, $p < .001$). The more a participant understood the task, the more that participant thought they executed the task correctly. As this was also found for the other conditions in the research done by Nederstigt and Hilberink-Schulpen (2019), the mean of the two variables will be used in all further analyses in the remainder of this study.

Participants were also asked to report their perceived Dutch and English proficiency by choosing whether their proficiency was *very good*, *good*, *moderate*, *not good* or *absent*. The English proficiency of all Dutch participants was considered to be sufficient to understand the video and the aim of this study is to first establish the effects of standardisation and adaptation on relatively proficient L2 speakers in general. Highly proficient English L2 speakers could prefer standardisation over adaptation, however, whereas less proficient English L2 speakers might prefer adaptation over standardisation. Therefore, the self-reported English proficiency of participants will be included as a possible moderator and the current research will attempt to make an initial tentative prediction towards the differences between levels of language proficiency. Self-reported proficiency has been shown to correlate with measured proficiency (Shameem, 1998; Marian, Blumenfeld & Kaushanskaya, 2007) and can therefore be used to make this initial prediction.

Participants were also asked to indicate the last time they have built LEGO. Frequent LEGO builders could be more dexterous and thus this factor will be considered to be a possible moderator, which might influence the relationship between the effectivity of the instructional video and the manipulation of language and subtitles. LEGO Dexterity was determined by asking the question *When was the last time you used LEGO?* to which participants were able to answer with a few specific time options.

Finally, participants were asked to report their preference between paper manual instructions and video instructions by answering three questions. Manual Preference was considered to be a

possible moderator as well, since a preference for paper manuals could reduce the effectiveness of the instructional video. Manual Preference was measured using three questions. The question started with *in comparison to a paper manual instruction, the instructional video is…* The participants then rated this comparison on the options *easier*, *more fun* and *more informative* by using a 5-point Likert scale that ranged from 1 (completely agree) to 5 (completely disagree). The reliability of Manual Preference consisting of three items was good: $\alpha = .84$ within the $\text{Dutch}_{\text{DutchSubs}}$ condition, as well as within the other three conditions: $\alpha = .86$. Consequently, the mean of the three items was used to calculate the compound variable Manual Preference, which was used in all further analyses.

*Procedure*

Participants were gathered from within the social circle of the researcher in addition to a large group of participants which had been collected by other researchers beforehand at the NEMO Science Museum in Amsterdam. There was no reward involved. The experiment was done in real life. Participants were either asked to come to an empty classroom within the Radboud University in Nijmegen, or they executed the experiment in a secluded space in their own home under supervision of the researcher. The experiment was done individually and took about fifteen minutes per participant in total. The instructional video took four minutes and 22 seconds to watch, and the questionnaire took about ten minutes. Participants first orally gave their consent to anonymously volunteer in this experiment. Subjects were told that if they wished to quit the experiment, they were allowed to do so any time. Participants were then informed that they were going to watch an instructional video on building a LEGO house. They were asked to imagine they were building an IKEA wardrobe and that it was very important for them to closely follow and execute the exact instructions they were given. The building bricks were already sorted by colour before the participants started watching the video on the researcher's laptop (see Figure 3 for the complete set-up). Participants were asked to build the LEGO house simultaneously while watching the video, without pausing, rewinding or fast-forwarding it. After this process was completed, participants filled in a Qualtrics questionnaire on the researcher's laptop. The questionnaire contained questions about the participants' attitude, their experience with LEGO and some demographic questions regarding their gender, their age, their level of education, their native language and their perceived English and Dutch

*Figure 3. The complete set-up of the experiment*

proficiency (see Appendix A for the full questionnaire). While participants filled in the questionnaire, the researcher filled in a form examining the building performance of the participant per layer of the LEGO house. This form could then later be recoded into a file constructing the dependent variables regarding the performance of the participants. Each participant was given a specific number, used on the form as well as in the questionnaire. This way participants' results could remain anonymous, while the data of the experiment could still be linked to the data of the questionnaire. At the end of the experiment participants were explained the goal of the research if they were interested and thanked for their participation. The procedure was the same for all subjects.

*Statistical Treatment*

Per dependent variable, a two-, three- or four-way analysis of variance was conducted to test the research question and the hypotheses.

## Results

*Performance with Language and Subtitles*

The two-way analysis of variance for Total Performance with between subject factors Language (Dutch versus English) and Subtitles (Dutch subtitles present or not) showed no significant main effect of Language ($F$ (1,144) = 2.96, $p$ = .088, $\eta^2$ = .02) and Subtitles ($F$ (1,144) < 1). There was also no significant interaction ($F$ (1,144) = 2.96, $p$ = .087, $\eta^2$ = .02).

Next, analyses for the three individual coded performance variables, namely correct brick size, correct colour use and correct position, were run.

The two-way analysis of variance for Brick Size Total with between subject factors Language (Dutch versus English) and Subtitles (Dutch subtitles present or not) showed no significant main effect of Language ($F$ (1,144) < 1) and Subtitles ($F$ (1,144) < 1). There was also no significant interaction ($F$ (1,144) = 3.78, $p$ = .054, $\eta^2$ = .03). It should be noted that the homogeneity of variance assumption was violated, however, it is beyond the scope of this thesis to run alternative analyses.

The two-way analysis of variance for Position Total with between subject factors Language (Dutch versus English) and Subtitles (Dutch subtitles present or not) showed no significant main effect of Language ($F$ (1, 144) < 1) and Subtitles ($F$ (1,144) = 2.28, $p$ = .133, $\eta^2$ = .02). There was also no significant interaction ($F$ (1,144) = 3.07, $p$ = .082, $\eta^2$ = .02).

The two-way analysis of variance for Colour Total with between subject factors Language (Dutch versus English) and Subtitles (Dutch subtitles present or not) showed a significant main effect of Language ($F$ (1,144) = 24.42, $p$ < .001, $\eta^2$ = .15). The colour performance was higher for the instructional video with *spoken English* ($M$ = 97.28, $SD$ = 7.46) than the instructional video with *spoken Dutch* ($M$ = 89.42, $SD$ = 11.71). There was no significant main effect of Subtitles ($F$ (1,144) < 1). There was also no significant interaction ($F$ (1,144) < 1). It should be noted that the homogeneity of variance assumption was violated, however, again it is beyond the scope of this thesis to run alternative analyses.

Table 3. Means and standard deviations for the Total Performance, Brick Size Total, Position Total and Colour Total (0 – 100%) per condition (Language (Dutch or English) and Subtitles (Present or Absent))

| Language | Dutch | | English | |
|---|---|---|---|---|
| Subtitles | Present | Absent | Present | Absent |
| | n = 30 | n = 39 | n = 45 | n = 34 |
| | M (SD) | M (SD) | M (SD) | M (SD) |
| Total Performance | 83.87 (13.17) | 89.12 (11.65) | 90.67 (9.37) | 89.12 (13.75) |

| | | | | |
|---|---|---|---|---|
| Brick Size Total | 93.03 (7.85) | 94.91 (9.56) | 97.09 (3.97) | 92.98 (13.95) |
| Position Total | 70.10 (26.10) | 82.34 (18.76) | 77.79 (21.25) | 76.88 (24.79) |
| Colour Total | 88.50 (10.92) | 90.13 (12.38) | 97.11 (6.95) | 97.50 (8.19) |

*Performance with Language, Subtitles and Type of Info*

As stated in the Method section, the video was also manipulated in type of information given per layer. This could be audio information only, visual information only or both. The following analyses include this within subject variable. For all three analyses with Type of Info the assumptions for homogeneity of variance were violated. Where possible alternative F-statistics such as Greenhouse-Geisser and Huynh-Feldt will be reported, any other corrections are beyond the scope of this thesis.

Total Performance

A repeated measures analysis with Type of Info as within-subject factor and Language and Subtitles as between-subject factors for the total performance showed a significant main effect of Type of Info ($F$ (1.27, 183.39) = 9.46, $p$ = .001, $\eta^2$ = .06). Irrespective of the language of the video and the presence or absence of subtitles, the total performance for the audiovisual layers ($M$ = 90.69) was significantly higher than for the audio layers ($M$ = 85.81; $p$ = .006, Bonferroni-correction) and the visual layers ($M$ = 85.63; $p$ < .001, Bonferroni-correction). The total performance for the audio layers and the visual layers did not significantly differ ($p$ = 1.00, Bonferroni-correction). A significant main effect for Language was found as well ($F$ (1, 144) = 6.80, $p$ = .010, $\eta^2$ = .05). The total performance was, irrespective of type of information and the presence or absence of subtitles, higher for the instructional video with spoken English ($M$ = 90.18) than the instructional video with spoken Dutch ($M$ = 84.58). There was no significant main effect for Subtitles ($F$ (1, 144) = 1.48, $p$ = .227, $\eta^2$ = .01). The interaction between Language and Subtitles was not significant ($F$ (1, 144) = 3.09, $p$ = .081, $\eta^2$ = .02), nor was the interaction between Type of Info and Subtitles ($F$ (1.27, 183.39) = 1.91, $p$ = .166, $\eta^2$ = .01) and Type of Info, Language and Subtitles ($F$ (1.27, 183.39) < 1). However, there was a significant interaction between Type of Info and Language ($F$ (1.27, 183.39) = 23.67, $p$ < .001, $\eta^2$ = .14).

Separate analyses were run for each Language to disentangle the significant interaction. A repeated measures analysis with Type of Info as within-subject factor for the *Dutch spoken* videos only for total performance showed a significant main effect of Type of Info ($F$ (1.25, 85.10) = 14.13, $p$ < .001, $\eta^2$ = .17). The total performance for the audiovisual layers ($M$ = 90.99 $SD$ = 11.94) was significantly higher than for the audio layers for the *Dutch spoken* videos ($M$ = 78.50, $SD$ = 26.68; $p$ < .001, Bonferroni-correction) and the visual layers ($M$ = 85.49, SD = 11.72; $p$ < .001, Bonferroni-correction. The total performance for the visual layers was also significantly higher than for the audio layers (p = .040, Bonferroni-correction).

The repeated measures analysis with Type of Info as within-subject factor for the *English spoken* videos only for total performance showed a significant main effect of Type of Info ($F$ (1.33, 103.38) = 17.16, $p < .001$, $\eta^2 = .18$). The total performance for the audiovisual layers ($M = 90.78$, $SD = 11.58$) was significantly higher than for the visual layers for the *English spoken* videos ($M = 86.11$, $SD = 12.13$; $p < .001$, Bonferroni-correction). The total performance for the audio layers ($M = 93.88$, SD = 16.70) was also significantly higher than for the visual layers (p < .001, Bonferroni-correction) for the *English spoken* videos. The audiovisual layers and the audio layers did not significantly differ from each other (p = .158, Bonferroni-correction).

The interaction between Type of Info and Language is apparently due to the effect that in both spoken languages the audiovisual layers outperform the visual layers. However, for the *Dutch spoken* videos the audiovisual layers outperform the audio layers as well whereas this is not the case in the *English spoken* videos. Besides this difference, the visual layers outperform the audio layers in the *Dutch spoken* videos, whereas this is the other way around in the *English spoken* videos.

Table 4. Means and standard deviations for the Total Performance (0 – 100%) per Type of Info (audiovisual layers, visual layers and audio layers)

| Language | Dutch | | English | |
|---|---|---|---|---|
| Subtitles | Present | Absent | Present | Absent |
| | n = 30 | n = 39 | n = 45 | n = 34 |
| | M (SD) | M (SD) | M (SD) | M (SD) |
| Audiovisual layers | 89.24 (11.83) | 92.33 (12.00) | 91.94 (9.56) | 89.26 (13.82) |
| Visual layers | 82.70 (10.38) | 87.65 (12.35) | 86.27 (10.39) | 85.89 (14.28) |
| Audio layers | 72.22 (30.43) | 83.33 (22.62) | 94.07 (14.72) | 93.63 (19.25) |

Colour Performance

A repeated measures analysis with Type of Info as within-subject factor and Language and Subtitles as between-subject factors for colour performance showed a significant main effect of Type of Info ($F$ (1.02, 147.25) = 70.20, $p < .001$, $\eta^2 = .33$). Irrespective of the language of the video and the presence of subtitles, the colour performance for the audiovisual layers ($M = 99.05$, $SD = 4.57$) was significantly higher than for the audio layers ($M = 71.96$, $SD = 44.88$; $p < .001$, Bonferroni-correction). The same holds for the visual layers ($M = 98.99$, SD = 5.58; $p < .001$, Bonferroni-correction). The colour performance for the audiovisual and visual layers did not differ ($p = 1.00$, Bonferroni-correction). A significant main effect for Language ($F$ (1, 144) = 27.05, $p < .001$, $\eta^2 = .16$) was found. The colour performance was, irrespective of type of information and the presence of subtitles, higher for the instructional video with *spoken English* ($M = 95.90$) than the instructional video with *spoken Dutch* ($M = 83.12$). There was no significant main effect for Subtitles ($F$ (1, 144)

< 1). The interaction between Type of Info and Subtitles was not significant ($F$ (1.02, 147.25) = 1.54, $p$ = .217, $\eta^2$ = .01), nor was the interaction between Language and Subtitles ($F$ (1, 144) < 1) and Type of Info, Language and Subtitles ($F$ (1.02, 147.25) < 1). However, the main effect of Type of Info was qualified by a significant interaction between Type of Info and Language ($F$ (1.02, 147.25) = 28.04, $p$ < .001, $\eta^2$ = .16).

To disentangle the significant interaction separate analyses were run for each Language. A repeated measures analysis with Type of Info as within-subject factor for the *Dutch spoken* videos only for colour performance showed a significant main effect of Type of Info ($F$ (1.03, 70.01) = 58.80, $p$ < .001, $\eta^2$ = .46). The colour performance for the audiovisual layers ($M$ = 98.55 $SD$ = 6.00) was significantly higher than for the audio layers for the *Dutch spoken* videos ($M$ = 52.90, $SD$ = 49.91; $p$ < .001, Bonferroni-correction). The same holds for the visual layers ($M$ = 98.55, SD = 6.85; $p$ < .001, Bonferroni-correction. The colour performance for the audiovisual and visual layers did not differ ($p$ = 1.00, Bonferroni-correction).

The repeated measures analysis with Type of Info as within-subject factor for the *English spoken* videos only for colour performance showed a significant main effect of Type of Info ($F$ (1.00, 78.48) = 9.43, $p$ = .003, $\eta^2$ = .11). The colour performance for the audiovisual layers ($M$ = 99.49, $SD$ = 2.72) was significantly higher than for the audio layers for the *English spoken* videos ($M$ = 88.61, $SD$ = 31.97; $p$ = .009, Bonferroni-correction). The same holds for the visual layers ($M$ = 99.37, SD = 4.17; $p$ = .009, Bonferroni-correction). The colour performance for the audiovisual and visual layers did not differ ($p$ = 1.00, Bonferroni-correction).

The interaction between Type of Info and Language is apparently due to the effect that in both spoken languages a similar pattern is found (the audio layer scores lower), but for the *Dutch spoken* video this difference is larger.

Table 5. Means and standard deviations for the Colour Performance (0 – 100%) per Type of Info (audiovisual layers, visual layers and audio layers)

| Language | Dutch | | English | |
|---|---|---|---|---|
| Subtitles | Present | Absent | Present | Absent |
| | n = 30 | n = 39 | n = 45 | n = 34 |
| | M (SD) | M (SD) | M (SD) | M (SD) |
| Audiovisual layers | 99.00 (4.03) | 98.21 (7.21) | 99.78 (1.49) | 99.12 (3.79) |
| Visual layers | 98.89 (4.23) | 98.29 (8.37) | 99.63 (2.48) | 99.02 (5.72) |
| Audio layers | 46.67 (50.74) | 57.69 (49.39) | 86.67 (34.38) | 91.12 (28.79) |

Brick Size Performance

A repeated measures analysis with Type of Info as within-subject factor and Language and Subtitles as between-subject factors for brick size performance showed a significant main effect of Type of Info ($F$ (1.36, 195.66) = 7.71, $p$ = .003, $\eta^2$ = .05). Irrespective of language of the video and the presence of subtitles, the brick size performance for the audiovisual layers ($M$ = 95.06, $SD$ = 9.84) was significantly higher than for the visual layers ($M$ = 92.76, $SD$ = 9.32; $p$ < .001, Bonferroni-correction). The same holds for the audio layers ($M$ = 96.96, SD = 16.73; $p$ = .001, Bonferroni-correction). The brick size for the audiovisual and audio layers did not differ ($p$ = .501, Bonferroni-correction). There was no significant main effect for Language ($F$ (1, 144) < 1), nor for Subtitles ($F$ (1, 144) < 1). The interaction between Type of Info and Subtitles was not significant ($F$ (1.36, 195.66) < 1), nor was the interaction between Language and Subtitles ($F$ (1, 144) = 3.36, $p$ = .069, $\eta^2$ = .02) and Type of Info and Language ($F$ (1.36, 195.66) = 1.72, $p$ = .190, $\eta^2$ = .01). Lastly, the interaction between Type of Info, Language and Subtitles ($F$ (1.36, 195.66) < 1) was also not significant.

Table 6. Means and standard deviations for the Brick Size Performance (0 – 100%) per Type of Info (audiovisual layers, visual layers and audio layers)

| Language | Dutch | | English | |
|---|---|---|---|---|
| Subtitles | Present | Absent | Present | Absent |
| | n = 30 | n = 39 | n = 45 | n = 34 |
| | M (SD) | M (SD) | M (SD) | M (SD) |
| Audiovisual layers | 94.31 (8.55) | 95.01 (10.55) | 97.39 (5.65) | 92.69 (13.51) |
| Visual layers | 90.68 (6.37) | 93.91 (8.54) | 94.65 (4.88) | 90.76 (14.87) |
| Audio layers | 93.33 (25.37) | 96.15 (17.71) | 100 (.00) | 97.06 (17.15) |

Position Performance

A repeated measures analysis with Type of Info as within-subject factor and Language and Subtitles as between-subject factors for position performance showed a significant main effect of Type of Info ($F$ (1.63, 234.60) = 66.03, $p$ < .001, $\eta^2$ = .31). Irrespective of language of the video and the presence of subtitles, the position performance for the audiovisual layers ($M$ = 78.52, $SD$ = 24.89) was significantly higher than for the visual layers ($M$ = 65.72, $SD$ = 28.08; $p$ < .001, Bonferroni-correction). The same holds for the audio layers ($M$ = 91.22, SD = 27.80; $p$ < .001, Bonferroni-correction). The position performance for the audio layers was also significantly higher than for the audiovisual layers ($p$ < .001, Bonferroni-correction). There was no significant main effect for Language ($F$ (1, 144) < 1), nor for Subtitles ($F$ (1, 144) = 3.11, $p$ = .080, $\eta^2$ = .02). The interaction between Type of Info and Subtitles was not significant ($F$ (1.63, 234.60) < 1), nor was the interaction

between Language and Subtitles ($F$ (1, 144) = 3.86, $p$ = .051, $\eta^2$ = .03) and Type of Info and Language ($F$ (1.63, 234.60) = 2.47, $p$ = .098, $\eta^2$ = .02). Lastly, the interaction between Type of Info, Language and Subtitles ($F$ (1.63, 234.60) = 1.34, $p$ = .262, $\eta^2$ = .01) was also not significant.

Table 7. Means and standard deviations for the Position Performance (0 – 100%) per Type of Info (audiovisual layers, visual layers and audio layers)

| Language | Dutch | | English | |
|---|---|---|---|---|
| Subtitles | Present | Absent | Present | Absent |
| | n = 30 | n = 39 | n = 45 | n = 34 |
| | M (SD) | M (SD) | M (SD) | M (SD) |
| Audiovisual layers | 74.42 (26.23) | 83.77 (21.70) | 78.65 (24.16) | 75.97 (27.93) |
| Visual layers | 58.52 (28.90) | 70.74 (28.35) | 64.53 (27.11) | 67.89 (28.08) |
| Audio layers | 76.67 (43.02) | 96.15 (17.71) | 95.56 (20.84) | 92.65 (25.02) |

*Attitude with Language and Subtitles*

The attitude was measured using five different dependent variables. For each dependent variable an ANOVA was run with Language and Subtitles as independent variables.

Task Difficulty

A two-way analysis of variance on Task Difficulty with between-subject factors Language and Subtitles did not show a significant main effect for Language ($F$ (1, 144) < 1), nor for Subtitles ($F$ (1, 144) < 1). The interaction was also not significant ($F$ (1, 144) < 1).

Attitude Video

A two-way analysis of variance on Attitude Video with between-subject factors Language and Subtitles did not show a significant main effect for Language ($F$ (1, 144) < 1), nor for Subtitles ($F$ (1, 144) < 1). The interaction was also not significant ($F$ (1, 144) < 1).

Attitude Audio

A two-way analysis of variance on Attitude Audio with between-subject factors Language and Subtitles did not show a significant main effect for Language ($F$ (1, 144) < 1), nor for Subtitles ($F$ (1, 144) < 1). The interaction was also not significant ($F$ (1, 144) < 1).

Attitude Task

A two-way analysis of variance on Attitude Task Q1 with between-subject factors Language and Subtitles did not show a significant main effect for Language ($F$ (1, 144) < 1), nor for Subtitles ($F$ (1, 144) < 1). The interaction was also not significant ($F$ (1, 144) = 1.42, $p$ = .235, $\eta^2$ = .01).

A two-way analysis of variance on Attitude Task Q2 with between-subject factors Language and Subtitles did not show a significant main effect for Language ($F$ (1, 144) < 1), nor for Subtitles ($F$ (1, 144) < 1). The interaction was also not significant ($F$ (1, 144) = 1.42, $p$ = .235, $\eta^2$ = .01).

A two-way analysis of variance on Attitude Task Q3 with between-subject factors Language and Subtitles did not show a significant main effect for Language ($F$ (1, 144) = 1.37, $p$ = .243, $\eta^2$ = .01), nor for Subtitles ($F$ (1, 144) < 1). The interaction was also not significant ($F$ (1, 144) = 1.77, $p$ = .185, $\eta^2$ = .01). It should be noted that the homogeneity of variance assumption was violated, however, it is beyond the scope of this thesis to run alternative analyses.

A two-way analysis of variance on Attitude Task Q4 with between-subject factors Language and Subtitles did not show a significant main effect for Language ($F$ (1, 144) < 1), nor for Subtitles ($F$ (1, 144) < 1). The interaction was also not significant ($F$ (1, 144) < 1). It should be noted that the homogeneity of variance assumption was violated, however, again it is beyond the scope of this thesis to run alternative analyses.

Attitude Subtitles

A two-way analysis of variance on Subtitles Understandability with between-subject factors Language did not show a significant main effect for Language ($F$ (1, 73) < 1).

The two-way analysis of variance on Subtitles Distraction with between-subject factors Language also did not show a significant main effect for Language ($F$ (1, 73) = 3.15, $p$ = .080, $\eta^2$ = .04).

Table 8. Means and standard deviations for the Attitude (measured on a 5-point Likert scale, 1 = a high attitude and 5 = a low attitude) per condition (Language (Dutch or English) and Subtitles (Present or Absent))

| Language | Dutch | | English | |
|---|---|---|---|---|
| Subtitles | Present | Absent | Present | Absent |
| | n = 30 | n = 39 | n = 45 | n = 34 |
| | M (SD) | M (SD) | M (SD) | M (SD) |
| Task Difficulty | 1.78 (.90) | 1.74 (.97) | 1.79 (.88) | 1.65 (.84) |
| Attitude Video | 2.17 (.81) | 2.22 (.88) | 2.16 (.85) | 2.14 (.66) |
| Attitude Audio | 2.29 (.80) | 2.29 (.89) | 2.18 (.68) | 2.31 (.81) |
| Attitude Task Q1 | 1.77 (.94) | 1.62 (.63) | 1.49 (.66) | 1.65 (.88) |

| | | | | |
|---|---|---|---|---|
| Attitude Task Q2 | 2.30 (1.09) | 1.97 (1.01) | 1.87 (1.08) | 2.12 (1.15) |
| Attitude Task Q3 | 1.90 (.89) | 2.21 (1.22) | 2.36 (1.15) | 2.18 (1.03) |
| Attitude Task Q4 | 2.83 (1.09) | 2.95 (1.41) | 2.82 (1.50) | 2.85 (1.31) |
| Subtitles Understandability | 2.33 (1.04) | | 2.47 (.84) | |
| Subtitles Distraction | 2.66 (.89) | | 3.03 (.90) | |

*Analyses combining the attitude data with the performance data*

Some additional analyses were run linking the experimental LEGO data with the attitudinal data retrieved from the questionnaire. Only relevant and/or significant effects will be reported to reduce the length of the results section. Significant main effects will not be explained when significant interactions are present for the same reason. When a three-way or four-way significant interaction is found, significant two-way interactions that involve the same variables will not be reported either. Again, for the analyses with Type of Info the assumptions for homogeneity of variance were violated. Where possible alternative F-statistics such as Greenhouse-Geisser and Huynh-Feldt will be reported, any other corrections are beyond the scope of this thesis.

Colour Performance and Subtitles Understandability

To see whether how well participants understood the subtitles influenced their performance, the participants were divided into two groups based on their scores to the Subtitles Understandability questions in the questionnaire. This was done based on the calculated mean (M = 2.41). Each participant scoring lower than this mean was placed in the group 'high understandability' and each participant scoring higher was placed in the group 'low understandability'. Next, analyses were run with Subtitles Understandability as between subject variable to see whether there were any interactions with this variable.

The repeated measures analysis with Type of Info as within-subject factor and Language and Subtitles Understandability as between-subject factors for colour performance showed a significant main effect of Type of Info ($F$ (1.01, 71.66) = 47.78, $p < .001$, $\eta^2 = .40$) and Language ($F$ (1, 71) = 17.76, $p < .001$, $\eta^2 = .20$). The main effect of Type of Info was also qualified by a significant interaction between Type of Info, Language and Subtitles Understandability ($F$ (1.01, 71.66) = 4.53, $p = .012$, $\eta^2 = .06$).

To unravel the significant three-way interaction separate analyses were run for both categories of Subtitles Understandability (low and high). A repeated measures analysis with Type of Info as within-subject factor and Language as between-subject factor for *high* understandability only did not show a significant interaction between Type of Info and Language ($F$ (1.01, 30.18) = 1.27, $p = .268$, $\eta^2 = .04$).A repeated measures analysis with Type of Info as within-subject factor and Language as

between-subject factor for *low* understandability only showed a significant main effect of Type of Info ($F$ (1.01, 41.53) = 44.51, $p < .001$, $\eta^2 = .52$). The analysis also showed a significant interaction between Type of Info and Language ($F$ (1.01, 41.53) = 25.16, $p < .001$, $\eta^2 = .38$).

To figure out the significant interaction between Type of Info and Language for *low* understandability, separate analyses were run for each Language. A repeated measures analysis with Type of Info as within-subject factor for the *English spoken* video only did not show a significant main effect of Type of Info ($F$ (1.01, 28.07) = 2.90, $p = .100$, $\eta^2 = .09$). A repeated measures analysis with Type of Info as within-subject factor for the *Dutch spoken* video did show a significant main effect of Type of Info ($F$ (1.02, 13.30) = 31.87, $p < .001$, $\eta^2 = .71$). The colour performance for the audio layers ($M = 28.57$) was significantly lower than for the audiovisual layers ($M = 97.86$; $p < .001$, Bonferroni-correction) and the visual layers (M = 98.81; p < .001, Bonferroni-correction). The colour performance for the audiovisual and visual layers did not differ ($p = 1.00$, Bonferroni-correction).

The interaction between Type of Info, Language and Subtitles Understandability is apparently due to the effect that for the *Dutch spoken* video the audio layer scores lower when the understandability regarding the subtitles is low, while this is not the case for the *English spoken* video.

Position Performance and Attitude Video

To check whether the participants' attitude towards the video influenced their performance, the participants were split into two groups based on the calculated mean (M = 2.17) of the Attitude Video questions in the questionnaire. Each participant scoring lower than this mean was placed in the group 'high attitude video' and each participant scoring higher was placed in the group 'low attitude video'. Next, analyses were run with Attitude Video as between subject variable to investigate whether there were any interactions with this variable.

A repeated measures analysis with Type of Info as within-subject factor and Language, Subtitles and Attitude Video as between-subject factors for position performance showed a significant main effect of Type of Info ($F$ (1.69, 236.29 = 68.35, $p < .001$, $\eta^2 = .33$) and Attitude Video ($F$ (1, 140) = 8.73, $p = .004$, $\eta^2 = .06$). A significant interaction between Type of Info, Language, Subtitles and Attitude Video was found as well ($F$ (1.69, 236.29) = 3.93, $p = .027$, $\eta^2 = .03$).

To disentangle the significant four-way interaction separate analyses were run for both categories of Attitude Video (low and high). A repeated measures analysis with Type of Info as within-subject factor and Language and Subtitles as between-subject factors for *high* Attitude Video only showed a significant main effect of Type of Info ($F$ (1.69, 119.72) = 27.96, $p < .001$, $\eta^2 = .28$). The interaction between Language and Subtitles was significant as well ($F$ (1.01, 30.18) = 1.27, $p = .268$, $\eta^2 = .04$). A repeated measures analysis with Type of Info as within-subject factor and Language as between-subject factor for *low* Attitude Video only also showed a significant main effect of Type

of Info ($F$ (1.70, 117.14) = 40.29, $p < .001$, $\eta^2 = .37$). The main effect of Type of Info was qualified by a significant interaction between Type of Info, Language and Subtitles ($F$ (1.70, 117.14) = 4.12, $p = .024$, $\eta^2 = .06$).

Apparently the four-way interaction is due to the fact that only for the participants with a *low* attitude regarding the video a significant three-way interaction between Type of Info, Language and Subtitles exists. Therefore, only this three-way interaction for *low* Attitude Video will be explored further.

For the significant three-way interaction between Type of Info, Language and Subtitles for *low* Attitude Video, separate analyses were run for subtitle presence and absence. A repeated measures analysis with Type of Info as within-subject factor and Subtitles as between-subject factor for *Dutch subtitles absent* only showed a significant main result of Type of Info ($F$ (1.76, 63.32) = 25.01, $p < .001$, $\eta^2 = .41$) and no significant interaction between Type of Info and Language ($F$ (1.76, 63.32) = 1.34, $p = .269$, $\eta^2 = .04$). A repeated measures analysis with Type of Info as within-subject factor and Language as between-subject factor for *Dutch subtitles present* also showed a significant main result of Type of Info ($F$ (1.58, 52.23) = 16.67, $p < .001$, $\eta^2 = .34$). In this case the analysis did show a significant interaction between Type of Info and Language ($F$ (1.58, 52.23) = 4.86, $p = .017$, $\eta^2 = .13$). To disentangle this interaction even further, separate analyses were run for each Language. A repeated measures analysis with Type of Info as within-subject factor for the *Dutch spoken* video showed a significant main effect of Type of Info ($F$ (1.40, 20.95) = 4.19, $p = .042$, $\eta^2 = .22$). The position performance for the audiovisual layers (M = 72.41) was significantly higher than for the visual layers (M = 52.80; p = .002, Bonferroni-correction). The position performance between the audio layers (M = 68.75) and the audiovisual layers did not differ (p = 1.00, Bonferroni-correction), neither did the position performance between the audio layers and the visual layers (p = .134, Bonferroni-correction). A repeated measures analysis with Type of Info as within-subject factor for the *English spoken* video showed a significant main effect of Type of Info ($F$ (1.45, 26.17) = 18.27, $p < .001$, $\eta^2 = .50$) as well. The position performance for the audio layers (M = 94.74) was significantly higher than for the audiovisual layers (M = 69.95; p = .010. Bonferroni-correction) and the visual layers (M = 54.01; p < .001, Bonferroni-correction). The position performance for the audiovisual layers was significantly higher than for the visual layers as well (p = .005, Bonferroni-correction).

The interaction between Type of Info, Language, Subtitles and Attitude Video is apparently due to the effect that for position performance when the attitude regarding the video is *low* and *subtitles* are *present*, the audiovisual layers scored higher than the visual layers in both the *Dutch spoken* and the *English spoken* video but this effect is stronger for the *English spoken* video as well

as to the effect that in the *English spoken* video, the audio layers outperformed both other layers whereas this was not the case in the *Dutch spoken* video.

Performance and Language Proficiency

Similar analyses for different language proficiency groups could not be run because in some conditions participants were too homogeneous concerning their proficiency and in some conditions data was missing.

## Discussion

The current study analysed whether the choice of language (L1 versus L2) and the presence of L1 subtitles influence the effectiveness of instructional videos. Dutch participants were investigated and the languages that were looked at were Dutch (L1) and English (L2). The expectation was that the use of a spoken L1 in instructional videos would be more effective than the use of a spoken L2 and that the presence of L1 subtitles, in contrast to no subtitles, would affect the effectiveness of instructional videos. The direction of this effect was not specified for L1 subtitles supporting a spoken L2, however, the presence of same language subtitles (SLS) was expected to improve the effectivity of instructional videos over the absence of subtitles. According to the results of the analyses, the hypotheses were largely not supported. The findings generally show that the use of a spoken L1 in comparison to a spoken L2 does not lead to differences in effectiveness of instructional videos and that the presence of L1 subtitles in contrast to the absence of subtitles does not impact the effectiveness of instructional videos either. This suggests that adaptation does not lead to noticeable differences in the effectivity of instructional videos, which would be an argument in favour of standardisation.

The expectation was that the use of a spoken L1 would lead to a higher effectivity of instructional videos, as opposed to the use of a spoken L2 (H1). Generally speaking, the results seem to contradict this expectation in the sense that the effectivity did not differ between both languages. However, the use of a spoken L2, as opposed to a spoken L1, did improve the colour performance. This suggests that in some areas the use of an L2 is more effective than the use of a spoken L1 in instructional videos, which is the opposite of the expectation of the current study. This expectation was based on previous research within the field of education, as well as the logical belief that consumers should have a better understanding of their L1 than their L2. The results of the current study, however, indicate that consumers may sometimes prefer to receive instructions in an L2 over an L1. The specific case in this study is the fact that the L2 turned out to be especially beneficial for the audio layers, for which participants had to listen closely, after which they consciously had to make the choice to use the colour presented verbally instead of the colour they perceived visually. The L2 thus seems to improve this conscious decision-making process, whereas the L1 does not. The L2 advantage effect may be at work in these specific cases. This theory proposes that the use of an L2 may lead to a reduction of automatic processes, resulting in more conscious decisions (Pavlenko, 2012). Earlier research has also shown that in product advertisements, the L2 is more effective to promote utilitarian products than the L1 (Caldwell-Harris & Aycicegi-Dinn, 2016). These theories suggest that the L2 may be more effective within the context of receiving practical information, like an instructional video. In specific cases a spoken L2 may therefore be preferred over a spoken L1,

which would again be an argument in favour of standardisation. Future research could investigate this matter to determine in what ways a spoken L2 could be more beneficial than a spoken L1 within the field of instructional videos.

The dual-coding hypothesis proposes that the comprehension of a video increases when both visual *and* verbal elements are used, as opposed to using only one of the two components (Paivio, 1990). In this study, for spoken L1 (Dutch), the dual-coding hypothesis indeed seems to be at work as the audiovisual layers outperformed both the visual and the audio layers in terms of total performance. However, the results for spoken L2 (English) only seem partly in line with the dual-coding hypothesis. The audiovisual layers did improve the total performance over the visual layers, which is in line with the dual-coding hypothesis. The total performance of the audio layers, however, did not differ from the audiovisual layers when the video was presented in the L2, which seems to contradict what the dual-coding hypothesis proposes. Even though the audiovisual layers should have been the most effective according to the dual-coding hypothesis, the earlier discussed L2 advantage effect may play a role in the effectiveness of the audio in an L2 in instructional videos.

Even though the presence of L1 subtitles, as opposed to the absence of those subtitles, was expected to affect the effectivity of instructional videos, the results of this study suggest there is no difference between adding subtitles and omitting them. This was found for L1 subtitles accompanying L2 spoken language (H3) as well as for L1 subtitles in combination with L1 spoken language (H2). According to the cognitive load (CL) theory, subtitles are demanding for extraneous CL, leading to a reduction of the available germane CL that can be used to actually understand the material. CL theory thus seems to suggest that the presence of subtitles is detrimental for instructional videos. The results of the current study, however, do not show any negative effects of the addition of L1 subtitling as opposed to omitting these subtitles, thus contradicting CL theory.

Looking specifically at *L1 subtitles accompanying a spoken L2*, Vulchanova, Aurstad, Kvitnes and Esthuis (2015) found that L1 subtitles significantly improved the comprehension of Norwegian students watching an English spoken cartoon over no subtitles. Their results seem to imply that L1 subtitles supporting a spoken L2 do indeed *not* lead to an increase in CL. Their research therefore seems to suggest that adding L1 subtitles may help the viewer to process and comprehend the spoken L2. However, the current results do not show an increase in effectivity of instructional videos when L1 subtitles are added to L2 spoken language at all, meaning that L1 subtitles do not seem to improve the comprehension of an L2 in instructional videos either. This could be due to the fact that subtitles may be neglected in instructional videos specifically, as instructional videos do not only require watching but also building the product. This building process may take up quite a lot of extraneous CL, leaving relatively less cognitive space to process the subtitles. The subtitles may have also been neglected because of the design of the experiment. Participants were not allowed to pause

the video, which may have forced participants to concentrate on the building process even more, leaving too little cognitive space to process the subtitles at all. Future research could therefore possibly give more conclusive results on the effect of the addition of subtitles by investigating the ability to pause and replay the instructional video. Real life conditions are reflected better when rewinding as well as being able to pause the instructional video are allowed and thus, this way, more insight into the exact workings of subtitles in instructional videos and its potential relation to CL theory could be gained.

Contradictory to what CL theory proposes, previous research also seems to suggest that *Same Language Subtitles* are useful for improving comprehension. Hirose and Kamei (1993) and Neuman and Koskinen (1992) investigated bilingual speakers with English as their L2 and found that English subtitles accompanying spoken English led to a higher comprehension than when subtitles were absent. As discussed, the results of the current research go against CL theory as well as against the results of previous research on SLS. This could be due to the fact that previous research looked at SLS in the L2, whereas the current research focused on SLS in the L1. SLS in the L2 may help to improve understanding and the processing of information, whereas SLS in the L1 may lead to redundancy, affecting the processing of those subtitles. However, since the results also contradict CL theory, and therefore also the redundancy effect that it proposes, the current research suggests that within instructional videos subtitles are not as demanding for extraneous CL as thought, which might be due to the aforementioned neglection of the subtitles in response to the demanding nature of the building process, or that subtitles are processed in a different way than CL theory (initially) proposes.

The redundancy effect as well as the split-attention effect both follow from CL theory and the dual-coding hypothesis. The redundancy effect occurs when language information is given in written *and* spoken form simultaneously, whereas the split-attention effect occurs when attention needs to be divided over animation and written text at the same time. Both effects are demanding for extraneous CL, leaving relatively less cognitive room for germane CL (Sweller, Ayres & Kalyuga, 2011; Sweller, 2002; Moreno & Mayer, 2002). According to these effects, subtitles should increase cognitive load, which lowers the effectiveness of the instructional video. The results of this study seem to contradict the prediction regarding subtitles that both the split-attention effect and the redundancy effect suggest. The combination of the attitudinal data with the experimental data could shed some more light on this matter. Even though the attitudinal data did not yield any significant results, combining this data with the experimental data enabled a more profound investigation of CL theory. The use of two distinct, separate methodological ways of measuring the same phenomenon thus seems to prove helpful. The results of this combination seem to suggest that for colour performance, consumers who find the subtitles difficult to understand, experience a redundancy effect when SLS are present, whereas the presence of L1 subtitles in addition to an L2 did not cause this effect. For SLS, the audio layers led

to a decrease in colour performance while this was not the case for L1 subtitles accompanying the L2. The fact that the same verbal information must be processed twice, once orally and once visually through the subtitles, seems to be a problem for SLS and not for L1 subtitles in addition to an L2 when looking at colour performance and consumers that had difficulty understanding the subtitles. The audiovisual layers and visual layers performed similarly, however, which might indicate that the addition of visuals helps to overcome the redundancy effect. A similar result was found for position performance, which again led to the implication that for consumers who dislike the video, SLS are processed as more redundant than L1 subtitles accompanying an L2. The audiovisual layers outperformed the visual layers a lot stronger when L1 subtitles were added to an L2 than when SLS were present, indicating that processing the language verbally and visually at the same time is harder in the latter case. In addition, the audio layers led to an increase in position performance when L1 subtitles were accompanying an L2, whereas this was not the case for L1 subtitles in addition to an L1. This again suggests that when consumers hear an L2 while they must process L1 subtitles, this does not lead to a redundancy effect when looking at position performance and consumers that dislike the video. The conclusions that can be drawn from these results should be interpreted with care, however, as they only comprise colour performance and position performance. Another potential limitation regarding the audio layers will be discussed towards the end of this discussion, as it may have affected the previously mentioned argument as well as parts of the discussion yet to follow.

The redundancy effect thus seems to apply only when SLS are used. As proposed earlier, the effectivity of L1 subtitles supporting a spoken L2 is a bit of a grey area. This research tentatively suggests that the redundancy effect is weaker or non-existent in this case because the information that must be processed twice is not fully identical, whereas it is when SLS are present. Perhaps a third mechanism, in addition to the verbal and the visual mechanism of the dual-coding hypothesis is used to process L1 subtitles supporting a spoken L2. This mechanism may process visual language separately from other visuals and spoken language. Another possible explanation could be that the verbal mechanism is used to process the written L1, while the spoken L2 is not processed at all due to it not being one's mother tongue.

A first tentative explanation of the processing of SLS is the fact that the redundancy effect may always arise when SLS are present, but it only seems to affect people in specific situations such that it affects the effectivity of only a certain part of the instructional video. This matches with the fact that the current research seems to suggest that the redundancy effect leads to cognitive overload in specific cases only, namely when participants dislike the video or find the subtitles hard to understand. The effect could also be explained in the opposite direction, meaning that the redundancy effect leads to cognitive overload for specific participants only, leading them to have a negative attitude towards the video or the subtitles. Consumers with a lower cognitive capacity beforehand,

may have a harder time understanding the subtitles due to this lower cognitive capacity or because this lower cognitive capacity causes cognitive overload. In both instances more cognitive space is already used up by intrinsic CL and less cognitive space is available for germane CL. Since germane CL is used to understand the given material, this leads to a decrease in effectiveness. Besides this, consumers who dislike the instructional video may have a hard time motivating themselves to understand the video, meaning that intrinsic CL is taking up more cognitive space. Conversely, specific participants could have experienced cognitive overload regardless, for example because of a lower cognitive capacity beforehand, leading to a negative video attitude. Again, in either explanation relatively less space is left for germane CL to understand the given information. The current research cannot show the directionality of the causation. Therefore, more research is necessary to provide more conclusive results regarding the presence or absence of the redundancy effect of L1 subtitles in addition to an L1 versus an L2, as well as to gain more insight into the exact workings of the differences in processing SLS in comparison to L1 subtitles accompanying a spoken L2 in relation to cognitive load.

The fact that the hypotheses of the current research were not borne out, could also be due to the design of the study or other additional factors playing a role in the domain of instructional videos. The instructional video used in the experiment contained three separate ways of giving information following each other. Each viewer was forced to switch between receiving audiovisual information and receiving visual or verbal information only. Even though this enabled further investigation of the discussed theories, it should be taken into account that each type of information change could have led to confusion or stress regardless of language choice and the addition of subtitles. This in turn may have had an effect on how well participants performed right after the transition of the type of information given. Further research could therefore attempt to replicate this study and instead use a between-subject design in which the three types of information are divided over separate groups.

Another limitation includes the fact that the audio condition was more limited than the other two conditions as it only consisted of two layers, whereas the audiovisual and the visual condition consisted of five and three layers respectively. The audio condition was also not completely verbal as visuals were still included, although providing wrong color information. Viewers were visually presented with a different colour than told verbally. This could have led to confusion, regardless of language and the addition of subtitles. However, the results still show that the $Dutch_{AudioOnly}$ condition performed worse on these layers than the $English_{AudioOnly}$ condition and that the $English_{DutchSubs}$ condition outperformed the $Dutch_{DutchSubs}$ condition as well. This means that apparently some factor in the $Dutch_{AudioOnly}$ condition as well as in the $Dutch_{DutchSubs}$ condition causes people to have a harder time to actually listen to the given information than when this information is given in *spoken English*. This factor may be the aforementioned L2 advantage effect. In addition, the previously discussed

argument surrounding the workings of the redundancy effect may also be a factor for the found difference between the two conditions with subtitles. The abovementioned discussion on H1 as well as the addition to CL theory resulting from the arguments surrounding colour performance therefore remain relevant regardless of this limitation. However, the given arguments surrounding the position performance, as well as the fact that for brick size no results were found, should be interpreted with care because the position and the size of the bricks were shown visually rather than verbally within the audio layers. This was due to the fact that the audio layers were designed surrounding colour performance.

Finally, an important suggestion for future research stems from the fact that this study was ultimately not able to investigate the effects of L2 proficiency. Dutch participants were presumed to have at least a basic understanding of the English language as 90% of the Dutch population speaks English as a foreign language (Eurobarometer, 2012; Edwards, 2014). The aim of this study was to establish the effects of standardisation and adaptation on relatively proficient L2 speakers, however, it would have been interesting to make an initial prediction on the differences between levels of L2 proficiency. Subtitles may for example prove more helpful for less proficient L2 speakers, whereas they may be perceived as a distraction for more proficient speakers. The current study was not able to explore this matter as some data regarding L2 proficiency from the previously gathered attitudinal data was missing and most groups turned out to have a very similar (relatively high) proficiency. Future research could therefore analyse the effects of L2 proficiency on language choice and subtitles presence. It would also be interesting to test whether adaptation or standardisation is preferred within countries with a lower general L2 proficiency than the Netherlands.

As opposed to previous studies and theories, this study suggests that in general instructional videos do neither become more effective by changing the choice of language (L1 versus L2), nor by adding or omitting L1 subtitles. This would be an argument in favour of standardisation within relatively proficient L2 countries, as the extra effort of adaptation does not seem to lead to any benefits. The findings do imply a slight advantage of spoken L2 over spoken L1 in some areas, however, meaning that the L2 advantage effect may play a role in instructional videos. These results give a slight indication that standardisation may even be preferred over adaptation in some specific cases. The current research has also added to the existing body of literature on the dual-coding hypothesis, CL theory and the redundancy effect. The findings imply that for specific parts of the instructional video, the redundancy effect may only affect consumers who have a negative attitude towards specific elements in the instructional video *and* when SLS are used, as opposed to L1 subtitles accompanying L2 language. The former therefore may show that the redundancy effect only leads to cognitive overload when a negative attitude is present, whereas the latter seem to be processed in a different way than thought. However, more research is necessary to examine this assumption.

References

Anton, M., & DiCamilla, F. J. (1999). Socio-Cognitive Functions of L1 Collaborative Interaction in the L2 Classroom. *The Modern Language Journal*, *83*(2), 233–247.

Bruhlmann, A. (2012). Does the L1 have a role in the foreign language classroom? A review of the literature. *Working Papers in Applied Linguistics and Tesol, 12*(2), 55–80. https://doi.org/10.7916/D89C791F

Caldwell-Harris, C. & Aycicegi-Dinn, A. (2016). Emotionality differences between a native and foreign language: Implications for cultural marketing strategies. *Journal of Cultural Marketing Strategy*, *2*(1), 9–20.

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2016). Comparing apples and oranges? A critical look at research on learning from statics versus animations. *Computers & Education*, *102*, 234-243. https://doi.org/10.1016/j.compedu.2016.09.004

Crystal, D. (2003). *English as a global language*. New York: Cambridge University Press.

Diao, Y., Chandler, P., & Sweller, J. (2007). The effect of written text on comprehension of spoken English as a foreign language. *The American Journal of Psychology*, *120*(2), 237-261.

*Education First*. (2020). *EF English Proficiency Index*. Ef.nl. https://www.ef.nl/epi/

Edwards, A. (2014). *English in the Netherlands: Functions, forms and atttiudes.* Cambridge: Cambridge University Press.

European Commission. (2014). *Special Eurobarometer 243: Europeans and their languages*. data.europa.eu. https://data.europa.eu/data/datasets/s518_64_3_ebs243?locale=en

Gök, O., Ersoy, P., & Börühan, G. (2019). The effect of user manual quality on customer satisfaction: the mediating effect of perceived product quality. *Journal of Product & Brand Management*, *28*(4), 475-488.

Herron, C., Morris, M., Secules, T., & Curtis, L. (1995). A Comparison Study of the Effects of Video-Based versus Text-Based Instruction in the Foreign Language Classroom. *The French Review*, *68*(5), 775-795.

Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press Inc.

Kruger, J. L., Hefer, E., & Matthew, G. (2013). Measuring the Impact of Subtitles on Cognitive Load: Eye Tracking and Dynamic Audiovisual Texts. *Proceedings of the 2013 Conference on Eye Tracking South Africa*. ACM Digital Library.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research, 50*. 940-967.

Mayer, R. E., & Anderson, R. B. (1991). Animations Need Narrations: An Experimental Test of a Dual-Coding Hypothesis. *Journal of Educational Psychology*, *83*(4), 484-490.

Mekheimer, M. A. A. G. (2011). The Impact of Using Videos on Whole Language Learning in EFL Context. *Arab World English Journal*, *2*(2), 5-39.

Moreno, R., & Mayer, R. E. (2002). Learning Science in Virtual Reality Multimedia Environments: Role of Methods and Media. *Journal of Educational Psychology, 94*(3), 598-610.

Namaziandost, E., Esfahani, F. R., & Hashemifarnia, A. (2018). The Impact of Using Authentic Videos on Prosodic Ability among Foreign Language Learners. *International Journal of Instruction*, *11*(4), 375-390.

Nederstigt, U., & Hilberink-Schulpen, B.J.H. (2019). *The use of English in instruction videos in non-English speaking countries.* Paper presented at DICOEN, Faculty of Arts at KU Leuven, Leuven.

Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, *27*(1). 95-106.

Paivio, A. (1990). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology*, *47*(6), 405–428. https://doi.org/10.1080/00207594.2012.743665

Tumbull, M. (1999). Multidimensional Project-Based Teaching in French Second Language (FSL): A Process-Product Case Study. *The Modern Language Journal. 83*(4), 548-568.

Selnes, F. (1993). An Examination of the Effect of Product Performance on Brand Reputation, Satisfaction and Loyalty. *European Journal of Marketing*, *27*(9), 19-35.

Shameem, N. (1998). Validating self-reported language proficiency by testing performance in an immigrant community: the Wellington Indo-Fijians. *Language Testing*, *15*(1), 86-108.

Sweller, J. (2002). Visualisation and Instructional Design.

Sweller, J. (2011). Cognitive Load Theory. In J.P. Mestre, & B.H. Ross (Eds.), *Psychology of Learning and Motivation* (55, pp. 37-76). Cambridge: Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Measuring Cognitive Load. In J. Sweller, P. Ayres, S. Kalyuga (Eds.), *Cognitive Load Theory* (1, pp. 71-85). New York: Springer. https://doi.org/10.1007/978-1-4419-8126-4

Vulchanova, M., Aurstad, L. M., Kvitnes, I. E., & Eshuis, H. (2015). As naturalistic as it gets: subtitles in the English classroom in Norway. *Frontiers in Psychology, 5.* https://doi.org/10.3389/fpsyg.2014.01510

# Appendix A

Q1. Deelnemersnummer:

Q2. Wat vond je van de taak? Ik heb de taak…

| | | |
|---|---|---|
| helemaal begrepen | ○ ○ ○ ○ ○ | helemaal niet begrepen |
| helemaal goed uitgevoerd | ○ ○ ○ ○ ○ | helemaal niet goed uitgevoerd |

Geef voor de volgende vragen aan wat je mening het beste weergeeft.

Q3. Ik vond de instructievideo

| | helemaal eens | | | | helemaal oneens |
|---|---|---|---|---|---|
| goed gestructureed | ○ | ○ | ○ | ○ | ○ |
| duidelijk | ○ | ○ | ○ | ○ | ○ |
| niet interessant | ○ | ○ | ○ | ○ | ○ |
| makkelijk te onthouden | ○ | ○ | ○ | ○ | ○ |
| van goede kwaliteit | ○ | ○ | ○ | ○ | ○ |
| goed in beeld gebracht | ○ | ○ | ○ | ○ | ○ |

Q4. Ik vond de taak in deze instructievideo

| | helemaal eens | | | | helemaal oneens |
|---|---|---|---|---|---|
| leuk om te doen | ○ | ○ | ○ | ○ | ○ |
| makkelijk om te doen | ○ | ○ | ○ | ○ | ○ |
| saai om te doen | ○ | ○ | ○ | ○ | ○ |
| moeilijker dan ik had verwacht | ○ | ○ | ○ | ○ | ○ |

Q5. De gesproken taal in de instructievideo was …

| | helemaal eens | | | | helemaal oneens |
|---|---|---|---|---|---|
| makkelijk te begrijpen | ○ | ○ | ○ | ○ | ○ |
| moeilijk te volgen | ○ | ○ | ○ | ○ | ○ |
| afleidend van de taak | ○ | ○ | ○ | ○ | ○ |
| ondersteunend aan de taak | ○ | ○ | ○ | ○ | ○ |
| te snel | ○ | ○ | ○ | ○ | ○ |
| te informatief | ○ | ○ | ○ | ○ | ○ |

Q6. De ondertiteling van de instructievideo was …

| | helemaal eens | | | | helemaal oneens |
|---|---|---|---|---|---|
| moeilijk te begrijpen | ○ | ○ | ○ | ○ | ○ |
| makkelijk te volgen | ○ | ○ | ○ | ○ | ○ |
| te langzaam | ○ | ○ | ○ | ○ | ○ |
| ondersteunend aan de taak | ○ | ○ | ○ | ○ | ○ |
| afleidend van de gesproken taal | ○ | ○ | ○ | ○ | ○ |
| afleidend van het beeld | ○ | ○ | ○ | ○ | ○ |

Q7. Wat vond je van de instructievideo in het algemeen?

Q8. In vergelijking met een papieren handleiding is de instructievideo

| | helemaal eens | | | | helemaal oneens |
|---|---|---|---|---|---|
| makkelijker | ○ | ○ | ○ | ○ | ○ |
| leuker | ○ | ○ | ○ | ○ | ○ |
| informatiever | ○ | ○ | ○ | ○ | ○ |

Q9. Stel dit was de handleiding voor het in elkaar zetten van een kast, wat had je liever?

- instructievideo
- papieren handleiding
- beide

Q10. Wanneer heb je voor het laatst met LEGO gebouwd?

- Afgelopen week nog
- Afgelopen maand nog
- Langer dan een jaar geleden
- Langer dan 5 jaar geleden
- Langer dan 10 jaar geleden

Q11. Welke van de volgende talen spreek je en hoe goed?

| | heel goed | goed | matig | niet goed | helemaal niet |
|---|---|---|---|---|---|
| Engels | | | | | |
| Duits | | | | | |
| Spaans | | | | | |
| Nederlands | | | | | |

Q12. Je bent

- man
- vrouw
- zeg ik liever niet

Q13. Hoe oud ben je?

Q14. Wat is je moedertaal?

- Nederlands
- Engels
- Duits
- anders, namelijk

Q15. Wat is je opleidingsniveau?

- MBO
- HBO
- WO
- Anders, namelijk: