

MASTER'S THESIS

Langetermijneffecten van instructievideo's op kennisbehoud: Een meta-analyse.

Vandevelde, Björn

Award date:
2023

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 05. Dec. 2023

Open Universiteit
www.ou.nl





Langetermijneffecten van Instructievideo's op Kennisbehoud: Een Meta-Analyse

Long-Term Effects of Instructional Videos on Retention: A Meta-Analysis

Björn Vandavelde

Master Onderwijswetenschappen, Open Universiteit

E-mailadres: vandavelde_bjorn@hotmail.com

Cursuscode en cursusnaam: OM9906 Masterscriptie

Naam begeleider: dr. K. Ackermans

Woordenaantal: 9.994 woorden

Datum: 21/01/2023

Samenvatting

Verschillende onderwijsniveaus maken steeds vaker gebruik van digitale leermaterialen, zoals instructievideo's (Henderson & Schroeder, 2021). Daarom is het belangrijk om te begrijpen hoe instructievideo's kunnen bijdragen tot betere leerresultaten. In de vakgroep Online leren en instructie binnen de Faculteit Onderwijswetenschappen van Open Universiteit loopt de onderzoekslijn "Naar evidence-based ontwerprichtlijnen voor instructievideo's". Eén van de doelen van de onderzoekslijn is onderzoeken hoe ontwerpvarianties in instructievideo's leren beïnvloeden op lange termijn. De huidige onderzoeksresultaten geven geen eenduidig beeld, wat kan liggen aan variaties in onderzoeksdesign of data-analytische keuzes (De Berg, 2022; Palte, 2022; Van der Spoel, 2022; Van Hese, 2022). Het doel van dit onderzoek was via een meta-analyse nagaan hoe ontwerpvarianties in instructievideo's een verschillende impact hebben op retentie. Door ruwe data uit eerdere onderzoeken te integreren in één Bayesiaans multilevel regressiemodel, konden de verschillende uitkomsten op data-analytisch vlak vergeleken worden. De resultaten konden bovendien aangeven of de data bewijs leverden voor op de literatuur gebaseerde hypothesen over ontwerpvarianties in instructievideo's. De mogelijkheid tot herbekijken, de integratie van oefeningen en het maken van tussentijdse toetsen lijken op basis van de regressieanalyse de condities te zijn die het meest positieve langetermijneffect hebben op kennisbehoud. Goed ontworpen instructievideo's staan dus op zichzelf niet garant voor kennisbehoud, maar hebben wel potentieel in combinatie met andere didactische ingrepen. Die bevinding kan ontwikkelaars en gebruikers van multimediale lesmaterialen helpen bij het ontwerpen en selecteren van instructievideo's die aansluiten op de leefwereld van studenten, en een positieve impact hebben op kennisbehoud.

Keywords: retentie, instructievideo's, meta-analyse, langetermijneffecten

Abstract

Various levels of education are increasingly making use of digital learning materials, such as instructional videos. That is why it is important to understand how instructional videos can contribute to better learning outcomes. The department of Online learning and instruction within the Faculty of Educational Sciences of the Open University developed the research line "Towards evidence-based design guidelines for instructional videos". One of the goals of the research line is to examine how design variations in instructional videos affect long-term learning. The current research results do not provide a clear picture, which may be a result of variations in research design or data-analytical choices (De Berg, 2022; Palte, 2022; Van der Spoel, 2022; Van Hese, 2022). The aim of this study was to investigate through a meta-analysis how design variations in instructional videos have a different impact on retention. By integrating raw data from previous studies into one Bayesian multilevel regression model, the different outcomes could be compared on a data-analytical level. The results could also indicate whether the data provided evidence for literature-based hypotheses about design variations in instructional videos. Based on the regression analysis, the possibility of reviewing, integrating exercises and making tests seem to be the conditions that have the most positive long-term effect on retention. This means that well-designed instructional videos do not in themselves guarantee knowledge retention, but do have potential in combination with other didactic interventions. These findings can help developers and users of multimedia teaching materials to design and select instructional videos that connect to the living environment of students and have a positive impact on knowledge retention.

Keywords: retention, instructional videos, meta-analysis, long-term effects

Inhoud

Samenvatting	2
Abstract	3
1. Inleiding	5
1.1 Probleemschets en Doel	5
1.2 Theoretisch Kader	7
1.3 Huidige Studie.....	16
2. Methode.....	19
2.1 Dataverzameling en Inclusiecriteria.....	19
2.2 Data-Analyse.....	22
3. Resultaten	25
3.1 Regressieanalyses.....	25
3.2 Vergelijking Tussen Modellen	31
4. Discussie.....	32
4.1 Beantwoording Onderzoeksvraag en Hypotheses.....	32
4.2 Beperkingen van het Onderzoek en Toekomstig Onderzoek.....	37
4.3 Praktische Implicaties	41
4.4 Conclusie.....	43
Referenties.....	45
Bijlage A	58
Bijlage B.....	61

Langetermijneffecten van Instructievideo's op Kennisbehoud: Een Meta-Analyse

1. Inleiding

1.1 Probleemschets en Doel

1.1.1 Probleemschets

De COVID-19 pandemie heeft in een korte periode wereldwijd verschillende domeinen van het dagelijks leven ingrijpend veranderd. Zo werden leerlingen op alle onderwijsniveaus geconfronteerd met schoolsluitingen en een plotse omschakeling naar afstandsonderwijs. In deze periode kregen leerlingen veelal een online variant van het lesprogramma aangeboden. Het Nederlands Ministerie van Onderwijs, Cultuur en Wetenschap wil ook na de COVID-19 pandemie middels digitaal afstandsonderwijs het onderwijs als geheel te verbeteren (Ministerie van Onderwijs, Cultuur en Wetenschap, 2020). Ook de Vlaamse regering wil technologische hulpmiddelen in het onderwijs aanmoedigen. Ze doet dat door de invoering van de Digisprong, een beleidsmaatregel die ervoor zorgt dat vanaf september 2022 elke Vlaamse leerling vanaf tien jaar over een laptop kan beschikken, zowel thuis als op school (Vlaanderen – Onderwijs en Vorming, 2021). Dat geeft scholen de kans om meer multimedia te integreren in de lessen. Universiteiten en hogescholen maken nu al op steeds grotere schaal gebruik van digitale leermaterialen, zoals instructievideo's (Jansen & Van Casteren, 2021). Het aanzienlijke online aanbod zorgt er ook voor dat studenten vaker instructievideo's bekijken, naast analoge lesmateriaal zoals boeken (Herreid & Schiller, 2013).

Instructievideo's kunnen ondersteuning bieden bij zowel fysieke lessen (Leo & Puzzi, 2016) als bij afstandsonderwijs (Mayer, 2018). Afstandsonderwijs brengt echter risico's met zich mee, waaronder minder actieve deelname aan het onderwijs (Ministerie van Onderwijs, Cultuur en Wetenschap, 2020) en verminderde motivatie (Holzer et al., 2021; Wang et al., 2019). Dalende motivatie kan een negatief effect hebben op leerresultaten. Leerlingen

presteren namelijk beter naarmate zij meer gemotiveerd zijn (Boekaerts et al., 2010). Ook bij fysieke lessen kunnen video's die in functie van de lesinhoud, leerstrategie en het doelpubliek niet adequaat ontworpen zijn, ongewenste leerbelemmeringen veroorzaken (Mayer & Moreno, 2003; Moreno & Mayer, 2007; Sweller & Chandler, 1994). Om die risico's zoveel mogelijk te beperken, is het belangrijk om te achterhalen op welke manier instructievideo's zo goed mogelijk zowel afstands- als fysiek onderwijs kunnen ondersteunen en hoe ze kunnen bijdragen tot betere leerresultaten.

Mayer (2021) heeft een reeks ontwerpprincipes ontwikkeld die moeten leiden tot het design van effectieve instructievideo's. Aandacht voor de langetermijneffecten van video's op kennisbehoud is daarbij essentieel. Het onderzoek naar het langetermijneffect van video's ontworpen op basis van die ontwerpprincipes in een authentieke context, zoals het klaslokaal, zit echter nog in een beginfase.

1.1.2 Doel

Verschillende scripties aan de Open Universiteit hebben onderzocht of het gebruik van video's een positief effect heeft op het vermogen om leerstof langdurig te onthouden. De resultaten geven momenteel geen eenduidig beeld. Het doel van dit onderzoek is door middel van een overkoepelende vergelijking van individuele scripties nagaan hoe hun specifieke onderzoeksontwerpen een verschillende impact hebben op retentie. Een meta-analyse maakt het mogelijk om ruwe data van de individuele scripties opnieuw te analyseren. Door de integratie van de data in één Bayesiaans multilevel model, kunnen uitkomsten van verschillende onderzoeken op data-analytisch vlak vergeleken en opnieuw geëvalueerd worden. De resultaten kunnen aangeven of er voor de vanuit de literatuur onderbouwde hypothesen een bewijs voor waarschijnlijkheid gevonden kan worden in de data. Zo kan een completer beeld gevormd worden over hoe ontwerpvarianties in instructievideo's leren beïnvloeden op lange termijn. De bevindingen uit dit onderzoek kunnen uitgevers en

gebruikers van multimediale lesmaterialen helpen bij het ontwikkelen en selecteren van video's die aansluiten op de voorkeur en leefwereld van studenten, en ook een positieve impact hebben op kennisbehoud.

1.2 Theoretisch Kader

Om te vatten hoe leren werkt, zal in dit theoretisch kader eerst op een algemener niveau besproken worden hoe leer- en informatieverwerkingsprocessen verlopen. Daarna wordt specifiek ingegaan op de voor- en nadelen van leren via multimedia en onderzoek naar langetermijneffecten van multimediaal leren. Ten slotte worden de voordelen besproken van een Bayesiaanse aanpak bij de analyse van resultaten die peilen naar langetermijneffecten van instructievideo's op retentie.

1.2.1 Het Leerproces

1.2.1.1 Retentie. Als onderdeel van het leerproces verwerken studenten grote hoeveelheden informatie tijdens of onmiddellijk na het instructiemoment. Leren vindt echter pas plaats als die informatie op een duurzame manier in het langetermijngeheugen terecht komt. Toch is informatie verzamelen zonder de kennis structureel te verankeren een veelvoorkomend fenomeen bij studenten. Ze verwerken grote hoeveelheden leerstof, maar verliezen vlug het vermogen om zich die stof te herinneren (Hopkins et al., 2016). Retentie is essentieel in het leerproces. Het kan omschreven worden als het vermogen om informatie in de loop van de tijd te onthouden en met succes uit het langetermijngeheugen op te halen (Engelbrecht et al., 2007; Hopkins et al., 2016). Het onvermogen om eerder opgedane kennis correct en vlot op te roepen kan een succesvolle voortzetting van de educatieve loopbaan bemoeilijken. Daarom is het belangrijk om te begrijpen hoe het leerproces verloopt, welke obstakels retentie in de weg staan en welke stappen het onthouden faciliteren.

1.2.1.2 Het Geheugen. Het geheugen is van essentieel belang bij het verwerken, bewaren en ophalen van kennis (Hopkins et al., 2016). Het bestaat uit het sensorisch

geheugen, het werkgeheugen en het langetermijngeheugen (Atkinson & Shiffrin, 1968). Eerst komt zintuiglijke input in het sensorisch geheugen terecht, waar de binnengekomen informatie kortstondig wordt vastgehouden. Via het sensorisch geheugen komt informatie vervolgens in het werkgeheugen terecht. Daar wordt de informatie verder verwerkt zodat ze opgeslagen kan worden in het langetermijngeheugen of meteen het antwoord op een vraag kan bieden. Ten slotte komt informatie in verwerkte vorm in het langetermijngeheugen terecht. Vanuit het langetermijngeheugen kan informatie opnieuw opgeroepen worden om via verwerking in het werkgeheugen als antwoord op een vraag te dienen (Driscoll, 1994). Het langetermijngeheugen bestaat uit kennisschema's, dat zijn structuren die kennisdelen aan elkaar koppelen. Nieuwe informatie die in het langetermijngeheugen terechtkomt, wordt verbonden met al bestaande kennisschema's. Op die manier ontstaat leren en kan een onbeperkte hoeveelheid informatie opgeslagen worden in het langetermijngeheugen (Schunk, 2014; Surma et al., 2019). Het leerproces kan echter beïnvloed worden door zowel hinderende als faciliterende omstandigheden.

1.2.1.3 Hindernissen in het Leerproces. De Duitse psycholoog Ebbinghaus heeft in 1880 al aangetoond hoe mensen na verloop van tijd informatie vergeten. Daarbij zag hij een samenhang tussen geheugen en tijd. Volgens de vergeetcurve van Ebbinghaus gebeurt vergeten eerst snel, om vervolgens te vertragen. In het eerste uur na de opname van informatie kunnen studenten nog vrij veel vergeten, maar informatie die na een week nog in het langetermijngeheugen zit, zullen ze zo goed als niet meer vergeten (Murre & Dros, 2015). Omdat retentie inhoudt dat studenten informatie na enige tijd nog steeds onthouden en beheersen (Surma et al., 2019), is het tijdens het leerproces dus essentieel dat vergeten voorkomen wordt en informatie het langetermijngeheugen bereikt.

Het informatieverwerkingsproces kan echter bemoeilijkt worden door onder meer overbelasting van het werkgeheugen (Sweller, 1994). De *cognitive load theory* (Chandler &

Sweller, 1991) gaat ervan uit dat de totale vereiste cognitieve belasting van leeractiviteiten de cognitieve capaciteit van het werkgeheugen niet mag overschrijden. Dan zou bij de lerende immers onvoldoende werkgeheugencapaciteit overblijven om informatie succesvol te integreren in het langetermijngeheugen (Mayer & Moreno, 2003). De cognitive load theory houdt dus in dat didactisch materiaal zodanig ontworpen moet worden dat het werkgeheugen optimaal wordt ingezet en niet overbelast geraakt.

Naast de manier waarop leermateriaal ontworpen is, kan ook de voorkennis van studenten een impact hebben op cognitieve belasting. Als het leerproces een onderwerp behandelt waarover een student al voorkennis heeft, is dat minder belastend voor het werkgeheugen dan wanneer de student nog geen eerdere kennis heeft. Voorkennis kan immers op een efficiënte manier in het langetermijngeheugen worden gestructureerd in kennisschema's (Sweller et al., 1998). De voorkennis in die schema's zorgt er tijdens het leren voor dat meer nieuwe informatie actief kan blijven in het werkgeheugen. Nieuwe informatie die bij een bestaand kennisschema past, wordt bovendien beter geïntegreerd. Wie nog geen of heel ruwe schema's over het leeronderwerp bezit, moet de nieuwe informatie nog apart verwerken zonder koppeling aan bestaande kennis. Dat is extra belastend voor het werkgeheugen.

1.2.1.4 Het Leerproces Faciliteren. Het leerproces optimaliseren kan door enerzijds inzicht te verwerven in hindernissen die retentie kunnen bemoeilijken. Anderzijds is het ook essentieel om zicht te krijgen op elementen die retentie net faciliteren. Onderzoek van Engelbrecht et al. (2007) heeft laten zien dat retentie ontstaat door interne factoren, zoals een diepgaande kennis van concepten, en externe factoren zoals leerstrategieën en onderwijsmethoden. Het onthouden van leerstof hangt immers samen met de manier waarop die wordt aangebracht. Surma et al. (2019) hebben aangegeven dat actieve leerstrategieën het behoud van kennis kunnen verbeteren. Daarnaast hebben Murre en Dros (2015) laten zien dat

studenten leerstof grondiger onthouden bij het spreiden van studeermomenten in de tijd, of *spaced practice*. Studenten die binnen een tijdspanne van een uur herhaaldelijk leerstof verwerken, zullen die een dag onthouden. Herhalen ze daags nadien de leerstof opnieuw, zullen ze die een week onthouden (Murre & Dros, 2015; Surma et al., 2019). Studenten kunnen dus een hoge retentie bereiken door leerstof meermaals te herhalen, bijvoorbeeld meteen na het instructiemoment, na een week en na een maand (Murre & Dros, 2015; Witt et al., 2011). Naast herhalen kan ook *retrieval practice* helpen om retentie te faciliteren (Dirkx et al., 2019). Retrieval practice is door middel van oefeningen of toetsen informatie oproepen uit het langetermijngeheugen. In functie van een hoge retentiegraad kunnen leraren dus inzetten op een combinatie van actieve leerstrategieën zoals spaced practice door middel van herhalen en retrieval practice in de vorm van bijvoorbeeld toetsen.

1.2.2 Instructievideo's

1.2.2.1 Multimediaal Leren. Eenmaal duidelijk is hoe het leerproces verloopt, welke obstakels retentie in de weg staan en welke stappen het onthouden faciliteren, is het belangrijk om inzicht te verwerven in hoe deze principes geïmplementeerd kunnen worden in multimediaal leren. Multimediaal leren betekent dat er kennis geconstrueerd wordt door de combinatie van woorden en afbeeldingen (Mayer & Pilegard, 2014). Daarbij kunnen woorden gesproken of in een presentatie geschreven zijn. Beelden kunnen op hun beurt stilstaand of bewegend zijn (Mayer & Moreno, 2003). Mayer (2002) heeft gesteld dat studenten door middel van multimedia beter leren dan louter van tekst. Multimediaal instructieontwerp gericht op menselijke informatieverwerking faciliteert betekenisvol leren (Mayer & Pilegard, 2014; Mayer & Moreno, 2003). Dat blijkt uit positieve uitkomsten in transfer- en retentietesten (Mayer, 2019). Het is dus essentieel om kennis te verwerven over hoe studenten leren van instructievideo's. In *The Cambridge Handbook of Multimedia Learning* (2005) heeft Mayer dat toegelicht via de *cognitive theory of multimedia learning* (CTML). Deze

theorie vertrekt vanuit twee principes, het modaliteitsprincipe en het dubbele coderingsprincipe.

Het modaliteitsprincipe stelt dat informatie via verschillende zintuigen het geheugen bereikt (Moreno, 2006). Bij een instructievideo bijvoorbeeld kunnen zowel de ogen als de oren aangesproken worden. Doordat simultaan twee afzonderlijke verwerkingsprocessen lopen in het werkgeheugen, vermindert de cognitieve belasting (Mayer, 2021; Mayer & Fiorella, 2014). Vooral bij een gebrek aan voorkennis bevordert dat het leerproces (Seufert et al., 2009).

Naast die opsplitsing op basis van sensorische input gaat Mayer ook uit van het dubbele coderingsprincipe van Paivio (1991). Dat principe stelt dat binnenkomende informatie gecodeerd en verwerkt wordt via enerzijds een verbaal systeem voor woorden en anderzijds een non-verbaal systeem voor onder andere beelden, geuren of smaken. Het verbale systeem kan zowel auditief als visueel woorden verwerken, terwijl het non-verbale systeem tegelijk andere impulsen zoals beelden kan verwerken. De simultane inzet van beide systemen versterkt de informatieopname en beperkt de cognitieve belasting. In de CTML heeft Mayer zowel het modaliteits- als het dubbele coderingsprincipe geïntegreerd. Een ingeperkte belasting van het werkgeheugen en optimalisering van het leerproces worden gefaciliteerd door het via meerdere zintuigen kunnen verwerken van zowel verbale als non-verbale informatie. Instructievideo's die zowel auditief als visueel woorden en afbeeldingen bevatten, zouden dus resulteren in een succesvolle integratie van nieuwe kennis.

1.2.2.2 Voordelen van Instructievideo's. Op basis van de gecombineerde inzet van het modaliteits- en dubbele coderingsprincipe hebben instructievideo's ten opzichte van leerstof in de vorm van een tekst een aantal voordelen op vlak van kennisverankering. Studenten zouden een audiovisuele presentatie zoals bijvoorbeeld een grafiek begeleid door een ingesproken verklaring beter vatten dan een louter visuele presentatie van een grafiek met

een geschreven verklaring (Ginns, 2005; Leahy et al., 2003; Mousavi et al., 1995). Leahy en Sweller (2016) hebben echter gesteld dat dit modaliteitsprincipe enkel een positief effect heeft bij korte audiovisuele fragmenten. Dat sluit aan op onderzoek dat stelt dat instructievideo's die maximum 6 minuten duren studenten het meest aanspreken (Brame, 2016; Clossen, 2018).

Instructievideo's kunnen daarnaast ook op een dynamische manier complexe processen en praktijkvoorbeelden duiden (Cooper & Higgins, 2015; Mitra et al., 2010; Vaganova et al., 2019). Bij een statische informatieweergave moet een student eerst informatie uit de tekst verwerken en die daarna afstemmen met de begeleidende beelden (Kühl et al., 2011; Schnotz & Lowe, 2008). Videovoorstellingen van dynamische processen vergen daarentegen minder cognitieve belasting omdat de student in zijn werkgeheugen zelf geen dynamisch cognitief model meer moet samenstellen uit afzonderlijke statische onderdelen (Hegarty & Kriz, 2008). Bij een statische weergave van dynamische processen bestaat tevens het risico dat een student het proces onvoldoende kan reconstrueren uit de verschillende statische onderdelen (Berney & Bétrancourt, 2016; Tversky et al., 2002). Dat sluit aan op studies die stellen dat een video wel adequaat informatie kan weergeven die mondeling of door een tekst onmogelijk op een afdoende manier kan worden beschreven (Rasi & Poikela, 2016; Schneps et al., 2010).

Los van de voordelen die rechtstreeks gerelateerd zijn aan de werking van het werkgeheugen, merken docenten ook dat instructievideo's tot betere leerresultaten leiden (Filius & Lam, 2010) omdat ze dicht bij de leefwereld van studenten liggen. De huidige generatie studenten verkiest immers geluid en video boven teksten (Granitz et al., 2021; Rickes, 2016; Szymkowiak et al., 2021). Studenten geven aan dat instructievideo's motiverend, bevredigend, geestelijk stimulerend en hulpvaardig zijn in het leerproces (Kay & Kletskin, 2012; Preston et al., 2010). Een video roept bovendien meer emoties op dan een tekst, waardoor studenten zich meer betrokken voelen bij de inhoud (Tarchi et al., 2021).

1.2.2.3 Nadelen van Instructievideo's. Er moet echter nagegaan worden of het gebruik van instructievideo's steeds afgestemd is op de specifieke leerinhoud. Wanneer dat niet het geval is, kan op vlak van retentie een hierboven geschetst voordeel in de praktijk immers een nadeel blijken te zijn. Als leerstof niet dynamisch is en een statische weergave voldoende zou zijn om de les te begrijpen, kunnen dynamische videovoorstellingen immers leiden tot meer cognitieve belasting omdat er onnodig veel informatie tegelijk wordt gepresenteerd (Arguel & Jamet, 2009; Sweller & Chandler, 1994). Daarnaast kan ook een nadelig redundantie-effect optreden (Mayer & Moreno, 2003; Sweller & Chandler, 1991). Dat houdt in dat als instructievideo's op hetzelfde moment ingesproken tekst laten horen en een identieke geschreven tekst laten zien, studenten die informatie als overbodig kunnen beschouwen. Exact dezelfde informatie die onnodig op twee manieren wordt gepresenteerd, kan cognitief overbelastend werken. Het is dus belangrijk om via het modaliteits- en dubbele coderingsprincipe niet dezelfde informatie dubbel aan te bieden, maar verschillende onderdelen leerinhoud simultaan op een wederzijds versterkende manier te presenteren.

De voorbijgaande aard van gepresenteerde gegevens in een videoweergave kan ten slotte ook als nadelig beschouwd worden. Het werkgeheugen kan immers maar voor een korte periode informatie vasthouden. Als een video aan een hoge frequentie complexe leerinhoud aanbiedt, kan het werkgeheugen relevante fragmenten missen omdat het de informatiestroom niet kan bijhouden. Het krijgt onvoldoende tijd om gegevens te verwerken en te integreren met voorkennis uit het langetermijngeheugen (Chandler, 2009; Leahy & Sweller, 2011; Mayer & Chandler, 2001; Moreno & Mayer, 2007). Dat kan zelfs leiden tot relatief slechtere leerresultaten ten opzichte van leren via tekst (Chen et al., 2017). Dit nadeel kan echter gecompenseerd worden door studenten via *self-pacing* controle te geven over het tempo van de video. Dat houdt in dat studenten de weergave kunnen aanpassen aan hun specifieke cognitieve capaciteiten (Mayer & Chandler, 2001; Moreno & Mayer, 2007) door

controleopties zoals pauzeren en terug- en doorspoelen te integreren in de video (Biard et al., 2018). De video door middel van segmentering opsplitsen in verschillende betekenisvolle segmenten kan daarnaast ook voorkomen dat het werkgeheugen overbelast wordt door de snelheid waarmee complexe informatie gepresenteerd wordt (Sweller et al., 2019).

1.2.2.4 Onderzoek naar Langetermijneffecten. Op basis van de CTML heeft Mayer (2021) een reeks ontwerpprincipes ontwikkeld die moeten leiden tot het design van effectieve instructievideo's. Mayer heeft echter aangegeven dat het effect van video's ontworpen op basis van de combinatie van die ontwerpprincipes, zoals het modaliteits- en dubbele coderingsprincipe, nog onderzocht moet worden in authentieke contexten zoals klaslokalen. Onderzoek naar langetermijneffecten van video's op kennisbehoud is daarbij essentieel. Studenten leren immers pas als ze kennis opslaan in het langetermijngeheugen. Daarom is het belangrijk om leerresultaten niet alleen meteen na een instructiemoment te meten, maar ook naar het langetermijneffect te peilen (Soderstrom & Bjork, 2015). De manier waarop verzamelde onderzoeksdata vervolgens geanalyseerd worden, heeft echter ook impact op de conclusies die geformuleerd kunnen worden.

1.2.3 Bayesiaanse Data-Analyse

Bij de analyse van resultaten die peilen naar het langetermijneffect van instructievideo's op retentie moet nagegaan worden welke data-analytische methode het meest aangewezen is. Een aantal argumenten pleiten voor een Bayesiaanse data-analyse. Een eerste argument is dat het gangbare gebruik van de nulhypothese-significantietest en de bijhorende p -waarden in verschillende reviews als oorzaak worden genoemd van de replicatiecrisis in psychologische wetenschappen (Nuzzo, 2014; Open Science Collaboration, 2015). Grootschalig onderzoek van 855 t -toetsen van Wetzels et al. (2011) heeft laten zien dat p -waarden en Bayesiaanse uitkomstmaten het vaak oneens zijn over effectgroottes. Dat doet vragen rijzen over de juistheid van de conclusies die onderzoekers trekken uit p -waarden. Een

nulhypothese-significantietest kan immers door middel van een p -waarde enkel een nulhypothese wel of niet weerleggen. Het niet-weerleggen van de nulhypothese stelt enkel dat er geen bewijs is voor een verschil tussen twee onderzoeksgroepen. Er kan niet nagegaan worden hoe goed het bewijs is dat er geen verschil is (Kubsch et al., 2021). Daarnaast wordt er bij weerlegging van een nulhypothese geen bewijs geleverd dat er wel een verschil zou zijn tussen onderzochte groepen. Nulhypothese-significantietesten zeggen immers niets over de alternatieve hypothese die stelt dat er een verschil is. Bayesiaanse analyses kunnen daarentegen wel een vergelijking maken tussen de waarschijnlijkheid van de nul- en de alternatieve hypothese (Wagenmakers, 2018).

Een tweede argument voor Bayesiaanse analyses is het gegeven dat nulhypothese-significantietesten veronderstellen dat data normaal verdeeld zijn (Field, 2018). In onderzoek naar de impact van instructievideo's op retentie is de afhankelijke variabele, met name toetsresultaten, echter vaak ongelijk verdeeld. Daardoor kan de assumptie van gelijke variantie geschonden worden. Een andere manier van analyseren is dan aan de orde. Ook bij Bayesiaanse analyses moeten data echter omzichtig worden geïnterpreteerd bij assumptieschending van normaliteit (van Doorn et al., 2021). Toch kunnen Bayesiaanse analyses op een voorzichtigere manier meer passende en beschrijvende informatie geven. Het effect van onafhankelijke variabelen wordt er immers uitgedrukt in minder strenge uitkomstmaten, met name de *Probability of Direction* (pd) en het percentage overeenkomst tussen het *High Density Interval* (HDI) en de *Region Of Practical Equivalence* (ROPE).

De pd is een index, variërend van 50% tot 100%, die op basis van de posterieure verdeling in een Bayesiaans regressiemodel de zekerheid weergeeft waarmee een effect positief of negatief is. Hoewel anders uitgedrukt, komt deze index vrij overeen met de frequentistische p -waarde (Makowski et al., 2019). Een pd van 97,5% komt daarbij overeen met een p -waarde van .05.

De ROPE laat op zijn beurt toe dat er een gebied rond de nulwaarde wordt gedefinieerd waarbinnen parameterwaardes vallen die beschouwd kunnen worden als praktisch gelijkwaardig aan nul (Kruschke, 2014). Kruschke (2018) heeft gesteld dat zo'n gebied rond de nulwaarde standaard kan worden ingesteld met een bereik van -0.1 tot 0.1. Het gedeelte van het door het geloofwaardigheidsinterval afgebakende HDI dat binnen de ROPE valt, kan gebruikt worden als index voor een Bayesiaanse vorm van nulhypothese-significantietesten. Als het HDI volledig buiten de ROPE valt, kan de nulhypothese verworpen worden. Als het HDI volledig binnen de ROPE valt, kan de nulhypothese niet verworpen worden. Bij een gedeeltelijke overlapping van het HDI en de ROPE blijft het onbeslist of de nulhypothese al dan niet verworpen wordt. Dat geeft een omzichtigere en meer tolerante uitkomst dan de scherpe verwerping van de nulhypothese op basis van een p -waarde (Wagenmakers, 2018).

Een derde argument voor Bayesiaanse analyses is de mogelijkheid om voorgaande kennis uit eerdere onderzoeken te integreren in de analyse. Het resultaat van Bayesiaanse analyses is immers opgebouwd uit enerzijds nieuwe verzamelde data en anderzijds beschikbare voorkennis. Die voorkennis kan geïntegreerd worden in de analyse door het selecteren van een prior-waarde. De uitkomst van Bayesiaanse analyses kan dus gezien worden als een actualisering van kennis die het resultaat is van een analytische integratie van voorkennis met nieuw empirisch bewijs (Kruschke & Liddell, 2017).

1.3 Huidige Studie

Een aantal scripties op het gebied van instructievideo's hebben onderzoek uitgevoerd waarbij het kennisbehoud na videolessen over een langere periode werd gemeten (De Berg, 2022; Palte, 2022; Van der Spoel, 2022; Van Hese, 2022). Ondanks die gemeenschappelijkheid werden in de scripties verschillende keuzes gemaakt op data-analytisch vlak. De verschillende uitkomstmaten maakten het dan ook moeilijk om resultaten

uit de onderzoeken te kunnen vergelijken. De gevonden effecten van deze onderzoeken gaven bovendien niet eenduidig één richting of sterkte aan. Dat maakte het lastig om op basis van de verschillende studies een correct beeld te vormen over de reële impact van instructievideo's op retentie. Daarom is een meta-analyse uitgevoerd met vier onderzoeken naar het langetermijneffect van instructievideo's op kennisbehoud.

Glass (1976) heeft meta-analyse gedefinieerd als de statistische analyse van resultaten uit individuele onderzoeken met als doel het integreren van de bevindingen. Die integratie vond binnen deze studie plaats door de ruwe data uit vier individuele onderzoeken samen te voegen in één Bayesiaans multilevel regressiemodel. Het multilevel model maakte het mogelijk om data uit verschillende longitudinale onderzoeken met onderling variërende meetmomenten te integreren in één model (Bauer & Curran, 2022). Daarbij zijn verschillende ontwerpvarianties onderling vergeleken. Bovendien werd een voorspelling gemaakt over welk effect die ontwerpkeuzes hebben op retentie. Er is nagegaan hoe de leerstrategie, de mogelijkheid tot het herbekijken van video's, de integratie van oefeningen in video's, de lengte van video's en de combinatie van video en tekst impact hadden op de leerresultaten. De onderzoeken die in de meta-analyse zijn opgenomen, zijn in Tabel 1 schematisch weergegeven.

Tabel 1*Onderzoeksdesigns van Onderzochte Scripties*

Auteur	Onderwijsniveau deelnemers	Experimentele conditie	Controle- conditie	Lesinhoud	Meetmoment	Tussendoor leerstrategieën i.f.v. retentie
De Berg	hbo	Kennisclip	Tekst	Employer branding	- Meteen - Na 6 weken	Nee
Van Hese	Voortgezet onderwijs	Video binnen flipped classroom	Tekst	Wiskunde	- Meteen - Na 1 week - Na 1 maand	Ja
Van der Spoel	Primair onderwijs	Video en tekst	Tekst	Woordenschat	- Meteen - Wekelijks - Na 5 weken	Ja
Palte	Primair onderwijs	Video	Directe instructie	Romeinse cijfers	- Meteen - Na 1 week - Na 5 weken	Ja

Vanuit de literatuur zijn hypothesen opgemaakt over wat de impact van die ontwerpkeuzes kon zijn op de onderzoeksresultaten. Een Bayesiaans multilevel model moest daarbij aangeven of er in de data ook bewijs voor waarschijnlijkheid gevonden kon worden voor die hypothesen. Dat leidde tot de volgende onderzoeksvraag: *Wat is het effect van ontwerpvariëaties in instructievideo's op kennisbehoud in langetermijninterventies?* Op basis van ontwerpvariëaties in de vier verzamelde onderzoeken op vlak van leerstrategie, de mogelijkheid tot het herbekijken van video's, de integratie van oefeningen in video's, de lengte van video's en de combinatie van video en tekst, werden de volgende hypothesen afgeleid:

H1: Instructievideo's die gecombineerd worden met op retentie gerichte leerstrategieën, leveren ten opzichte van lescondities zonder instructievideo betere

leerresultaten op dan instructievideo's die niet gecombineerd worden met op retentie gerichte leerstrategieën.

H2: Lescondities waarin het gebruik van instructievideo's en tekst gecombineerd wordt, leveren ten opzichte van lescondities zonder instructievideo betere leerresultaten op dan lescondities die enkel instructievideo's inzetten.

H3: Instructievideo's die herbekeken kunnen worden, leveren ten opzichte van lescondities zonder instructievideo betere leerresultaten op dan instructievideo's die slechts eenmaal bekeken kunnen worden.

H4: Instructievideo's waarin oefeningen geïntegreerd zijn, leveren ten opzichte van lescondities zonder instructievideo betere leerresultaten op dan instructievideo's waarin geen oefeningen geïntegreerd zijn.

H5: Instructievideo's korter dan 10 minuten, leveren ten opzichte van lescondities zonder instructievideo betere leerresultaten op dan instructievideo's langer dan 10 minuten.

2. Methode

2.1 Dataverzameling en Inclusiecriteria

Het doel van dit onderzoek was door middel van een overkoepelende vergelijking van individuele studies nagaan hoe specifieke onderzoeksontwerpen een verschillende impact hebben op retentie. De meest aangewezen onderzoeksmethode om geschikte studies te verzamelen, was een systematische review. Systematische reviews worden steeds vaker gebruikt om literatuur samen te vatten en onderzoeksgebieden af te bakenen (Crowther et al., 2010). Ze beperken bias door via een reproduceerbaar wetenschappelijk proces de literatuur te doorzoeken en de kwaliteit van individuele studies te evalueren. Indien mogelijk worden de resultaten uit verschillende onderzoeken vervolgens statistisch gecombineerd in een meta-analyse (Crowther et al., 2010).

De systematische review werd uitgevoerd in de database EBSCOhost. Om zicht te hebben op de meest recente onderzoeksresultaten, werden publicaties vanaf 2012 geselecteerd. Een volledig overzicht van alle zoektermen is weergegeven in Tabel 2.

Tabel 2

Trefwoorden Voor Systematische Review in EBSCOhost

Categorie	Trefwoorden
Leren	“Learning” OR “transfer” OR “retention” in titel of samenvatting
Onderzoeksmethode	“Longitudinal studies” OR “longitudinal research” OR “longitudinal method” in titel of samenvatting
Video	“Educational video” OR “instructional video” OR “multimedia” in alle velden

Na een eerste zoekronde werd de gevonden literatuur via de volgende inclusiecriteria beoordeeld:

(1) De studies moeten minstens één experiment bevatten waarbij minstens twee condities zijn vergeleken, namelijk een videoles als experimentele conditie en een reguliere les als controleconditie.

(2) Minstens één afhankelijke variabele moet het onthouden van de lesinhoud zijn, gemeten via een kennistoets.

(3) In functie van beantwoording van de hypothesen moet duidelijk zijn of er een mogelijkheid was tot het herbekijken van video's, of er oefeningen verwerkt waren in de video's, wat de videoduur was, of de videoconditie een combinatie van video en tekst inhield en of er extra leerstrategieën werden toegepast.

(4) De studie moet uitgevoerd zijn in een reguliere onderwijscontext, met als onderzoeksmiddel één of meer video's die de overdracht van nieuwe kennis inhielden.

Een eerste zoektocht leverde geen geschikte gepubliceerde studies op. De beschikbaarheid van geen of een klein aantal gepubliceerde studies kan suggereren dat het thema nog niet voldoende onderzocht is om bruikbare bevindingen op te leveren. Als het onderwerp echter van cruciaal en actueel belang is voor de samenleving, is een meta-analyse misschien alsnog aangewezen, ook al zijn er nog niet veel primaire onderzoeken (Cheung & Vijayakumar, 2016). Er moest dus op grond van andere beoordelingscriteria en vanuit een praktisch standpunt op zoek gegaan worden naar bruikbare studies.

Binnen de vakgroep Online leren en instructie van de Faculteit Onderwijswetenschappen van Open Universiteit hebben een aantal scripties onderzoek uitgevoerd waarbij kennisbehoud na videolessen over een langere periode werd gemeten (De Berg, 2022; Palte, 2022; Van der Spoel, 2022; Van Hese, 2022). Die verschillende scripties zijn gebundeld binnen een thesiskring over langetermijneffecten van instructievideo's op kennisbehoud. Ondanks enkele individuele verschillen in specifieke onderzoeksdesigns zijn de scripties op die manier inhoudelijk sterk verwant aan elkaar. In de vakgroep Online leren en instructie werden de scripties uitgevoerd binnen de onderzoekslijn "Naar evidence-based ontwerprichtlijnen voor instructievideo's". De Commissie Ethische Toetsing Onderzoek (cETO) van Open Universiteit heeft de ethische aspecten van de onderzoekslijn beoordeeld en goedgekeurd. Dat maakte het mogelijk om gebruik te maken van data die verzameld zijn in andere scripties die in de vakgroep Online leren en instructie zijn geschreven. Het inhoudelijke verwantschap tussen de scripties onderling, de link met het huidige onderzoek en de beschikbaarheid van data waren de praktische beoordelingscriteria op basis waarvan gekozen is om de meta-analyse uit te voeren met vier scripties uit de vakgroep Online leren en instructie.

Vier studies lijkt een vrij gering aantal. Davey et al. (2011) hebben echter 22.453 meta-analyses beoordeeld en ontdekt dat het mediane aantal onderzoeken drie was, terwijl het maximale aantal 294 was. Valentine et al. (2010) hebben daarnaast gesteld dat het nog steeds de moeite loont om een meta-analyse uit te voeren met slechts twee studies. Bij voldoende homogeniteit zijn meta-analytische schattingen immers nog steeds nauwkeuriger dan resultaten van individuele studies.

2.2 Data-Analyse

2.2.1 Multilevel Modeling

Het multilevel-model (MLM) is een regressiemodel dat ontwikkeld is om hiërarchisch gestructureerde data te analyseren. De structuur met meerdere niveaus, waarbij bijvoorbeeld verschillende studenten worden gegroepeerd in een klasgroep, kan echter ook toegepast worden op verschillende meetmomenten die gegroepeerd worden binnen één individu (Bauer & Curran, 2022). Het MLM is daarom geschikt om meta-analyses uit te voeren waarbij gebruikgemaakt wordt van verschillende onderzoeken met verschillende meetmomenten. Het MLM kan immers rekening houden met risico's die gepaard gaan met de samenvoeging van verschillende longitudinale onderzoeken, zoals gedeeltelijk ontbrekende gegevens, ongelijk verdeelde beoordelingen en individueel variërende meetmomenten (Bauer & Curran, 2022). Dat bood een concrete oplossing voor het feit dat deze meta-analyse vier onderzoeken samenvoegde die verschillende meetmomenten hanteerden.

2.2.2 Brms

2.2.2.1 Regressieanalyses. Dit MLM werd uitgevoerd in het R-package brms versie 2.18.0 (Bürkner, 2017) in RStudio versie 2022.07.2+576. In brms kunnen Bayesiaanse multilevel-modellen uitgewerkt worden met behulp van de probabilistische programmeertaal Stan. De keuze voor brms ondervindt zowel de voordelen van een Bayesiaanse analyse als die van een MLM. Het resultaat van een Bayesiaanse analyse, met name de posterior verdeling, is

immers opgebouwd uit enerzijds informatie uit de data en anderzijds beschikbare prior-informatie. Brms laat toe dat dit Bayesiaans principe via een MLM ook kan worden toegepast op complexere hiërarchische structuren. Een Bayesiaans MLM houdt immers rekening met complexe afhankelijkheden binnen en tussen variabelen en kan concreet binnen dit onderzoek aannemen dat het effect van tijd varieert tussen deelnemers. Daarom hoeven de gegevens niet normaal verdeeld te zijn voor een Bayesiaans MLM. Dat maakt het model geschikt voor onderzoek naar de impact van instructievideo's op retentie waarbij de afhankelijke variabele, met name toetsresultaten, vaak ongelijk verdeeld is. Een Bayesiaans MLM gaat immers standaard uit van een niet-lineaire Gaussiaanse verdeling (Bryk & Raudenbush, 1988; Bürkner, 2017; Seeger, 2004).

Bayesiaanse multilevel-modellen maken gebruik van Markov Chain Monte Carlo (MCMC)-algoritmen. In MCMC-algoritmen zijn de voorspelde retentiescores afhankelijk van toegevoegde covariaten en prior-informatie (Zeger & Qaqish, 1988). Dat betekent dat we aannemen dat de retentiescores in het model hoogstwaarschijnlijk afhankelijk zijn van ontwerpvariaties en tijd. Per hypothese werden daarom in brms twee regressiemodellen gegenereerd. Ten eerste werd een multilevel regressiemodel aangemaakt waarbij het effect van de variabele tijd kon variëren tussen deelnemers. Daarnaast werd een regressiemodel aangemaakt waarin het effect van tijd niet varieerde tussen deelnemers. Per ontwerpvariatie werd nagegaan wat de voorspelling was voor kennisbehoud op langere termijn. Bovendien werd bekeken hoe de ontwerpvariaties zich ten opzichte van elkaar verhielden.

De meta-analyse bevatte op basis van vier samengevoegde onderzoeken in totaal 279 deelnemers, samen goed voor 939 meetmomenten. Bij de berekening van de posterior verdeling werd daarom een geloofwaardigheidsinterval van 89% toegepast. Kruschke (2014) en McElreath (2018) hebben immers aangegeven dat bij minder dan 4.000 deelnemers een geloofwaardigheidsinterval van 89% stabielere resultaten oplevert. De schaalverschillen

tussen de vier studies werden gecorrigeerd, alle testcores werden daarbij gestandaardiseerd naar een score op 10. Als prior voor testscore werd gekozen voor een waarde van 7,4 met een standaarddeviatie van 2,1. Dat waren respectievelijk de gemiddelde testscore en standaarddeviatie van de volledige onderzoeksgroep. Om stabiele resultaten te genereren, werd voor elk van de vier Markov-ketens het aantal gedraaide iteraties verhoogd tot 40.000, waarvan 15.000 warm-up-iteraties en 25.000 post-burn-in-iteraties. Dat gebeurde op aangeven van een waarschuwingsboodschap in brms. Om tijdens het draaien van de analyse afwijkende transacties te voorkomen die zouden kunnen wijzen op een slecht passend model, werd het aantal adaptaties verhoogd tot een waarde van 0.99.

2.2.2.2 Vergelijking Tussen Modellen. Nadat verschillende regressiemodellen aangemaakt waren, kon het verschil in hun verwachte voorspellende nauwkeurigheid ingeschat worden. Via het R-package loo (Vehtari et al., 2017) werd per hypothese bekeken of de voorspellende waarde van het MLM gestegen was ten opzichte van het model zonder multilevelstructuur. Dat kon via de elpd of *expected log pointwise predictive density* (Vehtari et al., 2017). Die theoretisch verwachte log-puntsgewijze voorspellende dichtheid voor een dataset kan worden ingeschat met behulp van *leave-one-out cross-validation* (LOO). Dat is een inschattingmethode voor de puntsgewijze out-of-sample voorspellingsnauwkeurigheid van Bayesiaanse regressiemodellen. De methode evalueert de log-likelihood bij posterieure simulaties van de parameterwaarden. De elpd_loo geeft de Bayesiaanse LOO-schatting weer voor een regressiemodel en drukt dus de voorspellende waarde van het model uit. Elpd_diff geeft op zijn beurt het verschil in elpd_loo weer tussen twee regressiemodellen. Daarbij wordt steeds het verschil berekend tussen de elpd van het best passende model ten opzichte van het andere model. De elpd_diff van het meest passende model zal dus steeds 0 zijn. De elpd_diff van het minder passende model zal altijd negatief zijn (Vehtari et al., 2017).

3. Resultaten

3.1 Regressieanalyses

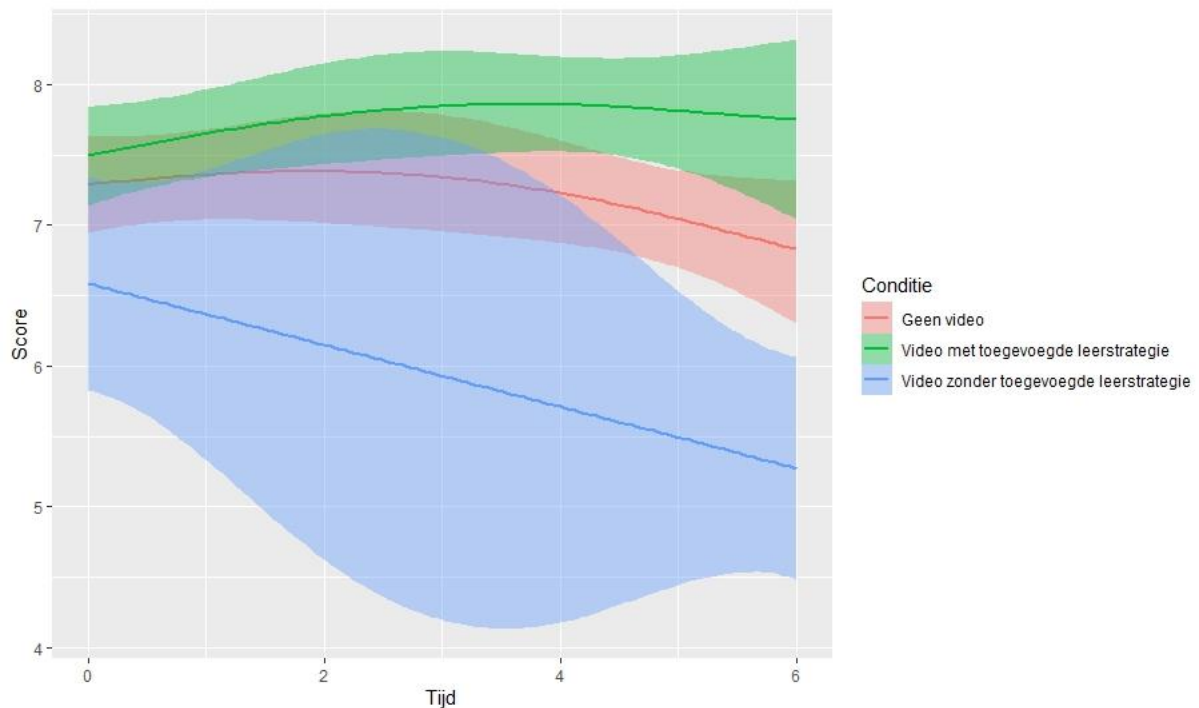
3.1.1 Uitkomstmaten

De regressieanalyses maakten een voorspelling over hoe de score van de deelnemers afhing van de conditie waarin ze zich bevonden en wat de samenhang was tussen score en het tijdsverloop na de interventie. Het effect van de onderzochte variabelen werd uitgedrukt in twee uitkomstmaten, de *pd* en het percentage overeenkomst tussen het HDI en de ROPE.

3.1.2 Hypothese 1

Het effect van tijd had een waarschijnlijkheid van 84,21% [*pd*] om negatief te zijn (*Mdn* = -0.03, 89% CI [-0.07, 0.02]) en kon beschouwd worden als niet significant (100% in ROPE). Het effect van instructievideo's met toegevoegde leerstrategieën had een waarschijnlijkheid van 99,72% [*pd*] om positief te zijn (*Mdn* = 0.56, 89% CI [0.24, 0.89]) en kon beschouwd worden als significant (0% in ROPE). Het effect van instructievideo's zonder toegevoegde leerstrategie had een waarschijnlijkheid van 99,53% [*pd*] om negatief te zijn (*Mdn* = -0.95, 89% CI [-1.53, -0.37]) en kon beschouwd worden als significant (0% in ROPE). De Rhat-waarde voor iedere onderzochte parameter bedroeg 1.000, wat wees op een sterke convergentie van het algoritme. Dat hield in dat er voldoende iteraties werden uitgevoerd en een sterke prior werd ingesteld (Gelman & Rubin, 1992). Figuur A1 in bijlage A geeft de details van de resultaten weer.

De regressieplot laat per conditie de voorspelling zien van de samenhang tussen score en het tijdsverloop na de interventie. Bovendien geeft de plot weer hoe de verschillende condities zich ten opzichte van elkaar verhouden (Figuur 1).

Figuur 1*Regressieplot Leerstrategieën*

Noot. De figuur presenteert het scoreverloop in de tijd per conditie. De score bij video's met toegevoegde leerstrategieën stijgt na zes weken, bij video's zonder toegevoegde leerstrategieën dalen de scores.

3.1.3 Hypothese 2

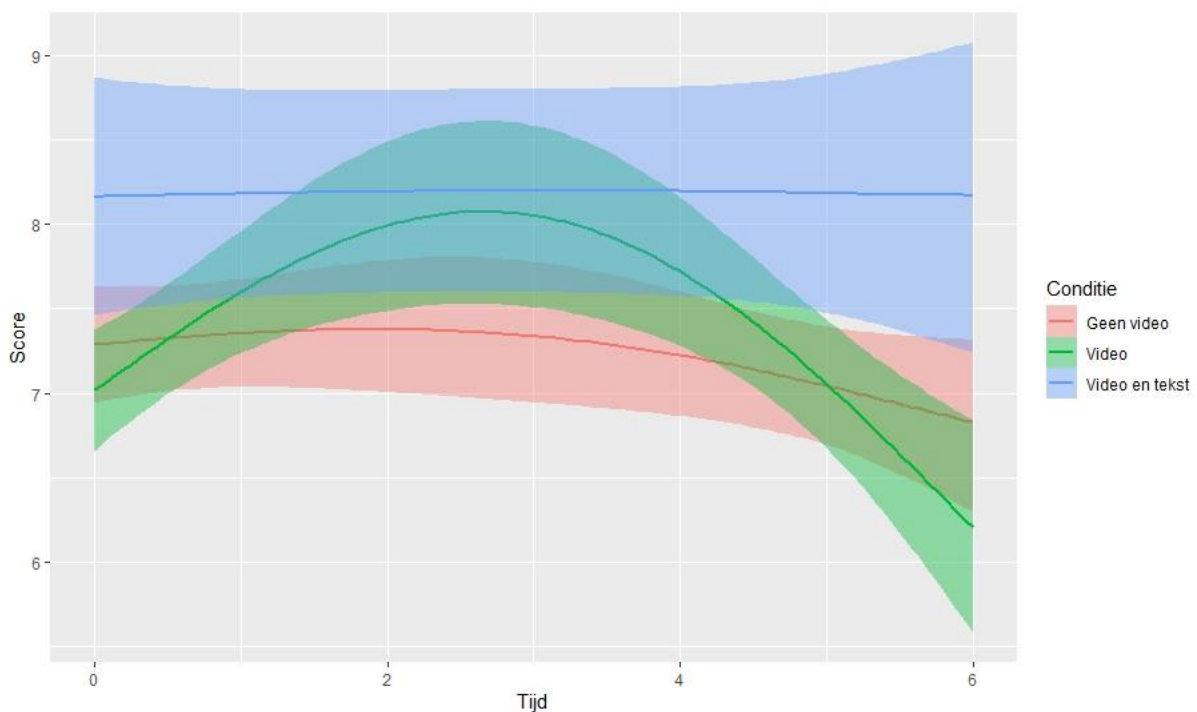
Het effect van tijd had een waarschijnlijkheid van 90.84% [*pd*] om negatief te zijn (*Mdn* = -0.03, 89% CI [-0.08, 0.01]) en kon beschouwd worden als niet significant (100% in ROPE). Terwijl het effect van instructievideo's zonder tekst een waarschijnlijkheid van 63.96% [*pd*] had om positief te zijn (*Mdn* = 0.07, 89% CI [-0.26, 0.42]), konden geen uitspraken gedaan worden over de significantie (38.62% in ROPE). Het effect van instructievideo's die gecombineerd werden met tekst, had een waarschijnlijkheid van 99.98% [*pd*] om positief te zijn (*Mdn* = 1.16, 89% CI [0.64, 1.69]) en kon beschouwd worden als

significant (0% in ROPE). De Rhat-waarde voor elke onderzochte parameter bedroeg opnieuw 1.000 (Bijlage A. Figuur A2).

De regressieplot geeft de voorspelling weer van hoe de score van de deelnemers afhangt van het tijdsverloop en de conditie waarin ze zich bevonden. De onderlinge verhouding tussen de condities is ook zichtbaar (Figuur 2).

Figuur 2

Regressieplot Video en Tekst



Noot. De figuur presenteert het scoreverloop in de tijd per conditie. De score bij lessen die video en tekst combineren blijft vrij egaal gedurende zes weken. De score bij lessen die enkel video's gebruikten stijgt eerst, om daarna af te nemen.

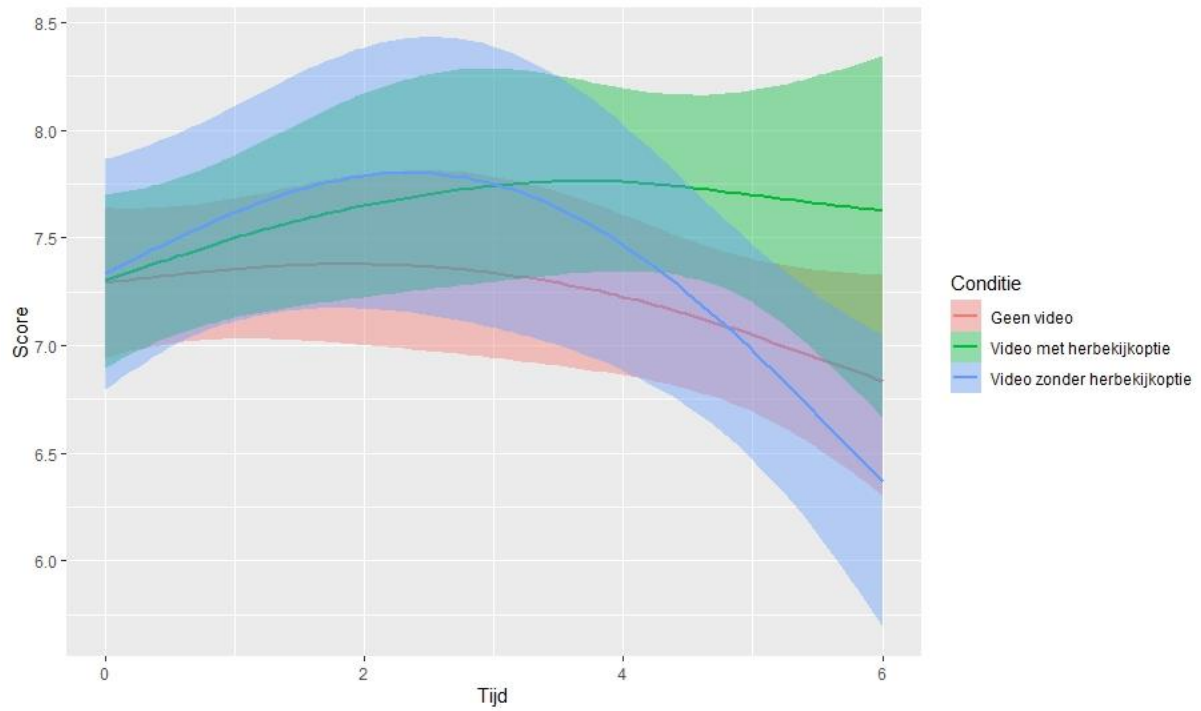
3.1.4 Hypotheses 3, 4 en 5

In hypothesen 3, 4 en 5 werd een regressieanalyse uitgevoerd met telkens dezelfde data verdeeld over drie conditiesgroepen die in ieder hypothese identiek waren. De hypothesen

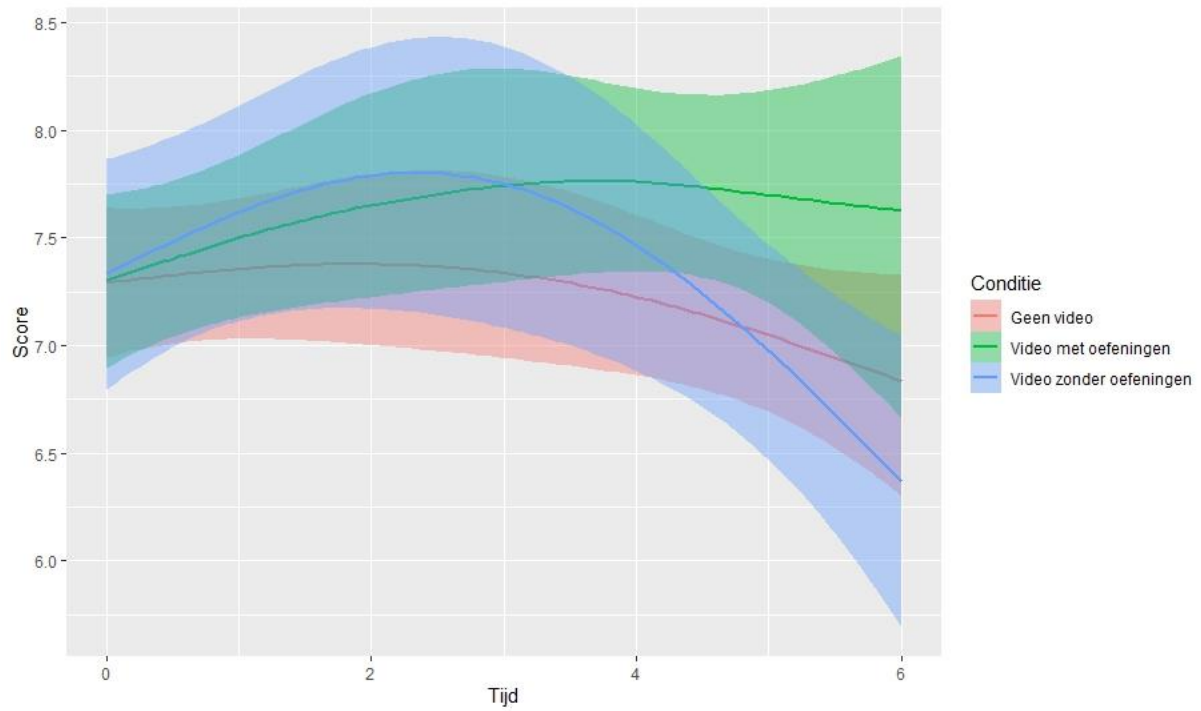
vergeleken dus dezelfde datagroepen, zij het telkens vanuit het perspectief van een ander ontwerpprincipe, met name de mogelijkheid tot video's herbekijken, de integratie van oefeningen in video's en de videolengte. De uitkomsten voor de drie hypothesen waren dan ook identiek en worden hieronder besproken voor zowel de derde (H3), vierde (H4) als vijfde (H5) hypothese.

Het effect van tijd had een waarschijnlijkheid van 87.57% [*pd*] om negatief te zijn (*Mdn* = -0.03, 89% CI [-0.07, 0.01] en kon beschouwd worden als niet significant (100% in ROPE). Het effect van korte instructievideo's (H5) zonder herbekijoptie (H3) en zonder oefeningen (H4) had een waarschijnlijkheid van 75.08% [*pd*] om positief te zijn (*Mdn* = 0.18, 89% CI [-0.25, 0.61]. Er konden echter geen uitspraken gedaan worden over de significantie (26.12% in ROPE). Terwijl het effect van lange instructievideo's (H5) met herbekijoptie (H3) en oefeningen (H4) een waarschijnlijkheid van 95.46% [*pd*] had om positief te zijn (*Mdn* = 0.36, 89% CI [0.02, 0.76], bleef de analyse onbeslist over de significantie (5.41% in ROPE). De Rhat-waarde voor elke onderzochte parameter bedroeg opnieuw 1.000. In Figuur A3, A4 en A5 van bijlage A zijn de details van de resultaten te raadplegen.

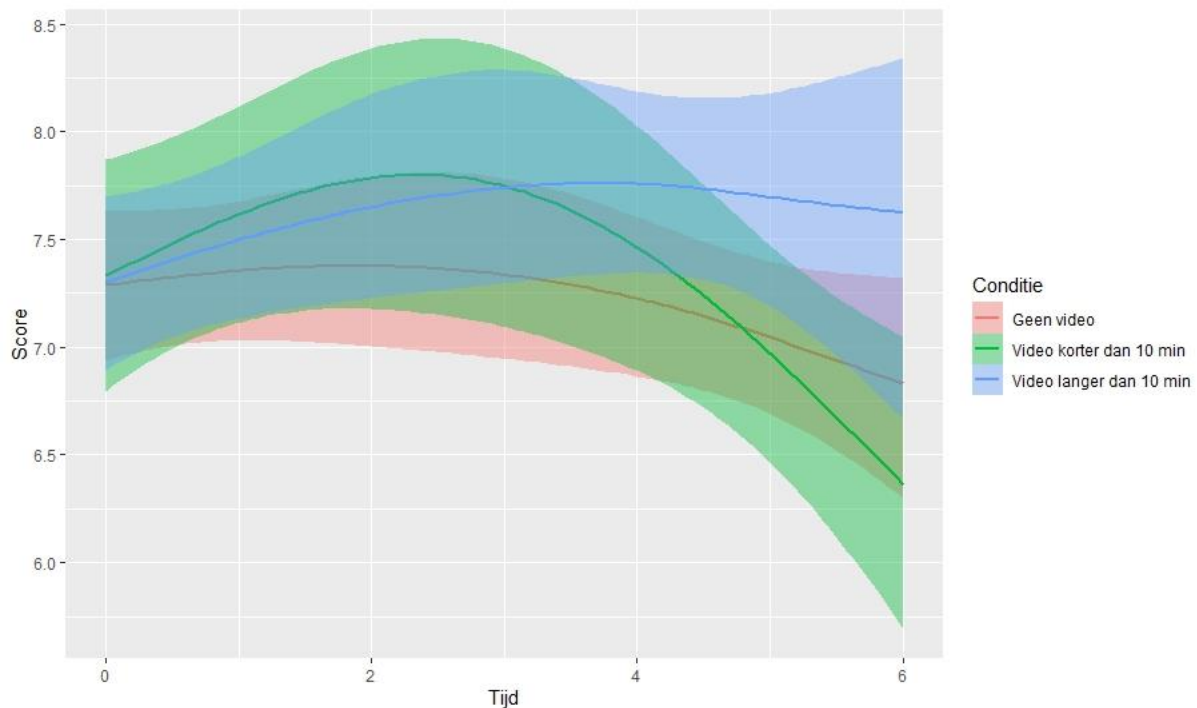
De regressieplot geeft per conditie de voorspelling weer van het verband tussen de score en het tijdsverloop na de interventie. Daarnaast laat de plot zien hoe de condities zich ten opzichte van elkaar verhouden (Figuur 3, 4 en 5).

Figuur 3*Regressieplot Herbekijkopties*

Noot. De figuur presenteert het scoreverloop in de tijd per conditie. De score bij video's met herbekijkoptie stijgt gedurende zes weken. De score bij video's zonder herbekijkoptie stijgt eerst, om daarna af te nemen.

Figuur 4*Regressieplot Oefeningen*

Noot. De figuur presenteert het scoreverloop in de tijd per conditie. De score bij video's met oefeningen stijgt gedurende zes weken. De score bij video's zonder oefeningen stijgt eerst, om daarna af te nemen.

Figuur 5*Regressieplot Videolengte*

Noot. De figuur presenteert het scoreverloop in de tijd per conditie. De score bij video's langer dan 10 minuten stijgt gedurende zes weken. De score bij video's korter dan 10 minuten stijgt eerst, om daarna af te nemen.

3.2 Vergelijking Tussen Modellen

In deze analyse werd ervoor gekozen om een multilevel regressiemodel aan te maken waarbij het effect van de variabele tijd kon variëren tussen verschillende deelnemers. Om na te gaan of dat regressiemodel nauwkeurige voorspellingen opleverde, werd een vergelijking gemaakt met een regressiemodel waarin het effect van tijd niet varieerde tussen deelnemers.

In de eerste hypothese leverde het multilevel regressiemodel veruit de nauwkeurigste voorspellingen op. Het regressiemodel waarin het effect van tijd niet varieerde tussen deelnemers scoorde opmerkelijk minder ($elpd_diff = -134.4$). In de tweede hypothese zorgde het regressiemodel waarin het effect van tijd niet varieerde tussen deelnemers opnieuw voor

de minst nauwkeurige voorspellingen ($elpd_diff = -137.7$). Ook in de derde, vierde en vijfde hypothese genereerde het multilevel regressiemodel veruit meer nauwkeurige voorspellingen dan het regressiemodel waarin het effect van tijd niet varieerde tussen deelnemers ($elpd_diff = -147.9$). Figuren B1, B2 en B3 in bijlage B geven de details van de resultaten weer.

4. Discussie

4.1 Beantwoording Onderzoeksvraag en Hypotheses

4.1.1 *Leerstrategieën*

De onderzoeksvraag luidde: wat is het effect van ontwerpvariëaties in instructievideo's op kennisbehoud in langetermijninterventies? Daarbij werden vijf hypothesen geformuleerd die zich elk toespitsten op een specifieke ontwerpvariëatie. In de eerste hypothese werd verwacht dat instructievideo's die gecombineerd werden met op retentie gerichte leerstrategieën, ten opzichte van lescondities zonder instructievideo betere leerresultaten opleverden dan instructievideo's die niet gecombineerd werden met op retentie gerichte leerstrategieën. Deze hypothese werd bevestigd door de meta-analyse. Instructievideo's met toegevoegde leerstrategieën scoorden meteen na de interventie beter dan lescondities zonder video, die op hun beurt beter scoorden dan instructievideo's zonder toegevoegde leerstrategie. De regressieanalyse liet zien hoe toetsresultaten na een periode van zes weken stegen voor instructievideo's met toegevoegde leerstrategie, daalden voor lescondities zonder video en nog sterker daalden voor instructievideo's zonder toegevoegde leerstrategie. In het onderzoek van Van der Spoel (2022) maakten de studenten gedurende zes weken iedere week een toets. Die wekelijkse herhaling kon gezien worden als een vorm van spaced practice en retrieval practice. Die twee strategieën kwamen ook terug in de studie van Van Hese (2022), waar een toets werd gemaakt meteen na de interventie, na een week en na een maand. Ook Palte (2022) integreerde op dezelfde manier spaced practice en retrieval practice in de interventie. Bij De Berg (2022) kwamen deze strategieën niet voor. De leerlingen maakten meteen na de

interventie een toets en zes weken later nog eens. Intussen werd de leerstof niet via tussentijdse toetsen of oefeningen opgepikt. Niettegenstaande de instructievideo van De Berg (2022) werd ontworpen volgens de ontwerpprincipes uit de CTML van Mayer (2021) en met aandacht voor een korte tijdsduur, bleken de verwachte positieve effecten daarvan toch niet op te wegen tegen het feit dat de leerstof gedurende de interventieperiode niet werd herhaald via tussentijdse oefeningen of toetsen. Lescondities zonder instructievideo maar met integratie van spaced practice en retrieval practice hadden een positievere impact op retentie dan lescondities met instructievideo, maar zonder toegevoegde leerstrategieën. Een kanttekening bij deze uitkomst is dat in deze meta-analyse de data in de conditie instructievideo zonder toegevoegde leerstrategie afkomstig waren uit slechts één onderzoek. De conditie instructievideo met toegevoegde leerstrategie bevatte daarentegen data uit drie onderzoeken. De specifieke resultaten uit het individuele onderzoek van De Berg (2022) hebben dus een grote impact op de uitkomst van deze hypothese, wat vraagt om enig voorbehoud bij het vormen van conclusies.

4.1.2 Combinatie van Tekst en Video

De tweede hypothese veronderstelde dat lescondities waarin het gebruik van instructievideo's en tekst gecombineerd werd, ten opzichte van lescondities zonder instructievideo betere leerresultaten opleverden dan lescondities die enkel instructievideo's inzetten. Deze hypothese werd ook bevestigd. De combinatie van video en tekst scoorde meteen na de interventie beter dan lescondities zonder video, die op hun beurt beter scoorden dan lescondities met enkel video's. De regressieanalyse bevestigde deze rangschikking ook na een verloop van zes weken. Een verklaring hiervoor kan zijn dat een tekst kan fungeren als ondersteuning en houvast bij het bekijken van leervideo's. Terwijl gepresenteerde gegevens in een videoweergave van voorbijgaande aard zijn, blijft informatie uit een tekst wel beschikbaar. Studenten ervaren via video's dus de voordelen van het modaliteits- en dubbele

coderingsprincipe, maar de simultane beschikbaarheid van een tekst vangt tegelijk één van de mogelijke nadelen van een instructievideo op. Het werkgeheugen kan immers maar voor een korte periode informatie vasthouden. Het vergankelijkheidseffect (Leahy & Sweller, 2011) treedt op als het werkgeheugen onvoldoende tijd krijgt om audiovisuele gegevens te verwerken en te integreren met voorkennis uit het langetermijngeheugen. Mogelijk verklaart dat de slechtere retentiescores van de videoconditie ten opzichte van de lesconditie zonder video. Bij deze hypothese moet echter opnieuw de bedenking gemaakt worden dat de data in de conditie waarin video en tekst gecombineerd werden, afkomstig waren uit één onderzoek, met name dat van Van der Spoel (2022). Dat noopt opnieuw tot voorzichtigheid bij het maken van conclusies.

4.1.3 Herbekijkopties, Geïntegreerde Oefeningen en Videolengte

Alvorens in te gaan op de resultaten van de derde, vierde en vijfde hypothese, is het belangrijk om te onderstrepen dat die hypothesen op vlak van conclusies niet los van elkaar bekeken kunnen worden. De hypothesen vergeleken immers dezelfde datagroepen, zij het telkens vanuit het perspectief van een ander ontwerpprincipe, met name de mogelijkheid tot video's herbekijken, de integratie van oefeningen in video's en de lengte van de video's. Bij de derde hypothese was de verwachting dat instructievideo's die herbekeken konden worden, ten opzichte van lescondities zonder instructievideo betere leerresultaten opleverden dan instructievideo's die slechts eenmaal bekeken konden worden. In de vierde hypothese werd verwacht dat instructievideo's waarin oefeningen geïntegreerd waren, ten opzichte van lescondities zonder instructievideo betere leerresultaten opleverden dan instructievideo's waarin geen oefeningen geïntegreerd waren. De vijfde hypothese veronderstelde ten slotte dat instructievideo's korter dan 10 minuten, ten opzichte van lescondities zonder instructievideo betere leerresultaten opleverden dan instructievideo's langer dan 10 minuten. De verschillen tussen de onderzochte datagroepen die werden vastgesteld in bijvoorbeeld de derde

hypothese, waren dus dezelfde verschillen die ook terug te vinden waren in de vierde en vijfde hypothese. Bij de bespreking van de resultaten is het daarom belangrijk om erbij stil te staan dat het vastgestelde verschil tussen de datagroepen niet noodzakelijk wijst op een effect van de specifieke ontwerpvariatie uit de hypothese. Het gevonden effect kan net zo goed het resultaat zijn van een andere ontwerpvariatie.

4.1.3.1 Herbekijkopties. Als we de resultaten bekeken vanuit het perspectief van de derde hypothese, die zich richtte op het herbekijken van video's, dan stelden we vast dat de hypothese bevestigd kon worden. Video's die herbekeken kunnen worden, scoorden volgens het regressiemodel na zes weken beter dan lescondities zonder video, die op hun beurt beter scoorden dan video's die niet herbekeken konden worden. Bij de impact van herbekijkbare video's moeten echter een paar specifieke ontwerpkeuzes van de betrokken onderzoeken in beschouwing genomen worden. In het onderzoek van Van Hese (2022) hadden studenten de vrije keuze om de video al dan niet te herbekijken. Slechts 25% van de deelnemers had de video echter meer dan één keer bekeken vóór het eerste meetmoment. De helft van de deelnemers had de video nog eens bekeken na het tweede meetmoment. De vraag stelt zich dus in welke mate de impact van de herbekijkoptie werkelijk doorwoog in de data uit het onderzoek van Van Hese (2022). In het onderzoek van Palte (2022) hadden deelnemers slechts deels een vrije keuze wat betreft het herbekijken van de video. Na het bekijken van een videosegment moesten ze een in de video geïntegreerde oefening maken. Op basis van het aantal fouten, kregen ze al dan niet het advies om het segment te herbekijken. Wie te veel fouten maakte, moest het segment sowieso verplicht herbekijken. Het was echter niet duidelijk hoeveel deelnemers de video hetzij vrijwillig of verplicht herbekeken hadden. Ook hier is het dus moeilijk te bepalen welk effect de herbekijkoptie had op de data.

4.1.3.2 Geïntegreerde Oefeningen. Bekeken we de resultaten vanuit het perspectief van de vierde hypothese, die zich focuste op integratie van oefeningen in instructievideo's,

dan zagen we dat ook die hypothese bevestigd kon worden. Video's waarin oefeningen geïntegreerd werden, scoorden volgens het regressiemodel na zes weken beter dan lescondities zonder video, die op hun beurt beter scoorden dan video's waarin geen oefeningen geïntegreerd werden. Zowel het onderzoek van Van Hese (2022) als dat van Palte (2022) bevatten oefeningen die sowieso door alle leerlingen in de videoconditie gemaakt werden. Dat maakt het aannemelijk dat het positief effect op retentie het resultaat is van geïntegreerde oefeningen. Die uitkomst was ook in de literatuur terug te vinden. Wachtler et al. (2016) hebben het effect onderzocht van oefeningen zoals interactieve vragen die in instructievideo's verwerkt waren. Ten opzichte van een controlegroep zonder instructievideo behaalden leerlingen die instructievideo's met interactieve vragen konden bekijken na een maand opmerkelijk betere resultaten. De uitkomst van deze hypothese lag dus in lijn met resultaten uit eerder onderzoek.

4.1.3.3 Videolengte. Bekeken we de resultaten ten slotte vanuit het perspectief van de vijfde hypothese, die de lengte van instructievideo's belichtte, dan vonden we geen bevestiging voor de hypothese. Video's langer dan 10 minuten, scoorden volgens het regressiemodel na zes weken beter dan lescondities zonder video, die op hun beurt beter scoorden dan video's korter dan 10 minuten. Die resultaten sloten niet aan op de literatuur, die stelde immers dat het modaliteitsprincipe enkel een positief effect heeft bij korte audiovisuele fragmenten (Leahy & Sweller, 2016). Eerdere onderzoeken gaven ook aan dat instructievideo's die maximum 6 minuten duren studenten het meest aanspreken (Brame, 2016; Clossen, 2018). Het was echter niet te bepalen in hoeverre het negatieve antwoord op deze hypothese het gevolg was van de lengte van de onderzochte video's. In deze meta-analyse maakten immers uitgerekend de onderzoeken met langere video's meer gebruik van andere ondersteuningsmiddelen die wel een positieve impact konden hebben op het retentieproces. Palte (2022) onderzocht video's van 25 minuten, Van Hese (2022) bestudeerde

video's van 15 minuten. In beide onderzoeken konden video's echter wel herbekeken worden en waren er oefeningen geïntegreerd in de video's. Dat kan een rol spelen in het feit dat deelnemers uit deze videocondities de beste retentieresultaten behaalden, ongeacht de lengte van de video's. Als we aan de andere kant de onderzoeken met kortere video's in beschouwing namen, dan zagen we dat in de studie van De Berg (2022) kennisclips van 5 minuten werden onderzocht. Toch gingen de leerresultaten na zes weken achteruit. De video's bevatten immers geen geïntegreerde oefeningen, konden niet herbekeken worden en waren niet omkaderd door andere leerstrategieën zoals tussentijdse toetsen. Van der Spoel (2022) onderzocht eveneens video's van 5 minuten die niet herbekeken konden worden en geen oefeningen inhielden. De leerlingen maakten gedurende zes weken wel wekelijks een toets, wat gezien kon worden als een vorm van spaced practice en retrieval practice. In deze meta-analyse blijkt de impact van oefeningen en herbekijkoefeningen op kennisbehoud dus groter te zijn dan de verwachte positieve effecten van een kortere videoduur. Hoe dan ook blijkt dat een onderzochte ontwerpvariantie zelden op zichzelf staat en bij de beoordeling van resultaten steeds samen met andere bewust of onbewust toegevoegde ontwerpvarianties in beschouwing genomen moet worden.

4.2 Beperkingen van het Onderzoek en Toekomstig Onderzoek

Er kunnen een aantal beperkingen in beschouwing genomen worden met betrekking tot het huidige onderzoek. Een eerste beperking betreft het aantal opgenomen studies. Er waren vier onderzoeken opgenomen in deze meta-analyse. Niettegenstaande Davey et al. (2011) en Valentine et al. (2010) hebben aangegeven dat meta-analyses met een beperkt aantal onderzoeken niet ongewoon zijn en bij voldoende homogeniteit nog steeds meer nauwkeurige schattingen opleveren dan individuele studies, blijft dat aantal relatief laag. Zolang de toevoeging van meer studies de homogeniteit niet hypothekeert, kunnen extra onderzoeken extra data aanleveren en ertoe bijdragen dat de samengevoegde resultaten van een meta-

analyse meer gerichte schattingen genereren over effecten in de populatie. Meer beschikbare studies geven ook meer beoordelingsruimte over welke onderzoeken geschikt zijn om opgenomen te worden in de meta-analyse. De beschikbaarheid van geen of een klein aantal gepubliceerde studies heeft bij deze concrete meta-analyse ertoe geleid dat er op grond van meer praktische beoordelingscriteria op zoek werd gegaan naar bruikbare studies. Daarbij werden vier onderzoeken gebundeld die enerzijds wel vergeleken konden worden op vlak van ontwerpvariaties, maar die anderzijds toch grondig van elkaar afweken op vlak van leeftijd van de onderzoeksgroep, aangereikte leerstof en manier van toetsen.

Dat leidt meteen tot de tweede beperking van de huidige meta-analyse. Meer beschikbare onderzoeken zouden meer ruimte geboden hebben om studies te bundelen die ook op vlak van andere niet-onderzochte variabelen overeenstemmen. Op die manier wordt de homogeniteit van de meta-analyse verhoogd en wordt ruis door variatie in niet-onderzochte variabelen zoals leeftijd of leerstof ingeperkt. Als er slechts een beperkt aantal empirische studies beschikbaar zijn die bovendien onderling variëren op een aantal punten die niet in beschouwing genomen zijn in het huidig model, kan de vraag gesteld worden of een meta-analyse überhaupt zinvol is. In de volgende paragraaf worden enkele verschillen tussen de geïncludeerde studies toegelicht.

Ten eerste bevatten de onderzoeken deelnemers uit het primair, voortgezet en hoger onderwijs. De manier waarop informatie in het werkgeheugen wordt verwerkt, varieert echter tussen die verschillende leeftijdsgroepen. Informatieverwerkingstheorieën benadrukken immers hoe de verwerkingslimiet van kinderen toeneemt door de natuurlijke groei van interne controle, zelfregulatie, werkgeheugen en automatisering naarmate ze ouder worden (Crone & Ridderinkhof, 2011). Daarbij wijzigt de absolute capaciteit van het werkgeheugen niet, maar de capaciteit functioneert wel efficiënter naarmate kinderen ouder worden (Case, 1991). Er

kan dus gesteld worden dat naargelang de leeftijd van leerlingen varieert, de manier waarop en de efficiëntie waarmee ze informatie uit instructievideo's verwerken, ook zal variëren.

Daarnaast onderscheiden de studies in deze meta-analyse zich ook van elkaar op vlak van aangereikte leerstof. Het effect van een instructievideo kan echter variëren tussen lessen wiskunde (Palte, 2022; Van Hese, 2022), taal (Van der Spoel, 2022) of vakinhoudelijke leerstof (De Berg, 2022). Volgens het modaliteits- en dubbele coderingsprincipe zouden de voordelen van multimedia immers meer doorwegen als puur visuele gegevens zoals meetkundige figuren of Romeinse cijfers gecombineerd kunnen worden met gesproken tekst.

De geïncludeerde onderzoeken variëren ook op vlak van de manier waarop ze via toetsen naar kennisbehoud peilden. De Berg (2022) en Van der Spoel (2022) gebruikten meerkeuzevragen, terwijl Palte (2022) en Van Hese (2022) open vragen stelden. Dat zijn twee methodes die peilen naar respectievelijk herkennen en herinneren, twee verschillende processen bij het ophalen van informatie uit het langetermijngeheugen (Driscoll, 1994). Aangezien de gebruikte toetsen een beroep doen op verschillende leerprocessen, kan dat ook een impact hebben op de resultaten.

Ondanks die kanttekeningen werd er toch voor gekozen om de meta-analyse uit te voeren met deze vier studies. De langetermijneffecten van instructievideo's op kennisbehoud vormen op dit moment immers een gat in de literatuur, terwijl het onderwerp toch maatschappelijk belangrijk en urgent is. Hoewel deze eerste bevindingen op basis van vier onderzoeken nog een aantal kanttekeningen inhouden, kunnen ze toch enig inzicht geven in de vraag of het huidige beperkte bewijs convergeert. Dat biedt een basis voor verder onderzoek. De uitvoering van deze meta-analyse vormt dus als het ware een proof of concept die aantoont dat de methodiek werkt, wat het draagvlak voor vervolgonderzoek kan vergroten.

De beperkingen van het huidige onderzoek leiden immers tot een aantal aanbevelingen voor toekomstige studies. Allereerst is het belangrijk dat er meer individuele onderzoeken

naar langetermijneffecten van instructievideo's op kennisbehoud uitgevoerd worden. Die onderzoeken leveren op zichzelf al extra inzichten op over online lessen en leervideo's, een groeiende maatschappelijke trend die vraagt om meer wetenschappelijk onderbouwde ontwerprichtlijnen. Bovendien bieden meer beschikbare onderzoeken ook kwantitatief en kwalitatief voordelen op bij de uitvoering van meta-analyses. Extra onderzoeken leveren extra data op die tot nauwkeurigere schattingen kunnen leiden. Daarnaast ontstaat ook de mogelijkheid om op basis van scherpere beoordelingscriteria onderzoeken op te nemen in meta-analyses. Dat houdt in dat een meta-analyse die zich bijvoorbeeld wil richten op het herbekijken van instructievideo's, onderzoeken includeert die zo veel mogelijk overeenstemmen op vlak van andere mogelijk impactvolle variabelen zoals leeftijd van de onderzoeksgroep, aangereikte leerstof of toegepaste toetsmethode.

Een laatste aanbeveling voor vervolgonderzoek betreft de eventuele uitvoering van een mixed methods onderzoeksdesign. Daarbij leveren enerzijds kwantitatieve data gegevens op die statistisch geanalyseerd kunnen. Anderzijds kunnen kwalitatieve data, zoals open-end interviews, verschillende perspectieven over het onderzoeksonderwerp opleveren (Creswell, 2014). Uit deze meta-analyse blijkt immers dat niet altijd eenduidig vast te stellen is welke ontwerpvariantie verantwoordelijk is voor een gevonden effect. Naast een kwantitatieve gegevensverzameling zouden studenten daarom ook bevraagd kunnen worden over welke aspecten van instructievideo's ze effectief en zinvol vinden. Eerder onderzoek heeft reeds aangegeven dat studenten verwachten dat video's leerstof bevattelijker maken (Miner & Stefaniak, 2018) en leerzamer zijn dan teksten (Harrison, 2019; Long et al., 2016). Bovendien blijkt dat ze video's boeiender en motiverender vinden dan boeken (Granitz et al., 2021). Hoewel onderzoeken het niet unaniem eens zijn over de vraag of een voorkeur voor instructievideo's ook resulteert in betere leeruitkomsten, staat wel vast dat leerlingen beter presteren naarmate zij meer gemotiveerd zijn (Boekaerts et al., 2010). Studenten hebben in

eerder onderzoek reeds aangegeven dat instructievideo's motiverend en geestelijk stimulerend zijn (Kay & Kletskin, 2012; Preston et al., 2010). Een mixed methods onderzoeksdesign waarin studenten enerzijds een toets afleggen na het bekijken van een instructievideo en anderzijds hun mening geven over ontwerpvarianties in de instructievideo, kan helpen verklaren waarom bepaalde ontwerpvarianties effect hebben. Ontwerpvarianties die studenten als zinvol ervaren, gaan immers ook impact hebben op hun motivatie, en zodoende ook op hun leerresultaten. Moreno (2005) en Mayer (2005; 2018; 2021) hebben bovendien aangegeven dat de huidige CTML uitgebreid zou kunnen worden met de mediërende rol van motivatie en metacognitieve strategieën zoals concentratie opbrengen, nieuwsgierig zijn en leerplezier hebben (Hattie, 2009). Kortom elementen waarnaar gepeild kan worden via mixed methods onderzoek.

4.3 Praktische Implicaties

Hoewel de hierboven uiteengezette belemmeringen aangegeven dat conclusies uit dit onderzoek met enig voorbehoud in beschouwing genomen moeten worden, kunnen toch enkele praktische implicaties opgesomd worden. Een eerste implicatie is van wetenschappelijke aard. Terwijl het langetermijneffect van instructievideo's op retentie maatschappelijk relevant en actueel is, is het onderwerp momenteel nog onvoldoende onderzocht. Deze meta-analyse is daarom uitgevoerd met data afkomstig uit een vrij beperkt aantal onderzoeken. Los van concrete uitkomsten en het voorbehoud waarmee die benaderd moeten worden, heeft dit onderzoek wel aangetoond dat de uitvoering van een meta-analyse haalbaar is en praktisch potentieel heeft. Bovendien is aangetoond dat het gebruik van een MLM daarbij ook rekening kan houden met praktische tijdsgebonden knelpunten die kunnen voorvallen wanneer longitudinale onderzoeken met verschillende meetpunten samengevoegd worden in één analyse. Deze meta-analyse kan dus beschouwd worden als een succesvolle proof of concept waarop toekomstig onderzoek kan voortbouwen.

Een tweede implicatie richt zich op concrete onderwijscontexten waarin instructievideo's ingezet worden. Het toenemend gebruik van instructievideo's leidt tot een grotere behoefte aan wetenschappelijk onderbouwde richtlijnen over het ontwerp van digitaal lesmateriaal. De ontwerpprincipes uit Mayers CTML kunnen daarbij een leidraad zijn, maar volstaan op zichzelf niet om kennisbehoud op lange termijn te garanderen. In het onderzoek van De Berg (2022) werd de instructievideo specifiek ontworpen op basis van Mayers ontwerpprincipes zoals overtolligheid, nabijheid, coherentie en signalering. Op die manier werd volgens de theorie cognitieve belasting beperkt en kon de informatie goed verwerkt worden. Toch gingen de toetsresultaten achteruit na zes weken. Palte (2022) en Van Hese (2022) hielden bij het ontwerp van hun instructievideo's ook rekening met Mayers ontwerpprincipes, maar combineerden die met extra leerstrategieën zoals herbekijkopties, geïntegreerde oefeningen en herhalingstoetsen. Dat levert volgens het regressiemodel na zes weken wel betere toetsresultaten op. Een video ontworpen volgens principes uit de CTML is dus niet louter effectief op basis van die principes alleen. Een positief effect wordt vooral bereikt door de implementatie van herbekijkmogelijkheden, oefeningen en tussentijdse toetsen. Leerkrachten kunnen hiermee rekening houden bij het lesontwerp.

Een laatste implicatie richt zich op onderzoek dat wordt uitgevoerd binnen de vakgroep Online leren en instructie van de Faculteit Onderwijswetenschappen van de Open Universiteit. Een van de doelen van deze vakgroep is het verstrekken van wetenschappelijk onderbouwde richtlijnen op basis waarvan het onderwijsveld efficiënt en effectief online lesmateriaal kan ontwerpen. Daarom zijn binnen de vakgroep verschillende scripties uitgevoerd waarbij het kennisbehoud na videolessen over een langere periode werd gemeten. Deze meta-analyse wou nagaan waaraan de verschillen in resultaten tussen die verschillende scripties konden liggen. Zo kunnen richtlijnen verstrekt worden voor toekomstig onderzoek naar langetermijneffecten van instructievideo's op retentie. Op basis daarvan kan een dubbele

implicatie geformuleerd worden voor de vakgroep Online leren en instructie. Enerzijds blijken naast het video-ontwerp ook leerstrategieën zoals herbekijkmogelijkheden, oefeningen en tussentijdse toetsen een impact te hebben op retentie. Toekomstige studies binnen de vakgroep kunnen daarmee rekening houden bij de formulering van onderzoeksvragen en het ontwerp van experimentele condities. Anderzijds leveren die toekomstige studies extra data op die de basis kunnen vormen voor nieuwe meta-analyses. Samen met de huidige studies kunnen nieuwe data tot meer nauwkeurige schattingen leiden over langetermijneffecten van ontwerpvariëaties in instructievideo's op retentie.

4.4 Conclusie

In deze meta-analyse zijn via Bayesiaanse regressiemodellen vijf hypothesen onderzocht die afgeleid waren uit de onderzoeksvraag naar wat het langetermijneffect is van verschillende ontwerpvariëaties van instructievideo's op retentie. De regressieanalyse geeft aan dat instructievideo's die gecombineerd worden met andere leerstrategieën na zes weken betere leerresultaten opleveren dan instructievideo's zonder toegevoegde leerstrategieën. Ook instructievideo's die gecombineerd worden met teksten scoren beter dan video's zonder tekst. Lange instructievideo's scoren na zes weken beter dan korte instructievideo's. Daarbij moet echter de bedenking gemaakt worden dat de lange video's herbekeken konden worden en ook geïntegreerde oefeningen bevatten. De mogelijkheid tot herbekijken, de integratie van oefeningen en het maken van tussentijdse toetsen lijken op basis van de regressieanalyse dan ook de condities te zijn die het meest positieve langetermijneffect hebben op kennisbehoud. Alle onderzochte videocondities bevatten instructievideo's die specifiek voor het onderzoek ontworpen werden op basis van Mayers ontwerpprincipes, of bestaande video's waarin Mayers ontwerpprincipes ook voorkwamen. Dat houdt in dat goed ontworpen instructievideo's op zichzelf niet garant staan voor kennisbehoud op lange termijn. In combinatie met andere didactische ingrepen tonen instructievideo's wel potentieel als

leermiddel. De instructie- en leerprincipes van onderwijsinstellingen en de didactische aanpak van leerkrachten spelen dus ook een rol. Die conclusie sluit aan op de vraag naar meer onderzoek naar het effect van instructievideo's in authentieke onderwijscontexten zoals klaslokalen. Mayers ontwerpprincipes zijn immers vooral gebaseerd op bevindingen uit gecontroleerde onderzoekomgevingen. Uitgevers van multimediale leermaterialen kunnen op basis van dit onderzoek instructievideo's ontwikkelen waarin Mayers ontwerpprincipes gecombineerd worden met herbekijkopties en oefeningen na ieder segment. Leerkrachten kunnen op hun beurt instructievideo's inzetten die enerzijds aansluiten op de voorkeur en leefwereld van studenten, en anderzijds in combinatie met oefeningen en tussentijdse toetsen ook een positieve impact hebben op kennisbehoud.

Referenties

Referenties gemarkeerd met een asterisk geven studies aan die zijn opgenomen in de meta-analyse.

Arguel, A., & Jamet, E. (2009). Using Video and Static Pictures to Improve Learning of Procedural Contents. *Computers in Human Behavior*, 25(2), 354-359.

<https://doi.org/10.1016/j.chb.2008.12.014>

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control process. In K. W. Spence & J. T. I. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic Press.

Bauer, D., & Curran, P. (2022, 6 oktober). *Intensive Longitudinal Data: A Multilevel Modeling Perspective* [Webinar]. American Psychological Association.

<https://centerstat.org/apa-ild/>

Berney, S., & Bétrancourt, M. (2016). Does Animation Enhance Learning? A Meta-Analysis. *Computers & Education*, 101, 150-167.

<https://doi.org/10.1016/j.compedu.2016.06.005>

Biard, N., Cojean, S., & Jamet, E. (2018). Effects of segmentation and pacing on procedural learning by video. *Computers in Human Behavior*, 89, 411-417.

<https://doi.org/10.1016/j.chb.2017.12.002>

Boekaerts M., Van Nuland, H. J. C., & Martens R. L. (2010). Perspectives on motivation: what mechanism energises students' behaviour in the classroom. In K. Littleton, C. Wood, & J. Kleine Staarman (Eds.), *International Handbook of Psychology in Education* (pp. 535-568). Emerald Group Publishing Limited.

Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE life sciences education*, 15(4), es6.

<https://doi.org/10.1187/cbe.16-03-0125>

- Bryk, A. S., & Raudenbush, S. W. (1988). Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model. *American Journal of Education*, 97(1), 65-108. <https://doi.org/10.1086/443913>
- Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Case, R. (1991). *The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge*. Lawrence Erlbaum Associates, Inc.
- Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, 8(4), 293-332. http://dx.doi.org/10.1207/s1532690xci0804_2
- Chandler, P. (2009). Dynamic Visualisations and Hypermedia: Beyond the “Wow” Factor. *Computers in Human Behavior*, 25(2), 389-392. <https://doi.org/10.1016/j.chb.2008.12.018>
- Chen, O., Woolcott, G., & Sweller, J. (2017). Using Cognitive Load Theory to Structure Computer-Based Learning Including MOOCs. *Journal of Computer Assisted Learning*, 33(4), 293-305. <https://doi.org/10.1111/jcal.12188>
- Cheung, M. W.-L., & Vijayakumar, R. (2016). A Guide to Conducting a Meta-Analysis. *Neuropsychology Review*, 26(2), 121-128. <https://doi.org/10.1007/s11065-016-9319-z>
- Clossen, A. S. (2018). Trope or trap? Roleplaying narratives and length in instructional video. *Information Technology & Librarians*, 37(1), 27-38. <https://doi.org/10.6017/ital.v37i1.10046>
- Cooper, D., & Higgins, S. (2015). The effectiveness of online instructional videos in the acquisition and demonstration of cognitive, affective and psychomotor rehabilitation

- skills. *British Journal of Educational Technology*, 46(4), 768-779.
<https://doi.org/10.1111/bjet.12166>
- Creswell, J.W. (2014). *Educational research: planning, conducting, and evaluating quantitative and qualitative research* (4e ed.). Pearson.
- Crone, E. A., & Ridderinkhof, K. R. (2011). The developing brain: from theory to neuroimaging and back. *Developmental Cognitive Neuroscience*, 1(2), 101-109.
<https://doi.org/10.1016/j.dcn.2010.12.001>
- Crowther, M., Lim, W., & Crowther, M. A. (2010). Systematic review and meta-analysis methodology. *Blood*, 116(17), 3140–3146. <https://doi.org/10.1182/blood-2010-05-280883>
- Davey, J., Turner, R. M., Clarke, M. J., & Higgins, J. P. (2011). Characteristics of meta-analyses and their component studies in the Cochrane database of systematic reviews: a cross-sectional, descriptive analysis. *BMC Medical Research Methodology*, 11, 160.
<https://doi.org/10.1186/1471-2288-11-160>
- *De Berg, E. (2022). *Kennisclip of Tekst? : Met Welk Medium Halen Hbo-Studenten Betere Toetsresultaten?* [Masterscriptie, Open Universiteit]. OU Student Theses.
<https://research.ou.nl/en/studentTheses/kennisclip-of-tekst>
- Dirkx, K. J. H., Joosten - ten Brinke, D., & Camp, G. (2019). *Ontwerprichtlijnen voor tussentijdse toetsen vanuit de geheugenpsychologie*.
<https://go.exlibris.link/VQHF9QY1>
- Driscoll, M. P. (1994). *Psychology of learning for instruction*. Allyn & Bacon.
- Engelbrecht, J., Harding, A., & Du Preez, J. (2007). Long-term retention of basic mathematical knowledge and skills with engineering students. *European journal of engineering education*, 32(6), 735-744. <https://doi.org/10.1080/03043790701520792>

- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (4e ed.). SAGE Publications Ltd.
- Filius, R., & Lam, I. (2010). Ervaringen met weblectures. *OnderwijsInnovatie*, 12(1), 30-34.
- Gelman A, Rubin DB (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4), 457–472. <https://doi.org/10.1214/ss/1177011136>.
- Ginns, P. (2005). Meta-Analysis of the Modality Effect. *Learning and Instruction*, 15(4), 313-331. <https://doi.org/10.1016/j.learninstruc.2005.07.001>
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3–8. <http://doi.org/10.2307/1174772>
- Granitz, N., Kohli, C., & Lancelotti, M. P. (2021). Textbooks for the YouTube generation? A case study on the shift from text to video. *Journal of Education for Business* 96(5), 299-307. <https://doi.org/10.1080/08832323.2020.1828791>
- Harrison, T. (2019). How distance education students perceive the impact of teaching videos on their learning. *Open Learning: The Journal of Open, Distance and e-Learning*, 35(3), 260-276. <https://doi.org/10.1080/02680513.2019.1702518>
- Hattie, J. (2009). *Visible learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. SAGE Publications.
- Hegarty, M., & Kriz, S. (2008). Effects of Knowledge and Spatial Ability on Learning from Animation. In R. K. Lowe & W. Schnotz (Eds.), *Learning with Animation: Research Implications for Design* (pp. 3-29). Cambridge University Press.
- Henderson, M. L., & Schroeder, N. L. (2021). A systematic review of instructor presence in instructional videos: Effects on learning and affect. *Computers and Education Open*, 2, 100059. <https://doi.org/10.1016/j.caeo.2021.100059>
- Herreid, C. F., & Schiller, N. A. (2013). Case studies and the flipped classroom. *Journal of College Science Teaching*, 42(5), 62-67.

Hojtink, H., Mulder, J., van Lissa, C., & Gu, X. (2019). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods*, 24(5), 539–556.

<https://doi.org/10.1037/met0000201>

Holzer J., Korlat S., Haider C., Mayerhofer M., Pelikan E., Schober B., Spiel C., Toumazi T., Salmela-Aro K., Käser U., Schultze-Krumbholz A., Wachs S., Dabas M., Verma S., Iliev D., Andonovska-Trajkovska D., Plichta P., Pyzalski J., Walter N., ... Lüftenegger M. (2021). Adolescent well-being and learning in times of COVID-19—A multi-country study of basic psychological need satisfaction, learning behavior, and the mediating roles of positive emotion and intrinsic motivation. *PLoS ONE*, 16(5), Artikel e0251352.

<https://doi.org/10.1371/journal.pone.0251352>

Hopkins, R. F., Lyle, K. B., Hieb, J. L., & Ralston, P. A. S. (2016). Spaced Retrieval Practice Increases College Students' Short- and Long-Term Retention of Mathematics Knowledge. *Educational psychology review*, 28(4), 853-873.

<https://doi.org/10.1007/s10648-015-9349-8>

Jansen, B., & Van Casteren, W. (2021). Digitale leermiddelen in het hoger onderwijs.

<https://www.versnellingsplan.nl/wp-content/uploads/2021/09/Rapport-zone-digitale-leermaterialen-in-hoger-onderwijs.pdf>

Kay, R., & Kletskin, I. (2012). Evaluating the use of problem-based video podcasts to teach mathematics in higher education. *Computers & Education*, 59(2), 619-627.

<https://doi.org/10.1016/j.compedu.2012.03.007>

Kruschke, J. K. (2014). *Doing Bayesian Data Analysis—2nd Edition*. Academic Press.

Kruschke, J. K. (2018). Rejecting or accepting parameter values in Bayesian estimation.

Advances in Methods and Practices in Psychological Science, 1(2), 270-280.

<https://doi.org/10.1177/2515245918771304>

Kruschke, J. K., & Liddell, T. M. (2017). Bayesian data analysis for newcomers.

Psychonomic Bulletin & Review, 1–23. <https://doi.org/10.3758/s13423-017-1272-1>

Kubsch, M., Stamer, I., Steiner, M., Neumann, K., & Parchmann, I. (2021). Beyond p-values:

Using Bayesian Data Analysis in Science Education Research. *Practical Assessment, Research, and Evaluation*, 26(4). <https://doi.org/10.7275/vzpw-ng13>

Kühl, T., Scheiter, K., Gerjets, P., & Gemballa, S. (2011). Can Differences in Learning

Strategies Explain the Benefits of Learning from Static and Dynamic Visualizations? *Computers & Education*, 56(1), 176-187.

<https://doi.org/10.1016/j.compedu.2010.08.008>

Leahy, W., Chandler, P., & Sweller, J. (2003). When Auditory Presentations Should and

Should Not Be a Component of Multimedia Instruction. *Applied Cognitive*

Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 17(4), 401-418. <https://doi.org/10.1002/acp.877>

Leahy, W., & Sweller, J. (2011). Cognitive Load Theory, Modality of Presentation and the

Transient Information Effect. *Applied Cognitive Psychology*, 25(6), 943-951.

<https://doi.org/10.1002/acp.1787>

Leahy, W., & Sweller, J. (2016). Cognitive Load Theory and the Effects of Transient

Information on the Modality Effect. *Instructional Science*, 44(1), 107-123.

<https://doi.org/10.1007/s11251-015-9362-9>

Leo, J., & Puzio, K. (2016). Flipped instruction in a high school science classroom. *Journal of*

Science Education and Technology, 25, 775–781. <https://doi.org/10.1007/s10956-016-9634-4>

Long, T., Logan, J., & Waugh, M. (2016). Students' perceptions of the value of using videos

as a pre-class learning experience in the flipped classroom. *TechTrends: Linking*

Research & Practice to Improve Learning, 60(3), 245-252.

<https://doi.org/10.1007/s11528-016-0045-4>

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework.

Journal of Open Source Software, 4(40), 1541. <https://doi.org/10.21105/joss.01541>

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85-139. [https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6)

Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 41-61). Cambridge University Press.

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, 63(8), 760-769.

<https://doi.org/10.1037/0003-066X.63.8.760>

Mayer, R. E. (2018). Thirty Years of Research on Online Learning. *Applied Cognitive Psychology*, 33(2), 152-159. <https://doi.org/10.1002/acp.3482>

Mayer, R. E. (2021). Evidence-based principles for how to design effective instructional videos. *Journal of Applied Research in Memory and Cognition*, 10(2), 229-240.

<https://doi.org/10.1016/j.jarmac.2021.03.007>

Mayer, R. E., & Chandler, P. (2001). When Learning Is Just a Click Away: Does Simple User Interaction Foster Deeper Understanding of Multimedia Messages? *Journal of educational psychology*, 93(2), 390. <https://doi.org/10.1037/0022-0663.93.2.390>

Mayer, R. E., & Fiorella, L. (2014). 12 principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. Cambridge University Press.

- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, 38(1), 43-52.
https://doi.org/10.1207/S15326985EP3801_6
- Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. Cambridge University Press.
- McElreath, R. (2018). *Statistical Rethinking: A Bayesian Course with Examples in R and STAN*. Routledge & CRC Press.
- Miner, S., & Stefaniak, J. E. (2018). Learning via video in higher education: An exploration of instructor and student perceptions. *Journal of University Teaching and Learning Practice*, 15(2), 5-19. <https://doi.org/10.53761/1.15.2.2>
- Ministerie van Onderwijs, Cultuur en Wetenschap. (2020, 24 december). *Wetenschappelijke handreiking over de risico's en kansen van afstandsonderwijs*. [Persbericht]
<https://www.onderwijsinspectie.nl/actueel/nieuws/2020/12/24/wetenschappelijke-handreiking-over-de-risicos-en-kansen-van-afstandsonderwijs>
- Mitra, B., Lewin-Jones, J., Barrett, H., & Williamson, S. (2010). The use of video to enable deep learning. *Research in Post-Compulsory Education*, 15(4), 405-414.
<https://doi.org/10.1080/13596748.2010.526802>
- Moreno, R. (2005). Instructional technology: promise and pitfalls. In L. M. PytlikZillig, M. Bodvarsson, & R. H. Bruning (Eds.), *Technology-Based Education: Bringing Researchers and Practitioners Together* (pp. 1–19). Information Age Pub.
- Moreno, R. (2006). Does the modality principle hold for different media? A test of the methods-affect-learning hypothesis. *Journal of Computer Assisted Learning*, 22(3), 149-158. <https://doi.org/10.1111/j.1365-2729.2006.00170.x>

- Moreno, R., & Mayer, R. (2007). Interactive Multimodal Learning Environments. *Educational Psychology Review*, 19(3), 309-326. <https://doi.org/10.1007/s10648-007-9047-2>
- Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing Cognitive Load by Mixing Auditory and Visual Presentation Modes. *Journal of educational psychology*, 87(2), 319. <https://doi.org/10.1037/0022-0663.87.2.319>
- Murre, J. M. J., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PloS ONE*, 10(7), e0120644-e0120644. <https://doi.org/10.1371/journal.pone.0120644>
- Nuzzo, R. (2014). Scientific method: Statistical errors. *Nature*, 506(7487), 150–152. <https://doi.org/10.1038/506150a>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716–aac4716. <https://doi.org/10.1126/science.aac4716>
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 45(3), 225-287. <https://doi.org/10.1037/h0084295>
- *Palte, I. (2022). *Effectieve Instructievideo als Alternatief voor de Directe Instructie: Het Langetermijneffect op Leerresultaten* [Masterscriptie, Open Universiteit]. OU Student Theses. <https://research.ou.nl/en/studentTheses/effectieve-instructievideo-als-alternatief-voor-de-directe-instru>
- Preston, G., Phillips, R., Gosper, M., McNeill, M., Woo, K., & Green, D. (2010). Web-based lecture technologies: Highlighting the changing nature of teaching and learning. *Australasian Journal of Educational Technology*, 26. <https://doi.org/10.14742/ajet.1038>

- Rasi, P., & Poikela, S. (2016). A Review of Video Triggers and Video Production in Higher Education and Continuing Education PBL Settings. *Interdisciplinary Journal of Problem-Based Learning*, 10(1). <https://doi.org/10.7771/1541-5015.1609>
- Rickes, P. C. (2016). Generations in flux: How Gen Z will continue to transform higher education space. *Planning for Higher Education*, 44(4), 21-45.
- Schneps, M. H., Griswold, A., Finkelstein, N., McLeod, M., & Schrag, D. P. (2010). Using Video to Build Learning Contexts Online. *Science*, 328(5982), 1119-1120. <https://doi.org/10.1126/science.1186934>
- Schnotz, W., & Lowe, R. K. (2008). A Unified View of Learning from Animated and Static Graphics. In R. K. Lowe & W. Schnotz (Eds.), *Learning with Animation: Research Implications for Design* (pp. 304-356). Cambridge University Press.
- Schunk, D. H. (2014). *Learning theories: An educational perspective*. Pearson Education Limited.
- Seeger, M. (2004). Gaussian processes for machine learning. *International Journal of Neural Systems*, 14(2), 69–106. <https://doi.org/10.1142/S0129065704001899>.
- Seufert, T., Schütze, M., & Brünken, R. (2009). Memory characteristics and modality in multimedia learning: an aptitude-treatment-interaction study. *Learning and Instruction*, 19(1), 28-42. <https://doi.org/10.1016/j.learninstruc.2008.01.002>
- Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, 10(2), 176-199. <https://doi.org/10.1177/1745691615569000>
- Surma, T., Vanhoyweghen, K., Sluijsmans, D., Camp, G., Muijs, D., & Kirschner, P. A. (2019). *Wijze lessen, 12 bouwstenen voor effectieve didactiek*. Ten Brink Uitgevers.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design.

Learning and Instruction, 4(4), 295-312. [https://doi.org/10.1016/0959-4752\(94\)90003-5](https://doi.org/10.1016/0959-4752(94)90003-5)

Sweller, J., & Chandler, P. (1991). Evidence for Cognitive Load Theory. *Cognition and instruction*, 8(4), 351-362. https://doi.org/10.1207/s1532690xci0804_5

Sweller, J., & Chandler, P. (1994). Why Some Material Is Difficult to Learn. *Cognition and instruction*, 12(3), 185-233. https://doi.org/10.1207/s1532690xci1203_1

Sweller, J., van Merriënboer, J. J., & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational psychology review*, 10(3), 251-296.

<https://doi.org/10.1023/a:1022193728205>

Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, 32, 261-292.

<https://doi.org/10.1007/s10648-019-09465-5>

Szymkowiak, A., Melović, B., Dabić, M., Jeganathan, K., & Singh Kundie, G. (2021).

Information technology and Gen Z: The role of teachers, the internet, and technology in the education of young people. *Technology in Society*, 65, 101565.

<https://doi.org/10.1016/j.techsoc.2021.101565>

Tarchi, C., Zaccoletti, S., & Madon, L. (2021). Learning from text, video, or subtitles: A comparative analysis. *Computers & Education*, 160, 104034.

<https://doi.org/10.1016/j.compedu.2020.104034>

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can It Facilitate?

International journal of human-computer studies, 57(4), 247-262.

<https://doi.org/10.1006/ijhc.2002.1017>

Vaganova, O., Rudenko, I., Markova, S., Smirnova, Z., & Kutepov, M. (2019). The use of educational video materials in educational process of a higher educational institution.

Amazonia Investiga, 8(22), 216-222.

<https://amazoniainvestiga.info/index.php/amazonia/article/view/308>

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, 35(2), 215–247. <https://doi.org/10.3102/1076998609346961>

*Van der Spoel, I. (2022). *Video's bij de Woordenschatles van Nieuwsbegrip in het Primair Onderwijs* [Masterscriptie, Open Universiteit]. OU Student Theses.

<https://research.ou.nl/en/studentTheses/videos-bij-de-woordenschatles-van-nieuwsbegrip-in-het-primair-ond>

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E.-J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, 28(3), 813–826.

<https://doi.org/10.3758/s13423-020-01798-5>

*Van Hese, N. (2022). *Het Effect van een flipped-Classroomdesign in de Wiskundeles op het Retentievermogen van Leerlingen Secundair Onderwijs* [Masterscriptie, Open

Universiteit]. OU Student Theses. <https://research.ou.nl/en/studentTheses/het-effect-van-een-flipped-classroomdesign-in-de-wiskundeles-op-h>

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413-1432. <https://doi.org/10.1007/s11222-016-9696-4>

Vlaanderen – Onderwijs en Vorming. (2021). *Digisprong*.

<https://onderwijs.vlaanderen.be/digisprong>

- Wachtler, J., Hubmann, M., Zöhrer, H., & Ebner, M. (2016). An analysis of the use and effect of questions in interactive learning-videos. *Smart Learning Environments*, 3(1), 13. <https://doi.org/10.1186/s40561-016-0033-3>
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, 25(1), 35–57. <https://doi.org/10.3758/s13423-017-1343-3>
- Wang C., Hsu H.-C. K., Bonem E. M., Moss J. D., Yu S., Nelson D. B., & Levesque-Bristol C. (2019). Need satisfaction and need dissatisfaction: a comparative study of online and face-to-face learning contexts. *Computers in Human Behavior*, 95, 114-125. <https://doi.org/10.1016/j.chb.2019.01.034>
- Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science*. <https://doi.org/10.1177/1745691611406923>
- Witt, J.-A., Glöckner, C., & Helmstaedter, C. (2011). Extended retention intervals can help to bridge the gap between subjective and objective memory impairment. *Seizure (London, England)*, 21(2), 134-140. <https://doi.org/10.1016/j.seizure.2011.10.007>
- Zeger, S. L., & Qaqish, B. (1988). Markov regression models for time series: A quasi-likelihood approach. *Biometrics*, 44, 1019–1031. <https://doi.org/10.2307/2531732>

Bijlage A**Bayesiaanse Regressieanalyses****Figuur A1***Resultaten Regressieanalyse Video's met en Zonder Leerstrategieën*

Parameter	<i>Mdn</i>	89% CI	<i>pd</i>	ROPE	% in ROPE	Rhat
Intercept	7.21	[6.95, 7.46]	100%	[-0.10, 0.10]	0%	1.000
Tijd	-0.03	[-0.07, 0.02]	84.21%	[-0.10, 0.10]	100%	1.000
Video met leerstrategie	0.56	[0.24, 0.89]	99.72%	[-0.10, 0.10]	0%	1.000
Video zonder leerstrategie	-0.95	[-1.53, -0.37]	99.53%	[-0.10, 0.10]	0%	1.000

Noot. CI = Credible Interval, ROPE = Region Of Practical Equivalence

Figuur A2*Resultaten Regressieanalyse Video's wel en Niet Gecombineerd met Tekst*

Parameter	<i>Mdn</i>	89% CI	<i>pd</i>	ROPE	% in ROPE	Rhat
Intercept	7.23	[6.98, 7.48]	100%	[-0.10, 0.10]	0%	1.000
Tijd	-0.03	[-0.08, 0.01]	90.84%	[-0.10, 0.10]	100%	1.000
Video zonder tekst	0.07	[-0.26, 0.42]	63.96%	[-0.10, 0.10]	38.62%	1.000
Video met tekst	1.16	[0.64, 1.69]	99.98%	[-0.10, 0.10]	0%	1.000

Noot. CI = Credible Interval, ROPE = Region Of Practical Equivalence

Figuur A3*Resultaten Regressieanalyse Video's met en Zonder Herbekijkoctie*

Parameter	<i>Mdn</i>	89% CI	<i>pd</i>	ROPE	% in ROPE	Rhat
Intercept	7.21	[6.96, 7.47]	100%	[-0.10, 0.10]	0%	1.000
Tijd	-0.03	[-0.07, 0.01]	87.57%	[-0.10, 0.10]	100%	1.000
Video zonder herbekijkoctie	0.18	[-0.25, 0.61]	75.08%	[-0.10, 0.10]	26.12%	1.000
Video met herbekijkoctie	0.39	[0.02, 0.76]	95.46%	[-0.10, 0.10]	5.41%	1.000

Noot. CI = Credible Interval, ROPE = Region Of Practical Equivalence

Figuur A4*Resultaten Regressieanalyse Video's met en Zonder Oefeningen*

Parameter	<i>Mdn</i>	89% CI	<i>pd</i>	ROPE	% in ROPE	Rhat
Intercept	7.21	[6.96, 7.47]	100%	[-0.10, 0.10]	0%	1.000
Tijd	-0.03	[-0.07, 0.01]	87.57%	[-0.10, 0.10]	100%	1.000
Video zonder oefeningen	0.18	[-0.25, 0.61]	75.08%	[-0.10, 0.10]	26.12%	1.000
Video met oefeningen	0.39	[0.02, 0.76]	95.46%	[-0.10, 0.10]	5.41%	1.000

Noot. CI = Credible Interval, ROPE = Region Of Practical Equivalence

Figuur A5*Resultaten Regressieanalyse Korte en Lange Video's*

Parameter	<i>Mdn</i>	89% CI	<i>pd</i>	ROPE	% in ROPE	Rhat
Intercept	7.21	[6.96, 7.47]	100%	[-0.10, 0.10]	0%	1.000
Tijd	-0.03	[-0.07, 0.01]	87.57%	[-0.10, 0.10]	100%	1.000
Video korter dan 10 min	0.18	[-0.25, 0.61]	75.08%	[-0.10, 0.10]	26.12%	1.000
Video langer dan 10 min	0.39	[0.02, 0.76]	95.46%	[-0.10, 0.10]	5.41%	1.000

Noot. CI = Credible Interval, ROPE = Region Of Practical Equivalence

Bijlage B**Vergelijking Tussen Multilevel Modellen****Figuur B1***Vergelijking Regressiemodellen met en Zonder Multilevelstructuur voor Hypothese 1*

Model	ELPD_diff	SE_diff
Model met multilevelstructuur	0.0	0.0
Model zonder multilevelstructuur	-134.4	18.8

Noot. ELPD_diff = difference in expected log pointwise predictive density, SE_diff = difference in standard error

Figuur B2*Vergelijking Regressiemodellen met en Zonder Multilevelstructuur voor Hypothese 2*

Model	ELPD_diff	SE_diff
Model met multilevelstructuur	0.0	0.0
Model zonder multilevelstructuur	-137.7	18.3

Noot. ELPD_diff = difference in expected log pointwise predictive density, SE_diff = difference in standard error

Figuur B3

Vergelijking Regressiemodellen met en Zonder Multilevelstructuur voor Hypotheses 3, 4 en 5

Model	ELPD_diff	SE_diff
Model met multilevelstructuur	0.0	0.0
Model zonder multilevelstructuur	-147.9	18.8

Noot. ELPD_diff = difference in expected log pointwise predictive density, SE_diff = difference in standard error