



rijksuniversiteit
groningen

faculteit gedrags- en
maatschappijwetenschappen

De effectiviteit van toetsreflectie op toetsresultaat door middel van
zelfbeoordeling van het leerproces en beoordeling van het eigen werk
van leerlingen in het voortgezet onderwijs.

R. van Uhm (S2384159)

Assessor: dr. D. D. N. M. Kostons

Tweede assessor: N. Renting

Rijksuniversiteit Groningen

Faculteit der Gedrags- en Maatschappijwetenschappen

Masterthesis Onderwijswetenschappen – track onderwijsinnovatie

juni 2023

Abstract

To expand cognitive knowledge, it is important to know the appropriate learning strategies and self-regulate the learning process when using them. In secondary education, little is known about the use of learning strategies. Among students in the first two grades of secondary education, at levels vmbo-tl, vmbo-tl-havo, havo en vwo ($n=149$), significant differences in students' approaches were found. To find so, a questionnaire was administered twice, considering *educational levels, grade* and the subjects *mathematics, history* and *Dutch*.

Because reflection is an important skill in the learning process, the questionnaire was administered after taking a test, and significant indications were found that this form of reflection contributes to an adjusted use of learning strategies, higher academic achievement ($p<.001$), and more accurate self-assessment of one's own work ($p<.01$).

However, the used questionnaire has its limitations. Further research with an adapted questionnaire and a larger sample size can determine whether the found differences are generalizable to students in secondary education.

Inleiding

Omdat het zelfregulerend leren veel aspecten omvat die invloed hebben op de kwaliteit van het leren van leerlingen, is er de afgelopen decennia is er in het onderwijsveld veel onderzoek naar uitgevoerd. Het zelfregulerend leren beschrijft cognitieve, metacognitieve, gedragsmatige, motivationele en emotionele aspecten van het leren (Panadero, 2017), welke worden verdeeld onder cognitieve, metacognitieve en motivationele leerstrategieën (Ten Dam & Vermunt, 2003). De vaardigheid om het leerproces zelf te reguleren draagt er aan bij dat leerlingen in de toekomst zelfstandig en effectiever nieuwe kennis kunnen verwerven en het leerproces, waar nodig, kunnen bijsturen (Boekaerts & Simons, 1995). Door het ontwikkelen van zelfregulerend vermogen leren leerlingen eigen verantwoordelijkheid te nemen, doelen te stellen en zelfstandig keuzes te maken in uiteenlopende contexten (Zimmerman, 2000). Ook in het onderwijscurriculum krijgt zelfregulerend leren in toenemende mate een belangrijke rol, waar het wordt omschreven als: ‘het zelfstandig handelen van leerlingen en de verantwoordelijkheid die zij nemen in bepaalde contexten’ (Curriculum.nu, 2018).

Er zijn in de onderwijswetenschappen meerdere modellen tot stand gekomen om het proces van het zelfregulerend leren te omschrijven, waarvan de modellen van Pintrich (2000) en Zimmerman (2000) de meest gebruikte modellen zijn (Panadero, 2017). In beide modellen spelen motivationele aspecten de belangrijkste rol in de zelfregulerende vaardigheden, opgevolgd door (meta)cognitieve aspecten. Deze aspecten zijn verdeeld over een cyclus met drie fasen van het leren: *voor het leren*, *tijdens het leren* en *na het leren*. *Voor het leren* stelt de leerling doelen en een plan van aanpak op, waarbij de leerling zichzelf motiveert en activeert (Zimmerman, 2000). *Tijdens het leren* dient de leerling controle te hebben op de uitvoering van de leertaak door de juiste leerstrategieën toe te passen, zich te concentreren op de taak en het leerproces te monitoren (Pintrich, 2000). *Na het leren* vindt reflectie op het proces en het product plaats, waarbij de leerling het eigen werk beoordeelt en vervolgstappen bepaalt. Dit

laatste proces beïnvloedt de eerste fase van een volgend leerproces, waardoor de cyclus rond is. Het reflecteren in de laatste fase leidt tot nieuwe doelen en een verbeterde aanpak in een nieuwe cyclus (Silver, 2013), maar vooral tot meer motivatie wanneer de leerling een soortgelijke taak in de toekomst aan gaat pakken (Zimmerman, 2000).

De reflectieve vaardigheid beperkt zich echter niet alleen tot de derde fase van de cyclus. Er bestaat een onderscheid tussen *reflection-in-action*, het reflecteren tijdens het proces, en *reflection-on-action*, het reflecteren op het proces na afloop (Schön, 1983; 1987). Het ontwikkelen van de reflectieve vaardigheid heeft daarom een positieve weerslag op het gehele leerproces: de voorbereiding, uitvoering, en de reflectie zelf. Dignath & Büttner (2008) vonden in hun onderzoek naar effectieve interventies voor het vergroten van metacognitie in het voortgezet onderwijs een grotere effectgrootte wanneer het onderdeel reflectieve vaardigheid is opgenomen in de interventie.

Daarnaast is inzicht in en beoordelen van het eigen kunnen essentieel voor zelfregulerend leren (Baars, 2014). Leerlingen kunnen immers het eigen leerproces niet reguleren wanneer ze het eigen leerproces niet kunnen beoordelen. Met name leerlingen met een relatief laag prestatieniveau op bepaalde gebieden vinden het lastig om de waarde van hun eigen vaardigheden op die specifieke gebieden in te schatten. Als leerlingen vaardiger worden in de leertaak zelf, zien ze in wat ze nog precies moeten leren om verder te komen en kunnen daardoor zichzelf beter beoordelen (Kruger & Dunning, 1999). Hoe meer een leerling zich ontwikkelt in de metacognitieve vaardigheden, hoe efficiënter het leerproces van die leerling zal zijn (Dignath & Büttner, 2008). Voor deze ontwikkeling is het nodig dat leerlingen zicht ontwikkelen op eigen capaciteiten, handelen en leervaardigheid: metacognitie. De cognitieve en metacognitieve aspecten gaan daarom hand in hand om tot zelfregulerend leren te komen.

Om cognitieve kennis uit te breiden, is het van belang de juiste aanpak te kennen en te gebruiken. Om leerstof goed te leren moet een leerling toewerken naar de situatie waarbij het

mogelijk is om de juiste leerstof terug te halen op het moment dat het nodig is: *cue availability*. De leerling moet herkennen wat er gevraagd wordt en welke geleerde informatie nodig is: *cue diagnosticity* (Karpicke et al., in press). Het geleerde moet in het geheugen worden aangevuld en georganiseerd, zodat kennis over het onderwerp zich steeds meer uitbreidt en de opgedane kennis zich meer en meer onderscheidt van andere dingen die bekend zijn over gerelateerde onderwerpen (Hunt & McDaniel, 1993). Deze studie richt zich op vier leerstrategieën die het meest effectief zijn bevonden om leerstof op een goede manier in het geheugen te kunnen opslaan, te kunnen ordenen, uit te breiden en vervolgens op commando te kunnen ophalen (Karpicke et al., in press).

- *Teaching and explaining*: het navertellen van de leerstof alsof je de docent bent (Fiorella & Mayer, 2014).
- *Retrieval practice*: het meermaals herhalen en overhoren van de stof. Niet alleen zelftesten en checken wat je beheerst, maar blijven herhalen (Karpicke & Nunes, 2015).
- *Spacing practice*: het verdelen van leermomenten over de tijd, door de stof iedere dag in kleine stukjes tot je te nemen (Toppino & Gerbier, 2014).
- *Interleaved practice*: het verdelen van leermomenten over de tijd, door verschillende soorten vragen en kennis te combineren, in plaats van alle kennis van dezelfde soort achter elkaar te plakken (Taylor & Roher, 2010)

Onderzoek naar de meest effectieve combinatie van de strategieën is nauwelijks gedaan (O'Day & Karpicke, 2020). Wel is er in eerder uitgevoerd onderzoek geen effect gevonden dat *repetitive reading* (het herhaaldelijk lezen van een tekst) een effectieve leerstrategie is (Dunlosky et al., 2013), terwijl leerlingen dit vaak wel als effectieve strategie beschouwen (Blasiman, Dunlosky & Rawson, 2017).

Er is steeds meer bekend over de aspecten die zelfregulerend leren in verschillende *leerjaren, niveaus* en *vakgebieden* van het onderwijs omvatten. Zo is er een positieve

samenhang tussen de kennis die leerlingen over een vakgebied hebben en de efficiëntie van zelfregulering binnen dat vakgebied (Kruger & Dunning, 1999). Dit verklaart ook waarom leerlingen van hogere leerjaren en in hogere onderwijsniveaus (havo en vwo) van het voortgezet onderwijs over het algemeen over verder gevorderde zelfreguleringsvaardigheden beschikken dan leerlingen van lagere leerjaren en lagere onderwijsniveaus (vmbo). Dezelfde leerstrategie kan verschillend worden uitgevoerd door leerlingen van verschillende leerjaren en onderwijsniveaus, waarbij er vooral verschillen bestaan in de vormgeving en inhoud van de strategie (De Boer et al., 2013). Met betrekking tot het verschil in de aanpak en het reguleren van het leren van leerlingen van verschillende *leerjaren, vakgebieden* en *onderwijsniveaus* is er in het voortgezet onderwijs nog te weinig bekend om gegronde uitspraken over te kunnen doen (Kostons et al., 2014). Er zijn enkele onderzoeken gedaan naar de verschillen in het zelfregulerend leren van leerlingen van het vmbo en havo/vwo (Van der Neut et al., 2005), maar weinig over de verschillen binnen de stromingen van het vmbo, of de gemengde stroming vmbo-tl/havo.

Om de wetenschappelijke inzichten over zelfregulering toepasbaar te maken voor de onderwijspraktijk is nodig om inzicht te krijgen in het verschil van aanpak van leerlingen van verschillende onderwijsniveaus, zodat leerlingen op de juiste manier ondersteund en gemonitord kunnen worden in de ontwikkeling van zelfregulerende vaardigheden (Moos & Ringdal, 2012; Onderwijsraad, 2016). Leerkrachten in het primair onderwijs lijken het zelfregulerend leren bij leerlingen meer te stimuleren dan leerkrachten in het voortgezet onderwijs, waarbij er in het voortgezet onderwijs sprake is van een gebrek aan instructie over de leerstrategieën die tot zelfregulerend leiden (Moos & Ringdal, 2012). Onderzoeken hebben zich vooral gericht op de werking van het gehele zelfregulerend proces, maar niet ingezoomd op de werking van specifieke processen, zoals zelfreflectie (Panadero, 2017).

In dit onderzoek is in kaart gebracht welke strategieën leerlingen van verschillende onderwijsniveaus in het voortgezet onderwijs toepassen wanneer zij zich voorbereiden op een toets. Door middel van een toetsreflectievragenlijst reflecteren leerlingen herhaaldelijk op de toetsvoorbereiding en het proces tijdens het maken van de toets. Er wordt gemeten of leerlingen het leergedrag aanpassen bij het volgende toetsmoment doordat ze op de aanpak hebben gereflecteerd aan de hand van de vragenlijst. Leerlingen schatten daarnaast welk cijfer ze denken te hebben gehaald op de toets om te meten of verwachtingen van de leerling over de kwaliteit van het werk aansluiten bij de werkelijkheid en om te meten of de schatting na het reflecteren op eerdere toetsmomenten accurater wordt.

De volgende twee onderzoeksvragen, met enkele deelvragen, staan centraal in dit onderzoek:

1. Welke verschillen zijn er in de keuze en in de aanpak van leerstrategieën tussen leerlingen van verschillende onderwijsniveaus, verschillende leerjaren en verschillende vakgebieden van het voortgezet onderwijs?
 - 1.1. Welke verschillen zijn er in de keuze en de aanpak van leerstrategieën tussen leerlingen van verschillende onderwijsniveaus?
 - 1.2. Welke verschillen zijn er in de keuze en de aanpak van leerstrategieën tussen leerlingen van verschillende leerjaren?
 - 1.3. Welke verschillen zijn er in de keuze en de aanpak van leerstrategieën van leerlingen tussen verschillende vakgebieden?
2. In hoeverre heeft het gebruik van de toetsreflectievragenlijst een positief effect op het schoolresultaat en het zelfbeoordelingsvermogen van leerlingen in het voortgezet onderwijs?
 - 2.1. Wat is het effect van de toetsreflectievragenlijst op het schoolresultaat van leerlingen in het voortgezet onderwijs?

2.2. In hoeverre bevordert de toetsreflectievragenlijst de nauwkeurigheid van de beoordeling van het eigen werk en het eigen proces van leerlingen in het voortgezet onderwijs?

2.3. In hoeverre passen leerlingen in het voortgezet onderwijs het leergedrag aan nadat zij hebben gereflecteerd op de aanpak van het leren aan de hand van de toetsreflectievragenlijst?

Er wordt verwacht dat er verschillen bestaan in de aanpak van leerlingen van verschillende niveaus en leerjaren (De Boer et al., 2013; Kostons et al., 2014; Van der Neut et al., 2005). Het is waarschijnlijk dat leerlingen van havo en vwo meer complexere leerstrategieën zullen gebruiken, zoals *interleaved practice* en *teaching and explaining*. Ook is de verwachting dat leerlingen naarmate ze in een hoger leerjaar komen, meer complexe leerstrategieën zullen gebruiken. Ook wordt er verwacht dat voor verschillende vakgebieden op een andere manier wordt geleerd. Dit verschil is eerder in het basisonderwijs al aangetoond (Kostons et al., 2014).

Daarnaast is het te verwachten dat de interventie een positief effect heeft op het academische resultaat van alle leerlingen, doordat er meer aandacht aan de aanpak van het voorbereiden en het maken van de toets wordt besteed door hierop te reflecteren. Leerlingen worden zich zo meer bewust van de punten die zij hebben laten liggen op de toets. Vergeleken met omliggende landen besteden Nederlandse leerlingen in het voortgezet onderwijs gemiddeld minder tijd aan toetsopgaven (Feskens et al., 2020). Bij het invullen van de toetsreflectievragenlijst staan de leerlingen stil bij eventuele overgeslagen opgaven en het geven van volledige antwoorden op vragen.

Ook wordt een accuratere zelfbeoordeling van het eigen werk verwacht, doordat leerlingen worden aangespoord kritisch naar het eigen werk te kijken en hier een beoordeling op te geven. Wanneer het daadwerkelijke resultaat afwijkt van het behaalde resultaat, gaat de

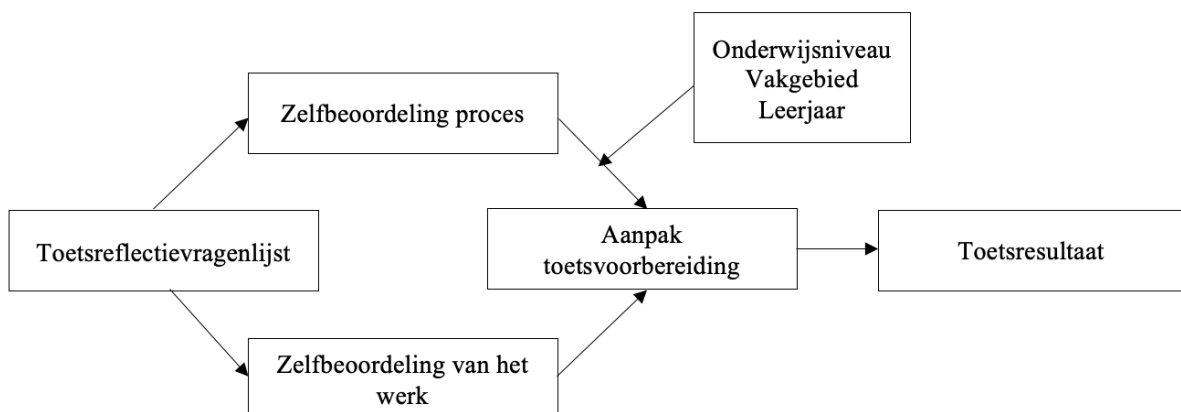
leerling na waar het verschil vandaan komt. De leerling breidt op die manier de kennis over het eigen werk en de aansluitingen bij de verwachtingen van de docent uit.

Daarnaast werken leerlingen aan een verbetering van het zelfregulerend vermogen, doordat zij uitgebreider reflecteren op het proces van het voorbereiden en maken van de toets dan dat ze voorheen deden (Dignath & Büttner, 2008). In de toetsreflectievragenlijst worden leerlingen bevraagd over het gebruik van de effectieve leerstrategieën. Bij veelvuldig negatief moeten beantwoorden van de vraag en een tegenvallend toetsresultaat, komen leerlingen tot realisatie dat er iets moet veranderen in de aanpak van het leren.

Het verwachte effect van de interventie wordt weergegeven in onderstaand conceptueel model.

Figuur 1

Conceptueel Model Invloed Toetsreflectievragenlijst op Toetsresultaat



Methode

Onderzoeksonwerp

Om de onderzoeksvraag te beantwoorden is een quasi-experimenteel onderzoek uitgevoerd met een *Pretest Posttest Control Group Design*, waarbij twee groepen met ieder een andere behandeling zijn gevolgd. De interventiegroep (R_1 , $n = 70$) krijgt les volgens de gebruikelijke situatie en reflecteert na iedere toets aan de hand van de toetsreflectievragenlijst

(X). De controlegroep (R_2 , $n = 79$) krijgt ook les volgens de gebruikelijke situatie maar gebruikt tweede de toetsreflectievragenlijst niet.

Steekproef en populatie

De populatie die onderzocht wordt zijn alle leerlingen uit het reguliere voortgezet onderwijs in leerjaar 1 en 2. Het onderzoek is uitgevoerd in bestaande schoolklassen van dezelfde school, dus participanten zijn niet random ingedeeld. Wel zijn de klassen als geheel random toegewezen aan een van de condities via een *Cluster Randomized Control Trial (CRCT)*. Alle leerlingen uit de klas doen vervolgens automatisch mee aan de toegewezen conditie, maar leerlingen weten niet dat zij onder verschillende condities hun taak uitvoeren. Omdat uitkomsten op individueel niveau met elkaar worden vergeleken, worden ten behoeve van de externe validiteit de omstandigheden zoveel mogelijk gelijk gehouden binnen de groepen: voor ieder vak hebben de deelnemende klassen daarom dezelfde docent, en ook de manier van lesgeven en toetsen is door deze docent zoals het zonder onderzoek ook zou zijn geweest.

Er hebben twee eerste klassen ($n = 39$) en vijf tweede klassen ($n = 110$) deelgenomen aan de studie. De leeftijd van de deelnemers binnen de onderzoeksgroep is dan ook tussen 12 en 15 jaar. In alle onderzoeksgroepen zitten zowel jongens ($n = 77$) als meisjes ($n = 72$).

De twee eerste klassen nemen deel aan het onderzoek met het vak wiskunde, waarbij zij door dezelfde docent worden onderwezen. Hiervan bestaat een klas volledig uit vmbo-kb leerlingen, de andere klas is een gecombineerde havo/vwo klas. De vijf deelnemende tweede klassen hebben deelgenomen aan het onderzoek met de vakken geschiedenis of Nederlands onder leiding van dezelfde docent van het desbetreffende vak. Een van deze klassen is een gecombineerde vmbo-bb/kb klas, twee klassen doet vmbo-tl niveau, een klas doet tl/havo niveau en een klas is een gecombineerde havo/vwo klas. Er is geen overlap van leerlingen in de klassen, waardoor wordt uitgesloten dat leerlingen een dubbele interventie zullen hebben.

In Tabel 1 zijn de kenmerken van de deelnemers over de verschillende conditiesgroepen weergegeven. Deze steekproef maakt het mogelijk om vast te stellen of er een verschil in aanpak zit tussen leerlingen van verschillende *onderwijsniveaus*, verschillende *leerjaren* en verschillende *vakgebieden*. Leerlingen die een toets hebben gemist of de toetsreflectievragenlijst niet hebben ingevuld zijn niet in het databestand opgenomen. Leerlingen die een vraag open hebben gelaten op de toetsreflectievragenlijst zijn wel meegenomen in de statistische analyses.

Tabel 1

Verdeling van Onderzoekspopulatie over de Conditiegroepen

	jaar		vak			niveau					
	1	2	wis	ges	ned	bb	kb	tl	tl-h	h	v
R ₁	22	48	22	26	22	0	0	22	14	22	12
R ₂	17	52	17	26	36	4	27	11	11	20	6
totaal	39	110	39	52	58	4	27	33	25	42	18

Met uitvoerende leerkrachten zijn afspraken gemaakt over de uitvoering van de interventie en het delen van data. Zij weten op voorhand aan welke groep de condities zijn toegekend, en passen de toegewezen interventies toe in de daarvoor geselecteerde klassen. Om er op toe te zien dat de interventies op dezelfde manier worden toegepast en de interne validiteit van het onderzoek gewaarborgd blijft, is een informatiedocument en een stappenplan opgesteld. Deze informeren de leerkrachten over de toepassing per conditiesgroep. De leerkrachten zijn in dit document ook geïnstrueerd over de gepseudonimiseerde registratie van toetsresultaten en de

resultaten op de toetsreflectievragenlijst. De instructie voor deelnemende leerkrachten is in bijlage A toegevoegd.

Instrument

De toetsreflectievragenlijst bestaat uit 10 vragen, waarvan een aantal deelvragen bevat. In totaal beantwoorden leerlingen 25 stellingen, doormiddel van een twee-, drie of vijfpuntsschaal en twee stellingen waar een getal ingevuld moet worden. De gebruikte toetsreflectievragenlijst is bijgevoegd in bijlage B.

Om tot de te meten constructen te komen is gebruik gemaakt van de ontwerpprincipes voor deeltaakoefening van het 4C/ID model (Hoogveld et al., 2017). De leerstrategieën die leerlingen kunnen gebruiken wanneer zij zich op een goede manier voorbereiden op de toets zijn eerst samengebracht in een vaardighedenhiërarchie. Er is daarbij gefocust op de vier effectieve leerstrategieën *teaching and explaining*, *retrieval practice*, *spacing practice* en *interleaved practice*. In de vaardighedenhiërarchie worden de diverse taakaspecten en hun onderlinge samenhang zichtbaar. Aan de hand van deze hiërarchie zijn prestatiedoelen opgesteld (Gagné, 1985). Als prestatiedoelen goed geformuleerd zijn, geven ze een heldere beschrijving van de te verwachten prestaties, in dit geval op het gebruik van leerstrategieën, na het doorlopen van het onderwijsprogramma (Hoogveld et al., 2017). Ten slotte kan met de leerling worden besproken welke leeractiviteiten het meest geschikt zijn om aan de verbeterpunten te werken (Kicken, 2008). De vaardighedenhiërarchie en prestatiedoelen die zijn opgesteld ter voorbereiding op het samenstellen van de toetsreflectievragenlijst zijn bijgesloten in bijlage C.

Een veelgebruikte vragenlijst om zelfregulerend leren in kaart te brengen is de *Motivated Strategies for Learning Questionnaire (MSLQ)* (Pintrich et al., 1991). Deze is ontwikkeld voor studenten in het hoger onderwijs en kent een aangepaste vorm voor leerlingen van het middelbaar onderwijs (Pintrich & De Groot, 1990). Deze vragenlijst bevat 41 items en

meet veel constructen die voor dit onderzoek niet van toepassing zijn. Om focus te houden op de constructen die voor dit onderzoek belangrijk zijn is een vragenlijst opgesteld met een aantal vragen uit de subsectie *Cognitive and meta-cognitive strategies* van de *MSLQ*. Deze vragen zijn omgezet naar items met antwoordschalen die passend zijn bij de prestatiedoelen van dit onderzoek. De gemeten constructen en vragen per construct zijn weergegeven in Figuur 2.

Omdat het hoofddoel van de vragenlijst is om te achterhalen welke leerstrategieën leerlingen gebruiken, is het voor vragen die zich op die constructen richten niet nodig een uitgebreide antwoordschaal aan te bieden. Uit een meta-analyse van de *MSLQ* is daarnaast gebleken dat het gebruik van stellingen met een waarderingsschaal er toe kan leiden dat er assumpties worden gedaan die niet geheel juist zijn. Zo kan het onduidelijk zijn of een negatieve waardering staat voor het niet hebben uitgevoerd van de taak of het niet goed hebben uitgevoerd van de taak (Credé & Phillips, 2011). Ook kan de hoeveelheid antwoordcategorieën per stelling het aantal overgeslagen vragen van respondenten beïnvloeden (Toepoel, 2008). Voor vragen die moeten meten welke leerstrategieën zijn gebruikt, vragen 6 en 8, is een tweepuntsschaal gekozen, waarbij de leerling *ja* of *nee* kan antwoorden. Voor vragen waar naast *ja* en *nee* ook een antwoord zoals *gedeeltelijk* mogelijk is, vragen 4 en 5, is een middencategorie toegevoegd en een driepuntsschaal gebruikt (Baarda, 2017). Voor vragen 1, 2, 3 en 9 is een vijfpuntsschaal gebruikt om meer genuanceerd te meten in welke mate de leerling zichzelf beoordeelt.

Voor antwoorden waarin participanten een schatting moeten verrichten, heeft een open karakter van de antwoordmogelijkheid de voorkeur (Toepoel, 2008). Voor de vragen 7 en 10 is daarom voor een open antwoordmogelijkheid waar een getal kan worden ingevuld gekozen.

Figuur 2

Gemeten Constructen en Vragen per Construct

Gemeten Construct	Vraag in de vragenlijst
teaching and explaining	4; 8d
retrieval practice	2; 5; 8c
spacing practice	7; 8a; 8b
interleaved practice	1; 3; 6; 8e
toetsreflectie	9; 10

Literatuur om de vaardighedenhiërarchie, prestatiedoelen en toetsreflectievragenlijst op te bouwen is verkregen via ERIC, waarbij de volgende zoektermen zijn gebruikt: *reflection or reflective or reflective practice, effective learning methods, intervention or strategies or best practices, reflection on action, secondary education or secondary school or high school, self-regulated learning or self regulated learning* en *self-regulated learning questionnaire*.

De toetsreflectievragenlijst dient naast meetinstrument ook als interventie: bij het invullen ervan reflecteren de leerlingen op de voorbereiding van en het proces tijdens het maken van een toets. Zij hebben hierdoor de mogelijkheid hun eigen leerproces bij een volgend toetsmoment bij te sturen. De resultaten van deze vragenlijst zijn gekoppeld aan de behaalde resultaten op de toets waarover de vragenlijst is ingevuld om te onderzoeken of het herhaaldelijk gebruiken van de toetsreflectievragenlijst een positief effect heeft op het toetsresultaat.

Om de betrouwbaarheid van de toetsreflectievragenlijst te meten is een betrouwbaarheidsanalyse uitgevoerd. Hiermee is de validiteit van de vragenlijst bekeken. Aan de hand van de factoranalyse is vastgesteld hoeveel constructen de vragenlijst meet. Daarnaast

is de correlatie tussen de vragen berekend om te meten of deze een significant verband met elkaar hebben.

Onderzoeksprocedure

Op 21 februari 2023 zijn de deelnemende docenten geïnstrueerd over de deelname van hun klassen aan het onderzoek en de aan de klassen toegewezen conditie. Tussen 6 maart en 18 mei 2023 hebben de docenten de resultaten van leerlingen bijgehouden en de voorgeschreven interventies uitgevoerd in de interventiegroepen R_1 en R_2 . De leerkrachten hebben voor de voormetingen de twee meest recent behaalde resultaten geregistreerd in het bijgeleverde databestand. De lijst is vervolgens tweemaal aangevuld na het afnemen van een nieuwe toets.

De leerlingen uit alle klassen hebben gedurende de onderzoeksperiode volgens de gebruikelijke manier les gehad. Leerlingen uit de klassen die zijn toegewezen aan de interventieconditie hebben gedurende die periode tweemaal direct na het maken van een toets de toetsreflectievragenlijst ingevuld op hun iPad, via het digitale platform van de school. Het moment van afnemen van de toetsen en de vragenlijst, is afhankelijk van het door de docent geplande toetsmoment. Het eerste toetsmoment binnen de onderzoeksperiode heeft in week 10, 13 of 14 plaatsgevonden. Het tweede toetsmoment heeft voor alle klassen in week 20 plaatsgevonden. Het invullen van de toetsreflectievragenlijst neemt iedere keer 5 tot 10 minuten in beslag.

Data-analyse

Aan de hand van de toetsreflectievragenlijst is vastgesteld welke leerstrategieën leerlingen hebben gebruikt. Om de eerste onderzoeksvraag met de beide deelvragen te beantwoorden, worden de ingevulde toetsreflectievragenlijsten met elkaar vergeleken per *onderwijsniveau, leerjaar en vakgebied*. Met een correlatietabel is de samenhang van de items en groepsvariabelen gemeten. De samenhang wordt significant gevonden wanneer $p < .001$.

Daarnaast is met deze vragenlijst gemeten of leerlingen na het gebruik van de reflectievragenlijst het leergedrag aanpassen in de voorbereiding op latere toetsmomenten. Daarvoor is de vragenlijst bij alle leerlingen in interventiegroep R₁ tweemaal afgenomen, met daartussen een toetsmoment. Met een *Paired Samples t-test* is gemeten of de gemiddelden en standaarddeviaties hetzelfde zijn bij de tweede afname ten opzichte van de eerste. Indien de toetsreflectievragenlijst leerlingen aanzet tot verandering in gedrag bij de toetsvoorbereiding, is een significant verschil in gemiddelden tussen de metingen verwacht. Voor deze test geldt de volgende hypothese: $H_0: \mu_{XA1} = \mu_{XA2}$; $H_a: \mu_{XA1} \neq \mu_{XA2}$.

Om de tweede hoofdvraag en de eerste daarbij behorende deelvraag te beantwoorden, is de groei van toetsresultaten van de klassen over vier toetsmomenten gemeten. De groei van de toetsresultatenresultaten voor de interventie is in kaart gebracht aan de hand van twee eerder behaalde resultaten in de twee groepen, die gebruikt zijn als voormeting (O₁ en O₂). Deze resultaten zijn voor de start van de onderzoeksperiode behaald in hetzelfde schooljaar, tussen week 43 van 2022 en week 5 van 2023. De voormeting is vergeleken met de toets die na de toetsreflectievragenlijst is afgenomen (O₄). Eerst is aan de hand van een *One-Way ANOVA* gemeten in welke mate de groepen voorafgaand aan de interventie gelijk zijn aan elkaar. Vervolgens is een *Repeated Measures ANCOVA* uitgevoerd om te meten of er een significante groei is tussen de voormetingen en de meting na de interventie. Met de *Repeated Measures ANCOVA* is vervolgens gemeten of het gevonden verschil ook over de tijd heen significant wordt verwacht. Met een *Post Hoc LSD* is bepaald welke verschillen er tussen de groepen zijn gevonden.

Verwacht wordt dat leerlingen uit R₁, die volgens de toetsreflectievragenlijst op de gemaakte toets reflecteren, na verloop van tijd betere resultaten zullen behalen op volgende toetsen. Leerlingen hebben dan een beter beeld verkregen van hoe ze beter kunnen leren en hoe antwoorden beter beantwoord moeten worden. Aan de hand van de controlegroep R₂ wordt

gemeten welke vooruitgang leerlingen zouden hebben zonder interventie op eenzelfde toets. In de nulhypothese wordt verwacht dat er in alle groepen eenzelfde gemiddelde groei wordt gevonden tussen de pretests (pr) en posttest (po) van de twee interventiegroepen: $H_0: \mu_{1pr} - \mu_{1po} = \mu_{2pr} - \mu_{2po}$. In de alternatieve hypothese wordt verwacht dat er wel een verschil wordt gevonden tussen de gemiddelde cijfers op de posttest ten opzichte van de pretest, en dat dit verschil in beide interventiegroep anders is: $H_a: \mu_{1pr} - \mu_{1po} \neq \mu_{2pr} - \mu_{2po}$.

Naast een verbeterd toetsresultaat wordt verwacht dat leerlingen in R_1 vaardiger worden in het beoordelen van het niveau van het eigen werk. Leerlingen die regelmatig reflecteren zullen op termijn beter kunnen inschatten wat de kwaliteit van hun werk is en welk cijfer hier bij past. Om dit te meten stelt de toetsreflectievragenlijst als laatste de vraag welk cijfer de leerling verwacht te behalen (*vc*). Het cijfer dat de leerling verwacht te halen, wordt vergeleken met het daadwerkelijk behaalde cijfer (*bc*). Bij de tweede meting vindt hetzelfde proces plaats. De verwachting is dat leerlingen in groep R_1 tijdens de tweede meting een betere schatting zullen maken dan tijdens de eerste meting. Dit is onderzocht aan de hand van een *Repeated Measures ANOVA*. De volgende hypothesen zijn opgesteld ter beantwoording van de desbetreffende deelvraag: $H_0: \mu_{vc1} - \mu_{bc1} = \mu_{vc2} - \mu_{bc2}$; $H_a: \mu_{vc1} - \mu_{bc1} \neq \mu_{vc2} - \mu_{bc2}$.

In de periode van dataverzameling zijn in alle klassen twee toetsen afgenomen. Er zijn daarom twee meetmomenten na de voortesten. Op deze manier wordt de voortest (O_1) niet alleen met de natest (O_4) vergeleken, maar er zijn ook twee tussenmetingen (O_2 ; O_3) die gebruikt kunnen worden om effect te meten. Het verloop van het onderzoek in de groepen is in onderstaand model weergegeven.

Figuur 3

Onderzoeksontwerp

NR_1	O_1	O_2	X	O_3	X	O_4
NR_2	O_1	O_2		O_3		O_4

Alle leerlingen uit de deelnemende klassen doen mee aan het onderzoek. Alleen als van leerlingen alle toetsresultaten en interventiemomenten bekend zijn, worden hun gegevens gebruikt in het onderzoek. Bij één of meer missende waarden worden de data van deze leerling niet in het onderzoek meegenomen. Omdat leerlingen worden geacht deel te nemen aan het aangeboden schoolprogramma en, om te bevorderen naar het volgende leerjaar, geen cijfers mogen missen wordt verwacht dat er weinig tot geen missende waarden worden gevonden.

Resultaten

Om de onderzoeksvragen te kunnen beantwoorden zijn de in de methode omschreven testen in SPSS uitgevoerd. Eerst is de betrouwbaarheid en validiteit van het gebruikte instrument, de toetsreflectievragenlijst, gemeten en samengevat in de volgende alinea. In de daarop volgende alinea's worden de resultaten per hoofdvraag per hoofdvraag weergegeven.

Instrument

De toetsreflectievragenlijst kent betrouwbaarheidsmaten: $\alpha = .503$; $\lambda_2 = .599$. Dit kan een bedreiging vormen voor de interne validiteit van het onderzoek. De betrouwbaarheid wordt iets hoger als de twee ratiovariabelen, vragen 7 en 10, uit de test worden gehaald: $\alpha = .583$; $\lambda_2 = .679$.

Daarnaast toont een factoranalyse van de vragenlijst aan dat de vragenlijst negen verschillende factoren meet. Deze factoren zijn veelal de deelvragen van een van de hoofdvragen samen, van welke er in totaal tien zijn.

Met Spearman's r_s correlatietest voor ordinale schalen is gemeten welke items significant correleren. In Tabel 2 is de correlatie tussen de variabelen *onderwijsniveau*, *leerjaar* en *vakgebied* en de items van de vragenlijst weergegeven, waarbij alleen items zijn weergegeven waar ten minste één significante correlatie is gevonden van $p < .01$. In Tabel 3 zijn de ordinale

items van de vragenlijst waar onderling ten minste één significante correlatie van $p < .001$ is gevonden weergegeven. De ratio-variabelen, vragen 7 en 10, zijn hierin niet meegenomen omdat de correlatie van deze niet met Spearman's r_s gemeten kunnen worden. De correlatie tussen deze twee ratiovariabelen is gemeten met Pearson's r : $n = 59$; $p = .616$; $r = -.070$.

Tabel 2

Correlatie Tussen Variabelen en Items van de toetsreflectievragenlijst

	onderwijsniveau	leerjaar	vakgebied
2 vragen gesteld	.296	-.432**	-.384*
4b leerdoelen in samenvatting	-.317	.384*	.273
5a oefentoets gemaakt	.303	-.343	-.256*
5b oefentoets nagekeken	.284	-.322*	-.319
5c oefentoets geanalyseerd	.376*	-.445**	-.440**
7 aantal dagen begonnen ^a	.317*	-.491**	-.452**
9a toetsvragen goed gelezen	.391*	-.250	-.364*
9c toetsvragen volledig beantwoord	.366*	-.351*	-.372*
10 schatting behaald cijfer ^a	-.313	.306	-.323*

^a Gemeten met Pearson's r * $p < .01$. ** $p < .001$.

Tabel 3*Beschrijvende Statistieken en Correlatie Tussen Items van de Toetsreflectievragenlijst*

Variabele	<i>N</i>	<i>M</i>	<i>SD</i>	1b	1c	3	4a	4b	5a	5b	5c	6d	8d	9a	9b	9c
1b taakwerk nagekeken ^c	57	3.42	1.281	-												
1c taakwerk verbeterd ^c	56	3.19	1.407	.725**	-											
3 samenvatting gemaakt ^c	61	2.95	.805	-.165	-.207	-										
4a begrippen in samenvatting ^b	61	2.80	1.077	.000	.023	.621**	-									
4b leerdoelen in samenvatting ^b	61	2.62	.934	.044	.075	.466**	.699**	-								
5a oefentoets gemaakt ^b	60	2.10	.896	-.069	-.021	-.329*	-.262	-.344	-							
5b oefentoets nagekeken ^b	60	2.03	.901	.014	.123	-.352*	.256	.348**	.850**	-						
5c oefentoets geanalyseerd ^b	61	2.03	.875	-.044	.086	-.236	-.191	-.250	.769**	.707**	-					
6d leerdoelen laten overhoren ^a	59	1.89	.504	.147	.004	.640**	.534**	.647**	-.412*	-.449**	-.226	-				
8d leerstof uitgelegd ^a	61	1.36	.659	-.032	.102	.142	.178	.159	-.406*	-.274	-.276	.057	-			
9a toetsvragen goed gelezen ^c	61	4.05	.669	.299	.287	.156	.247	-.024	.108	.221	.191	.084	-.68	-		
9b toetsvragen goed begrepen ^c	61	4.28	.669	.164	.221	.094	.064	-.105	-.043	.081	.082	-.170	.283	.414**	-	
9c volledig beantwoord ^c	61	4.37	.756	.203	.369*	.065	.011	-.097	.243	.339*	.271	-.107	-.051	.463**	.454**	-

^a Item met tweepuntsschaal^b Item met driepuntsschaal^c Item met vijfpuntsschaal* $p < .01$. ** $p < .001$.

Verschillen in leerstrategie

Onderwijsniveaus

De toetsreflectievragenlijst is gemaakt door leerlingen uit vier verschillende niveaus: vmbo-tl, vmbo-tl/havo, havo en vwo. Een zevental items uit de vragenlijst levert in de frequentieverdeling een verschillende aanpak van toetsvoorbereiding tussen leerlingen van verschillende leerniveaus op. Een aantal daarvan zijn significant bevonden aan de hand van een *Chi-Squares* test. De gevonden aantallen zijn in percentages weergegeven in Tabel 4, per *onderwijsniveau*. Op de rest van de items zijn over de leerniveaus redelijk gelijke uitkomsten gevonden. Deze zijn dan ook niet in Tabel 4 opgenomen.

Tabel 4

Verschillen in Gebruik Leerstrategieën tussen Onderwijsniveaus in Percentages

	1b kijkt taakwerk na (75% van het werk)*	3 maakt samenvatting (gedeeltelijk)	4b gebruikt leerdoelen in samenvatting (gedeeltelijk)**	5a maakt oefentoets (gedeeltelijk)	5b kijkt oefentoets na (gedeeltelijk)	5c gebruikt oefentoets voor vervolgstappen (gedeeltelijk)*	7 gemiddeld aantal dagen van tevoren begonnen
vmbo-tl	25,0 (50,0)	15,8 (57,9)	5,3 (68,4)	31,6 (10,5)	26,3 (21,1)	15,8 (31,6)	2,9
tl-havo	18,2 (18,2)	63,3 (27,3)	63,6 (18,6)	45,5 (18,2)	36,4 (18,2)	45,4 (9,1)	4,3
havo	11,1 (16,7)	26,3 (26,3)	5,3 (47,4)	50,0 (22,2)	55,6 (5,6)	36,8 (36,8)	3,9
vwo	33,3 (50,0)	25,0 (25,0)	8,3 (25,0)	58,3 (33,3)	50,0 (41,7)	75,0 (8,3)	5,3
totaal niveaus	21,1 (33,3)	32,8 (31,1)	16,4 (44,3)	45,0 (20,0)	41,7 (20,0)	39,3 (24,6)	4,0

* $p < .05$. ** $p < .01$.

Leerjaren

De toetsreflectievragenlijst is ingevuld door leerlingen van twee leerjaren: de eerste klas en de tweede klas. Zeven items uit de vragenlijst leveren in de frequentieverdeling een verschillende aanpak van toetsvoorbereiding tussen leerlingen van verschillende leerjaren op. Een aantal daarvan zijn significant bevonden aan de hand van een *Chi-Squares* test. De gevonden aantallen zijn in percentages weergegeven in Tabel 5, per *leerjaar*. Op de rest van de items zijn over de vergeleken leerniveaus vrijwel gelijke uitkomsten gevonden. Deze zijn dan ook niet in Tabel 5 opgenomen.

Tot slot schatten alle leerlingen in de eerste klas in dat ze voldoende op de toets hebben gescoord. Twee van hen scoorden uiteindelijk een onvoldoende. In de tweede klas zijn er 5 leerlingen die inschatten dat ze een onvoldoende hebben gehaald, terwijl er in totaal 15 onvoldoendes zijn gevallen.

Tabel 5

Verschillen in Gebruik van Leerstrategieën tussen Leerjaren in Percentages

	3 maakt samenvatting (gedeelte)ljk)*	4a gebruikt begrippen in samenvatting (gedeelte)ljk)*	4a gebruikt leerdoelen in samenvatting (gedeelte)ljk)*	5a maakt oefentoets (gedeelte)ljk)*	5c gebruikt oefentoets voor vervolgstappen (gedeelte)ljk)*	7 gemiddeld aantal dagen van tevoren begonnen ^a	8b verspreiden van leerstof over dagen*
leerjaar 1	22,7 (18,2)	31,8 (13,6)	9,1 (27,3)	61,9 (28,6)	68,2 (18,2)	5,6	90,9
leerjaar 2	33,3 (46,2)	33,3 (41,0)	20,5 (53,8)	35,9 (15,4)	23,1 (28,2)	3,1	64,1
totaal leerjaren	29,5 (36,1)	32,8 (31,1)	16,4 (44,3)	45,0 (20,0)	39,3 (24,6)	4,0	73,8

^a in aantal dagen

* $p < .05$. ** $p < .01$.

Vakken

De aanpak van leerlingen per vak is op een aantal items anders in de frequentieverdeling, waarvan een aantal significant is bevonden aan de hand van een *Chi-Squares* test. De gevonden aantallen zijn in percentages weergegeven in Tabel 6, per *vakgebied*. Op de rest van de items zijn over de vergeleken vakken vrijwel gelijke uitkomsten gevonden en zijn niet in Tabel 6 opgenomen.

Tabel 6

Verschillen in Gebruik van Leerstrategieën tussen Vakken in Percentages

	2b taakwerk nagekeken (gedeeltelijk)*	3 maakt samenvatting (gedeeltelijk)*	4b gebruikt leerdoelen in samenvatting (gedeeltelijk)*	5a maakt oefentoets (gedeeltelijk)*	6e leerdoelen check leerstof*	7 gemiddeld aantal dagen van tevorens begonnen ^a	8b verspreiden van leerstof over dagen*
wiskunde	27,7 (36,4)	22,7 (18,2)	9,1 (27,3)	61,9 (28,6)	90,9	5,6	90,9
geschiedenis	15,8 (15,8)	50,0 (35,0)	35,0 (40,0)	40,0 (20,0)	90,0	3,3	55,0
Nederlands	25,0 (50,0)	15,8 (57,9)	5,3 (68,4)	31,6 (10,5)	63,2	2,9	73,7
totaal vakken	21,1 (33,3)	29,5 (36,1)	16,4 (44,3)	45,0 (20,0)	82,0	3,9	73,8

^a in aantal dagen

* $p < .05$. ** $p < .01$.

Algeheel beeld gebruikte leerstrategieën

In de toetsreflectievragenlijst zijn leerlingen bevraagd over vier leerstrategieën die wetenschappelijk bewezen effectief zijn voor het voorbereiden op een toets. In Tabel 7 is samengevat hoe vaak strategieën gebruikt zijn onder de leerlingen die de toetsreflectievragenlijst hebben ingevuld.

Tabel 7*Percentage Leerlingen die Evidence-Based Leerstrategie Gebruikt*

Leerstrategie	Percentage leerlingen
leerstof opdelen (<i>interleaved practice</i>)	45
leerstof over dagen verspreiden (<i>spacing practice</i>)	73,8
leerstof over dagen herhalen (<i>retrieval practice</i>)	65,6
leerstof uitleggen aan iemand (<i>teaching and explaining</i>)	29,5
op meer dan één manier leren (<i>interleaved practice</i>)	59

Effect van de toetsreflectievragenlijst***Effect van de toetsreflectievragenlijst op schoolresultaat***

Bij de uitvoering van een One-Way ANOVA is geen bewijs gevonden dat de interventiegroep ($M_{O1}=6.17$, $SD_{O1}=1.467$; $M_{O2}=5.65$; $SD_{O2}=1.470$) en de controlegroep ($M_{O1}=6.45$, $SD_{O1}=1.285$; $M_{O2}=5.63$; $SD_{O2}=1.349$) significant van elkaar verschillen op beide pretests ($F(147)=1.575$, $p=.211$). De *test of homogeneity of variances* toont aan dat aan de assumptie voor gelijke varianties is gedaan ($p=.119$). Ook is gecontroleerd of de toetsresultaten van O_1 , O_2 , O_3 en O_4 normaal zijn verdeeld aan de hand van een histogram en een Q-Q Plot per toets.

De uitvoering van een *One-Way ANCOVA met post-hoc LSD* toont aan dat er gelijke varianties zijn tussen de testen van de interventiegroep en de controlegroep ($p=.918$). Er is geen bewijs gevonden dat op O_4 beter is gescoord door de interventiegroep dan de controlegroep ($p=.187$), rekening houdend met de pretests. Ook is er geen significant verschil gevonden tussen O_1 en O_4 ($p=.282$). Wel is er een significant verschil gevonden tussen O_2 en O_4 ($p=<.001$).

Effect van de toetsreflectievragenlijst op de nauwkeurigheid van zelfbeoordeling

Doormiddel van een *Paired-Samples t-test* is de nauwkeurigheid van het schatten van het cijfer door de leerlingen berekend. Er is een significant verschil gevonden in de schatting van O₃, het eerste meetmoment, en het daadwerkelijk behaalde resultaat: $t(55)=2.261, p<.001$. Het verschil tussen de schatting van de eerste toets en het daadwerkelijk behaalde cijfer is: $M=.457, SD=1.513$. Er is ook een significant verschil gevonden in de schatting van O₄, het tweede meetmoment, en het daadwerkelijk behaalde resultaat: $t(60)=1.816, p=.01$. Het verschil tussen de schatting van de tweede toets en het daadwerkelijk behaalde cijfer is: $M=.300, SD=1.290$. In Tabel 8 zijn de beschrijvende statistieken weergegeven per variabele.

Tabel 8

Beschrijvende Statistieken Schatting en Daadwerkelijk Resultaat per Variabele

	Schatting O ₃		Resultaat O ₃		Schatting O ₄		Resultaat O ₄	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Niveau								
vmbo-tl	6.19	.712	5.81	1.276	6.19	.853	5.74	1.208
vmbo-tl/havo	6.64	.606	6.09	1.637	5.93	1.102	6.76	1.166
havo	6.33	.834	5.77	1.975	6.55	.734	6.32	1.209
vwo	6.49	.826	5.97	1.648	6.38	.902	6.18	1.300
Leerjaar								
1	6.82	.787	6.23	1.661	6.78	.799	6.23	1.324
2	6.29	.794	5.87	1.641	6.16	.888	6.16	1.299
Vakgebied								
wiskunde	6.82	.787	6.23	1.661	6.78	.799	6.23	1.324
geschiedenis	6.40	.882	5.91	2.029	6.12	.939	5.91	2.029
Nederlands	6.19	.712	5.84	1.12	6.19	.853	5.85	1.212
Totaal	6.49	.826	6.04	1.690	6.38	.902	6.08	1.261

Vershil in toetsvoorbereiding na toetsreflectievragenlijst

In een *Paired Samples t-test* zijn alle testitems van de eerste en de tweede ronde aan elkaar gekoppeld. Drie items tonen een significant ander beeld van de tweede toetsvoorbereiding ten opzichte van de eerste keer, waarvan de beschrijvende statistieken in Tabel 9 zijn weergegeven.

Tabel 9

Beschrijvende Statistieken Significant Veranderde Items

	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>r</i>	<i>Two-Sided p</i>
1b Taakwerk nagekeken	-.245	.806	-2.215	52	.752	.031
4a begrippen in samenvatting	.246	.786	2.360	56	.762	.022
6c leerdoelen beeldvorming toets	.246	.474	3.915	56	.379	<.001

Discussie

Verschillen in leerstrategie

Ter beantwoording van de eerst hoofdvraag wordt eerst antwoord gegeven op de drie bijbehorende deelvragen die samen antwoord geven op de gestelde hoofdvraag: *Welke verschillen zijn er in de keuze en in de aanpak van leerstrategieën tussen leerlingen van verschillende onderwijsniveaus, verschillende leerjaren en verschillende vakgebieden van het voortgezet onderwijs?*

Onderwijsniveau

Welke verschillen zijn er in de keuze en de aanpak van leerstrategieën tussen leerlingen van verschillende onderwijsniveaus?

Op zeven aspecten is een verschil in aanpak van het leren per onderwijsniveau gevonden. Leerlingen uit de klassen vmbo-tl en vwo geven aan dat zij het taakwerk het taakwerk voor minstens 75% nakijken. Dat is significant meer ($p < .05$) dan leerlingen uit vmbo-tl/havo en havo doen. Leerlingen van het vmbo verwerken de stof het liefst op een concrete manier, waarbij zij het liefst leerstof oefenen en herhalen (Van der Neut et al., 2005). Het taakwerk maken en nakijken sluit daar goed op aan. Leerlingen van het vwo geven daarentegen wel significant vaker ($p < .05$) aan dat zij de oefentoets gebruiken om vervolgstappen voor het leren te bepalen.

Verder geven leerlingen uit de klassen vmbo-tl en vmbo-tl/havo significant ($p < .001$) aan meer gebruik te maken van leerdoelen in de samenvatting die zij maken. Leerlingen van het vwo doen dit het minst.

Tot slot bereiden leerlingen van vmbo-tl zich het minst ver (2,9) van tevoren op de toets voor, waar leerlingen van het vwo aangeven de meeste dagen (5,3) van tevoren beginnen.

Leerjaar

Welke verschillen zijn er in de keuze en de aanpak van leerstrategieën tussen leerlingen van verschillende leerjaren?

Op zes items van de vragenlijst is een significant verschil ($p < .05$) gevonden tussen de aanpak van leerlingen in het eerste en tweede leerjaar.

In het tweede leerjaar wordt vaker een volledige samenvatting gemaakt ter voorbereiding op de toets ten opzichte van het eerste leerjaar. Van de leerlingen die een samenvatting maken, zijn er in de tweede klas aanzienlijk meer leerlingen die de begrippen verwerken in de samenvatting en leerdoelen gebruiken bij het maken van de samenvatting.

Daarentegen wordt in de eerste klas veel vaker een oefentoets gemaakt. Leerlingen uit de eerste klas gebruiken vervolgens de uitkomst van de oefentoets veel vaker om

vervolgstappen in het toetsvoorbereidingsproces te bepalen. Ook beginnen leerlingen uit de eerste klas eerder met leren en zij verspreiden de leerstof vaker over meerdere dagen.

Vakgebied

Welke verschillen zijn er in de keuze en de aanpak van leerstrategieën van leerlingen tussen verschillende vakgebieden?

De leerstrategie *spaced practice*, waarbij leerlingen de stof over meerdere dagen opdelen, is effectief bevonden voor het vak wiskunde. Als leerlingen dezelfde hoeveelheid wiskunde oefenen, maar dan verdeeld over meerdere momenten, behalen leerlingen een hoger cijfer op de toets (Karpicke et al., in press). In de steekproef wordt aangegeven dat leerlingen bij het vak wiskunde meer dagen van te voren beginnen met leren, hoewel dit verschil niet significant is gemeten, is het toch gemiddeld 2,3 dagen eerder dan voor het vak geschiedenis en gemiddeld 2,7 dagen eerder dan voor het vak Nederlands. Ook het koppelen van de betekenis van nieuwe informatie aan al bekende informatie is voor wiskunde een bewezen effectieve aanpak, waarbij leerlingen bijvoorbeeld te werk kunnen gaan door zichzelf de nieuwe stof uit te leggen (Donker et al., 2014). De strategie om de stof aan iemand anders uit te leggen wordt echter door 81,8% van de leerlingen niet gekozen in de toetsreflectievragenlijst voor het vak wiskunde. Wel wordt voor het vak wiskunde significant ($p < .05$) vaker aangegeven dat de oefentoets wordt gemaakt. De strategie *retrieval practice* lijkt voor het vak wiskunde vaak toegepast te worden.

De leerstrategie *teaching and explaining* is een geschikte strategie voor het vak geschiedenis, omdat leerlingen op die manier concepten moeten identificeren en deze informatie moeten organiseren om het over te brengen (Karpicke et al., in press). In de steekproef wordt niet met een significant verschil aangegeven dat deze strategie veel wordt gebruikt bij het vak geschiedenis. Sterker nog: dit is de minst gekozen strategie in de groep van geschiedenisleerlingen. De meest gekozen strategie is het herhalen van leerstof over meerdere

dagen. Daarnaast geven de meeste leerlingen van het vak geschiedenis aan dat zij een samenvatting maken voor het leren, en is er een significant verschil ($p < .05$) gevonden dat leerlingen voor het vak geschiedenis hier leerdoelen bij gebruiken. Wanneer leerlingen in de samenvattingen aan de hand van de leerdoelen concepten uitleggen en samenbrengen, passen zij wel een vorm van *teaching and explaining* toe in de samenvatting.

Algeheel beeld verschillen in leerstrategie

Welke verschillen zijn er in de keuze en in de aanpak van leerstrategieën tussen leerlingen van verschillende onderwijsniveaus, verschillende leerjaren en verschillende vakgebieden van het voortgezet onderwijs?

De resultaten van de toetsreflectievragenlijst hebben uitgewezen dat er binnen verschillende onderwijsniveaus verschillende leerstrategieën worden gebruikt. Leerlingen van het vmbo-tl geven aan bij voorkeur strategieën te gebruiken waarbij leerstof herhaald wordt. Dit sluit aan bij eerdere bevindingen over deze doelgroep (Van der Neut et al., 2005). Ook leerlingen van het vwo geven met een significant verschil ten opzichte van de andere groepen de voorkeur aan deze strategie. Dit is in eerder onderzoek nog niet aangetoond.

Daarnaast geven leerlingen in het eerste leerjaar significant vaker aan de oefentoets te maken bij het voorbereiden, waar leerlingen uit het tweede leerjaar aangeven samenvattingen te maken.

Voor het vak geschiedenis wordt significant vaker een samenvatting gemaakt, hoewel *teaching and explaining* als effectieve leerstrategie wordt beschouwd (Karpicke et al., in press). Voor het vak wiskunde wordt *retrieval practice* het vaakst aangegeven als gebruikte leerstrategie. Deze wordt in het onderwijsveld als effectieve strategie beschouwd voor dit vak, vooral in combinatie met *spaced practice* (Karpicke et al., in press). Ook *teaching and explaining* is een effectieve leerstrategie voor het vak wiskunde (Donker et al., 2014), maar deze wordt overwegend niet gekozen door leerlingen bij dit vak.

Effect van de toetsreflectievragenlijst

Ter beantwoording van de tweede hoofdvraag wordt eerst antwoord gegeven op de drie bijbehorende deelvragen die samen antwoord geven op de gestelde hoofdvraag: *In hoeverre heeft het gebruik van de toetsreflectievragenlijst een positief effect op het schoolresultaat en het zelfbeoordelingsvermogen van leerlingen in het voortgezet onderwijs?*

Effect van de toetsreflectievragenlijst op schoolresultaat

Wat is het effect van de toetsreflectievragenlijst op het schoolresultaat van leerlingen in het voortgezet onderwijs?

Er is bewijs gevonden dat op de tweede toets na de interventie beter is gescoord door de interventiegroep dan de controlegroep in vergelijking met de tweede pretest ($p < .001$), maar niet in vergelijking met de eerste pretest ($p = .282$). Op de tweede pretest is significant lager gescoord dan op de eerste pretest, waardoor het niet met zekerheid te zeggen is of het verbeterde toetsresultaat is te verklaren uit de interventie. De gevonden verschillen zijn niet op de variabelen gesplitst, waardoor er niets te zeggen is over de verschillen tussen *onderwijsniveaus*, *leerjaren* en *vakgebieden*.

Effect van de toetsreflectievragenlijst op de nauwkeurigheid van zelfbeoordeling

In hoeverre bevordert de toetsreflectievragenlijst de nauwkeurigheid van de beoordeling van het eigen werk en het eigen proces van leerlingen in het voortgezet onderwijs?

Op de eerste toets waarbij leerling een schatting hebben gemaakt is een significant ($p < .001$) verschil gevonden in de nauwkeurigheid van de schatting, waarbij de leerlingen zich met gemiddelde van 0,457 overschatten op het cijfer. Er is ook een significant verschil gevonden in de tweede schatting en het daadwerkelijk behaalde resultaat ($p = .01$), waarbij de leerlingen zich met een gemiddelde van 0,300 overschatten op het cijfer. Tijdens de tweede schatting hebben de leerlingen een nauwkeurigere schatting gemaakt van het cijfer dat ze verwachten te halen.

Verskil in toetsvoorbereiding na toetsreflectievragenlijst

In hoeverre passen leerlingen in het voortgezet onderwijs het leergedrag aan nadat zij hebben gereflecteerd op de aanpak van het leren aan de hand van de toetsreflectievragenlijst?

Drie items van de vragenlijst zijn op de tweede vragenlijst significant anders ingevuld dan op de eerste vragenlijst, hoewel de verschillen klein zijn. Leerlingen geven bij de tweede keer aan minder vaak het taakwerk te hebben nagekeken. Wel geven leerlingen aan dat zij vaker begrippen in de samenvatting hebben gebruikt en leerdoelen hebben gebruikt om zich een beeld te vormen van de toets.

Algeheel effect van de toetsreflectievragenlijst

In hoeverre heeft het gebruik van de toetsreflectievragenlijst een positief effect op het schoolresultaat en het zelfbeoordelingsvermogen van leerlingen in het voortgezet onderwijs?

Het schoolresultaat na de interventie is in vergelijking met een van de twee pretests significant verbeterd. Omdat dit niet voor beide pretests geldt, is het niet met zekerheid te zeggen of het gevonden verschil aan de interventie toe te schrijven is.

Na het gebruiken van de toetsreflectievragenlijst is het gedrag van leerlingen in kleine, maar significante, mate veranderd. Leerlingen schatten het cijfer op de toets met gemiddeld 0,157 punt nauwkeuriger. Daarnaast hebben leerlingen aangegeven voor de tweede toets een andere strategie te hebben gekozen: er wordt minder vaak taakwerk nagekeken, maar vaker gebruik gemaakt van begrippen in de samenvatting en leerdoelen om een beeld te krijgen van de toets.

Beperkingen

De betrouwbaarheidsmaten van de toetsreflectievragenlijst zijn niet hoog, maar wel hoger dan $\lambda_2 = .600$. Dit kan een bedreiging vormen voor de interne validiteit van het onderzoek. De vragen los meten wel de studie-aanpak en het gebruik van leerstrategieën van

leerlingen, maar zijn samen niet betrouwbaar genoeg om sterke uitspraken te kunnen doen over de toetsvoorbereiding van leerlingen.

De matige betrouwbaarheidsmaten kunnen het gevolg zijn van het kleine aantal participanten die de vragenlijst heeft gebruikt ($n = 70$). De populatie die onderzocht wordt, alle leerlingen uit het reguliere voortgezet onderwijs in leerjaar 1 en 2, bestaat in schooljaar 2021/2022 uit 377.854 leerlingen (Centraal Bureau voor de Statistiek, 2022). De Sample size calculator van Qualtrics berekent voor deze populatiegrootte een ideale sample size van 384 als je een betrouwbaarheidsinterval van 95% hanteert. De steekproef van dit onderzoek ($n = 149$) is kleiner en kan daarom een bedreiging voor de externe validiteit van het onderzoek zijn. De vragenlijst zou beter getest kunnen worden op betrouwbaarheid als de vragenlijst onder een grotere groep uitgezet zou worden, over een meer representatieve steekproef op verschillende scholen.

Daarnaast toont een factoranalyse van de vragenlijst aan dat de vragenlijst negen verschillende factoren meet, en zeven wanneer men de twee ratiovariabelen uit de vragenlijst haalt. Iedere hoofdvraag van de vragenlijst lijkt daarom een losse factor te meten wat ook een bedreiging voor de betrouwbaarheid van het instrument vormt. Vervolgonderzoek naar de losse factoren, waarbij er aan de hand van meer vragen onderzoek wordt gedaan, kan meer uitwijzen over de factoren van toetsvoorbereiding.

Suggesties voor vervolgonderzoek

In dit onderzoek zijn significante verschillen gevonden in schoolresultaat, schatting van het cijfer en aanpassing van leerstrategieën na de interventie. De vragenlijst die is gebruikt kent beperkingen en zou daarom in aangepaste vorm een duidelijker beeld kunnen schetsen van het effect van de vragenlijst op het zelfregulerend leren van leerlingen. De factoren die de vragenlijst meet dienen bij voorkeur op meer vragen gebaseerd te zijn, waarbij de antwoordmogelijkheden op tenminste een vijfpuntsschaal worden aangeboden (Hofmans &

Mairesse, 2007). Ook verdient het de aanbeveling om de vragenlijst onder een grotere en meer representatieve populatie uit te zetten. Op die manier kan er een beter beeld gevormd worden van de gebruikte leerstrategieën per *onderwijsniveau, leerjaar* en *vakgebied*.

Literatuur

- Baarda, B. (2017). *Basisboek methoden en technieken: kwantitatief praktijkgericht onderzoek op wetenschappelijke basis*. Groningen: Noordhoff Uitgevers.
- Baars, M. (2014). *Instructional Strategies for Improving Self-Monitoring of Learning to Solve problems*. Rotterdam: Erasmus Universiteit.
- Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31, 445-457. [https://doi.org/10.1016/S0883-0355\(99\)00014-2](https://doi.org/10.1016/S0883-0355(99)00014-2)
- Boekaerts, M., & Simons, P.R. (1995). *Leren en instructie*. Assen: Van Gorkum.
- Centraal Bureau voor de Statistiek. (2022, 22 april). *VO; leerlingen, onderwijssoort in detail, leerjaar 2003/'04-2021/'22* [dataset]. Geraadpleegd van <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80040ned/table?fromstatweb>
- Curriculum.nu. (2018). *Handreiking brede vaardigheden*. Amersfoort: SLO.
- De Boer, A. H., Donker-Bergstra, A. S., Kostons, D. D. N. M., Korpershoek, H., & Van der Werf, M. P. C. (2013). *Effective strategies for self-regulated learning: A meta-analysis*. Groningen: GION.
- Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, 21(4), 337-346. <https://doi.org/10.1016/j.lindif.2011.03.002>
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3, 231-264. <https://doi.org/10.1007/s11409-008-9029-x>
- Dignath, C., & Büttner, G. (2013). Assessing How Teachers Enhance Self-Regulated Learning: A Multiperspective Approach. *Journal of Cognitive Education and Psychology*, 12(3), 338-358. <https://doi.org/10.1891/1945-8959.12.3.338>

- Donker, A. S., De Boer, H., Kostons, D., Dignath-Van Ewijk, C. C., & Van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, 11, 1-26.
<https://doi.org/10.1016/j.edurev.2013.11.002>
- Feskens, R. C. W., Lek, K., Maassen, A. M., & Meelissen, M. R. M. (2020). *PISA-2018 De verdieping: Toetsgedrag van Nederlandse leerlingen in PISA-2018: Een vergelijking met omliggende landen*. Arnhem: Stichting Cito instituut voor Toetsontwikkeling.
- Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology*, 39(2), 75-85.
<https://doi.org/10.1016/j.cedpsych.2014.01.001>
- Gagné, R. M. (1985). *The conditions of learning*. New York: Holt, Rinehart & Winston.
- Hofmans, J., & Mairesse, O. (2007). Impact of the Number of Response Categories on Linearity and Self-Anchoring Scales. *Methodology - European Journal of Research Methods for the Behavioural and Social Sciences*, 3(4), 160-169.
<https://doi.org/10.1027/1614-2241.3.4.160>
- Hoogveld, B., Janssen-Noordman, A. & Van Merriënboer, J. J. G. (2017). *Innovatief onderwijs ontwerpen. De ontwerpprincipes van het 4CID-model*. Groningen: Noordhoff Uitgevers.
- Karpicke, J. D., & O'Day, G. M. (in press). Elements of Effective Learning. In M.J. Kahana & A.D. Wagner (Eds.), *Oxford Handbook of Human Memory, Volume II: applications*. Oxford: Oxford University Press
- Kicken, W. (2008). *Portfolio use in vocational education. Helping students to direct their learning*. Academische proefschrift. Heerlen: Open Universiteit Nederland.

- Kostons, D. D. N. M., Donker, A. S., & Opdenakker, M.-C. (2014). *Zelfgestuurd leren in de onderwijspraktijk. Een kennisbasis voor effectieve strategie-instructie*. Groningen: GION onderzoek/onderwijs.
- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77, 1121-1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Moos, D. C., & Ringdal, A. (2012). Self-regulated learning in the classroom: a literature review on the teacher's role. *Education Research International*, 1, 1-15. <https://doi.org/10.1155/2012/423284>
- Nunes, L. D., & Karpicke, J. D. (2015). Retrieval-based learning: Research at the interface between cognitive science and education. In R. A. Scott & S. M. Kosslyn (Eds.), *Emerging Trends in the Social and Behavioral Sciences* (pp. 1-16). New York, NY, United States: John Wiley & Sons, Inc.
- O'Day, G. M., & Karpicke, J. D. (2020). Comparing and combining retrieval practice and concept mapping. *Journal of Educational Psychology*. Advance online publication. <https://doi.org/10.1037/edu0000486>
- Onderwijsraad. (2016). *De volle breedte van onderwijskwaliteit. Van smal beoordelen naar breed verantwoorden*. Den Haag: Drukkerij Excelsior.
- Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00422>
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulation skills and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74-98. <https://doi.org/10.1016/j.edurev.2017.08.004>

- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In: M. Boekaerts, P. Pintrich, & M. Zeidner, *Handbook of Self-regulation* (pp. 452-502). San Diego, CA: Academic Press.
- Pintrich, P. R., & De Groot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.
- Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A Manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Michigan: National Centre for Research to Improve Postsecondary Teaching and Learning.
- Silver, N. (2013). Reflective Pedagogies and the Metacognitive Turn in College Teaching. In: M. Kaplan, N. Silver, D. LaVaque-Manty, & D. Meizlish, *Using Reflection and Metacognition to Improve Student Learning: Across the Disciplines, Across the Academy* (pp. 10-21). Sterling, VA, United States: Stylus Publishing.
- Schön, D. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. San Francisco/London: Jossey-Bass.
- Schön, D. (2016). *The reflective practitioner: How professionals think in action*. London/New York: Routledge.
- Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology*, 24(6), 837-848. <https://doi.org/10.1002/acp.1598>
- Ten Dam, G., & Vermunt, J. (2003). De leerling. In: N. Verloop & J. Lowyck (Eds.), *Onderwijskunde* (151-193). Groningen/Houten: Wolters-Noordhoff.
- Toepoel, V. (2008). *A closer look at web questionnaire design*. [Doctoral Thesis, Tilburg University]. CentER, Center for Economic Research.
- Toppino, T. C., & Gerbier, E. (2014). About practice: Repetition, spacing, and abstraction. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 60, pp. 113-189). San Diego, CA, United States: Elsevier Academic Press.

- Van der Neut, I., Teurlings, C., & Kools, Q. (2005). *Inspelen op leergedrag van vmbo-leerlingen*. Tilburg: IVA.
- Winne, P. H. (2018). Theorizing and researching levels of processing in self-regulated learning. *British Journal of Educational Psychology*, 88, 9-20.
<https://doi.org/10.1111/bjep.12173>
- Winne, P., & Perry, N. (2000). Measuring self-regulated learning. In: M. Boekaerts, P. Pintrich, & M. Zeidner, *Handbook of Self-regulation* (pp. 531-566). San Diego, CA: Academic Press.
- Wood, E., Pressley, M., & Winne, P. H. (1990). Elaborative interrogation effects on children's learning of factual content. *Journal of Educational Psychology*, 82(4), 741-748.
<https://doi.org/10.1037/0022-0663.82.4.741>
- Zimmerman, B. J. (2000). Attaining self-regulation: a social cognitive perspective. In: M. Boekaerts, P. Pintrich, & M. Zeidner, *Handbook of Self-Regulation* (pp. 13-40). San Diego, CA: Academic Press.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner. *Theory into Practice*, 41, 54-70.
https://doi.org/10.1207/s15430421tip4102_2

Bijlagen

Bijlage A: Leerkrachtinstructie deelname onderzoek

Beste deelnemend docent,

Fijn dat je een bijdrage wilt leveren aan het onderzoek naar de effectiviteit van toetsreflectievragen. In deze brief licht ik je kort in over het praktische verloop van het onderzoek.

Je geeft les aan twee of drie verschillende klassen voor hetzelfde vak in hetzelfde leerjaar. Iedere klas krijgt de komende 8 weken een andere onderzoeksconditie toegekend: 1) de normale situatie of 2) het invullen van een digitale reflectievragenlijst op It's Learning.

Jij verwerkt alle data gepseudonimiseerd: het is de bedoeling dat ik niet kan achterhalen welke leerling achter welk pseudoniem zit. Alleen jij hebt daartoe de sleutel. Op deze manier bevat de data geen persoonsgegevens, waardoor geen toestemming van ouders nodig is. De data wordt na het onderzoek geanalyseerd en verwerkt in een masterthesis. De gepubliceerde data is niet herleidbaar naar specifieke personen.

Deelname aan het onderzoek is vrijwillig. Je hebt het recht op ieder moment te stoppen met de deelname aan dit onderzoek. Je hoeft in dat geval niet uit te leggen waarom je niet meer deel wilt nemen. Dit zal geen negatieve gevolgen hebben. Je kunt altijd vragen stellen over het onderzoek: nu, tijdens het onderzoek en na afloop. Dat kan door mij aan te spreken of via onderstaande contactgegevens te bereiken.

Ik besef dat dit onderzoek inzet van je vraagt. Ik ben je dan ook erg dankbaar dat je mee wilt doen aan dit onderzoek. Mede dankzij jouw inzet, kunnen we meer inzicht krijgen in de manier waarop leerlingen zich voorbereiden op de toets, hoe zij hierop reflecteren en wat zij hier vervolgens mee doen. Hopelijk zijn de ingevulde vragenlijsten ook waardevol voor jouw directe lespraktijk.

Met vriendelijke groet,

Ruth van Uhm

06-42059216

r.van.uhm@student.rug.nl

Instructie dataregistratie toetsreflectie

Drie klassen, twee condities

Voor het onderzoek ga je in twee of drie verschillende klassen, twee verschillende condities toepassen. Van tevoren koppel je de klassen aan een van de drie conditie.

- 1) **Normale situatie:** je behandelt de toetsen en het nabespreken daarvan zoals gewoonlijk.
- 2) **Vragenlijst:** deze klas maakt zo snel mogelijk na de toets de digitale vragenlijst op It's Learning. Meer hoeft je met de leerlingen niet te bespreken, al is het niet erg als dit op natuurlijke wijze wel voorkomt.

Het invullen van de dataregistratie

- per klas één formulier
- leerlingen zijn gepseudonimiseerd weergegeven (bijvoorbeeld met nummers, cijfers, combinaties, kleuren, dieren, random namen etc.)
- toetsresultaten
 - graag de afnamedatum van de toets noteren
 - toets 1 en toets 2 betreffen resultaten die leerlingen hebben behaald voordat deze interventie van start ging. Het gaat om de meest recente resultaten
 - toets 3 en toets 4 zijn resultaten waarop een van de condities is gehanteerd. Mocht er ruimte zijn, zou toets 5 ook kunnen worden ingevuld. Dat is voor het onderzoek geen vereiste.

Algemene opmerkingen

- Onderzoeker neemt in één keer (bij voorkeur de meivakantie) het compleet ingevulde registratieformulier in ontvangst.
- Onderzoeker ontvangt ook in één keer de resultaten van de digitale vragenlijst, gepseudonimiseerd met dezelfde 'code'. Als het goed is, kunnen de resultaten in een Excel-bestand worden gedownload en kunnen namen makkelijk worden vervangen voor het gekoppelde pseudoniem.
- Alleen de lesgevende docent heeft beschikking over de zelfgekozen 'sleutel' om van leerling naar pseudoniem te vertalen.

Dateregistratie toetsreflectie

niveau:

vak:

conditie: normaal / vragenlijst

	datum →	voor het onderzoek		tijdens het onderzoek		
		toets 1	toets 2	toets 3	toets 4	toets 5
leerling	m/v ↓	toets 1	toets 2	toets 3	toets 4	toets 5
1	m	7,5	6,4	7,6	8,5	5,6

Dataregistratie toetsreflectie

niveau: vak: conditie: normaal / vragenlijst

	datum →	voor het onderzoek		tijdens het onderzoek		
		toets 1	toets 2	toets 3	toets 4	toets 5
leerling	m/v ↓					
1	m	7,5	6,4	7,6	8,5	5,6

Dateregistratie toetsreflectie

niveau:

vak:

conditie: normaal / vragenlijst

		voor het onderzoek		tijdens het onderzoek		
	datum →					
leerling	m/v ↓	toets 1	toets 2	toets 3	toets 4	toets 5
1	m	7,5	6,4	7,6	8,5	5,6

Sleutel leerling-pseudoniem

Op dit formulier kan je voor jezelf de pseudoniemen registreren die je aan je leerlingen hebt gekoppeld. Het staat je vrij dit op een andere manier aan te pakken.

leerling	pseudoniem	leerling	pseudoniem	leerling	pseudoniem
Pietje Bel	1				

Bijlage B: De toetsreflectievragenlijst

Toetsreflectie

In deze enquête blik je terug op jouw voorbereiding op de toets, en hoe het maken van de toets is gegaan.

Als je de enquête eerlijk invult, leer je zelf het meest!

1

De eerste 8 vragen gaan over jouw voorbereiding op de toets.

Ik heb al het taakwerk over de toetsstof volledig...

	niks	25%	de helft	75%	alles
...gemaakt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...nagekeken	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...verbeterd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2

Ik heb vragen over de leerstof gesteld aan de docent.

	helemaal oneens	oneens	neutraal	eens	helemaal eens
Ik heb vragen over de leerstof gesteld aan de docent.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3

Ik heb een samenvatting gemaakt over de toetsstof.

	helemaal oneens	oneens	neutraal	eens	helemaal eens
Ik heb een samenvatting gemaakt over de toetsstof.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4

In mijn samenvatting...

	nee	gedeeltelijk	allemaal	ik heb geen samenvatting gemaakt
... heb ik de begrippen uitgelegd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... heb ik de leerdoelen gebruikt als leidraad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5

Ik heb de oefentoets...

- | | nee | gedeeltelijk | ja |
|---|-----------------------|-----------------------|-----------------------|
| ... gemaakt | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| ... nagekeken | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| ... gebruikt om te zien waar ik nog meer voor moest leren/oefenen | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

6

Ik heb de leerdoelen gebruikt...

- | | nee | ja |
|---|-----------------------|-----------------------|
| ... om mezelf te overhoren | <input type="radio"/> | <input type="radio"/> |
| ... om me te laten overhoren | <input type="radio"/> | <input type="radio"/> |
| ... om een beeld te krijgen van wat je op de toets kan verwachten | <input type="radio"/> | <input type="radio"/> |
| ... in mijn samenvatting | <input type="radio"/> | <input type="radio"/> |
| ... om te checken of ik alle stof heb doorgenomen | <input type="radio"/> | <input type="radio"/> |

7

Hoeveel dagen van te voren ben jij begonnen met de voorbereiding op de toets?

8

Ik heb de leerstof...

- | | ja | nee |
|---|-----------------------|-----------------------|
| ... in (korte) stukjes opgedeeld voor het leren | <input type="radio"/> | <input type="radio"/> |
| ... over meerdere dagen verspreid | <input type="radio"/> | <input type="radio"/> |
| ... op verschillende dagen herhaald | <input type="radio"/> | <input type="radio"/> |
| ... uitgelegd aan iemand anders om extra te oefenen | <input type="radio"/> | <input type="radio"/> |
| ... op meer dan één manier uit mijn hoofd geleerd | <input type="radio"/> | <input type="radio"/> |

9

Deze vraag gaat over het maken van de toets.

Tijdens het maken van de toets heb ik...

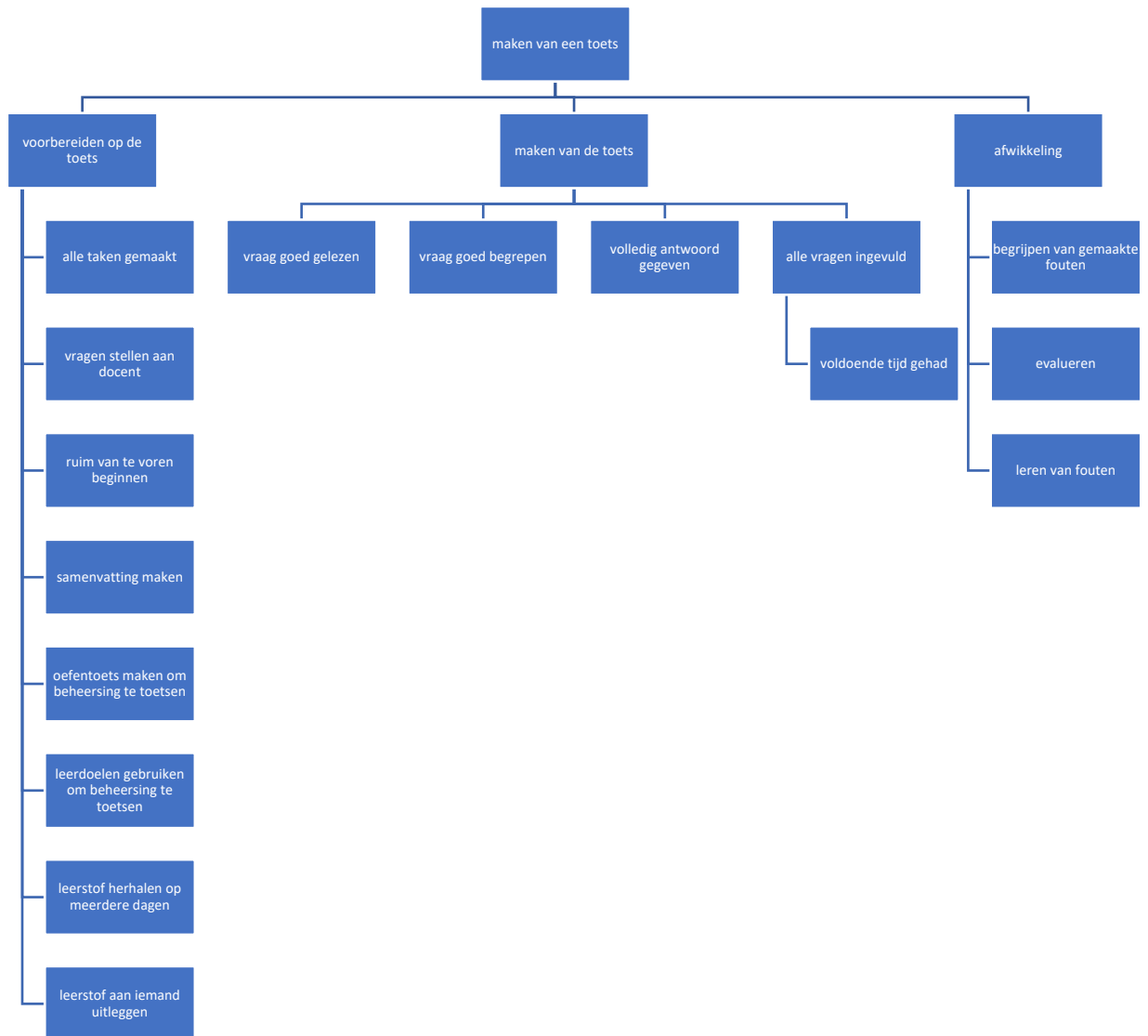
	helemaal oneens	oneens	neutraal	eens	helemaal eens
... de vragen op de toets goed gelezen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... de vragen op de toets goed begrepen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... de vragen op de toets volledig beantwoord	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... op alle vragen een antwoord ingevuld	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... voldoende tijd gehad om te toets te maken	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10

Ik denk dat ik het volgende cijfer heb gehaald:

Bijlage C: Totstandkoming toetsreflectievragenlijst

1) Vaardighedenhiërarchie



2) Prestatiedoelen in scoringsrubriek

	Te beoordelen samenstellende vaardigheid/aspecten	Prestatiedoelen	Beoordelingsschaal
voorbereiding op de toets	alle taken gemaakt	De student heeft het taakwerk over de toetsstof is volledig: 1a) gemaakt; 1b) nagekeken; 1c) verbeterd.	5-puntsschaal: 1 = niks 5 = alles
	vragen stellen aan docent	2) Indien nodig heeft de student vragen over de toetsstof gesteld aan de docent.	5-puntsschaal: 1 = helemaal oneens 5 = helemaal eens
	samenvattingen maken	3) De student heeft de toetsstof verwerkt in een samenvatting	5-puntsschaal: 1 = helemaal oneens 5 = helemaal eens
	samenvattingen maken	De student heeft in de samenvatting: (controlevraag op vraag 3) 4a) de begrippen uitgelegd; 4b) de leerdoelen gebruikt als leidraad.	3-puntsschaal: 1 = nee 2 = gedeeltelijk 3 = ja 4 = geen samenvatting gemaakt
	oefentoets maken om beheersing te toetsen	De student heeft de oefentoets: 5a) gemaakt; 5b) nagekeken 5c) hier conclusies uit getrokken over de stof die extra aandacht verdient.	3-puntsschaal: 1 = nee 2 = gedeeltelijk 3 = ja
	leerdoelen gebruiken om beheersing te toetsen	De student heeft de leerdoelen gebruikt: 6a) om zichzelf te overhoren; 6b) om zich te laten overhoren; 6c) om een beeld te krijgen van de toets; 6d) in de samenvatting (controlevraag op vraag 4b) 6e) om te checken of alle stof is doorgenomen	2-puntsschaal: 1 = nee 2 = ja

	leerstof herhalen op meerdere dagen	7) De student is ... aantal dagen van tevoren begonnen met leren.	open vraag
	leerstrategieën	De student heeft de leerstof... 8a) in (korte) stukjes opgedeeld voor het leren (<i>spacing practice</i>) 8b) over meerdere dagen verspreid (<i>spacing practice</i>) 8c) op verschillende dagen herhaald (<i>retrieval practice</i>) 8d) uitgelegd aan iemand anders (<i>teaching and explaining</i>) 8e) op meer dan één manier uit het hoofd geleerd (<i>interleaved practice</i>)	2-puntsschaal: 1 = nee 2 = ja
maken van de toets	vragen goed gelezen	9a) De student heeft de vragen goed gelezen.	5-puntschaal: 1 = helemaal oneens; 5 = helemaal eens
	vragen goed begrepen	9b) De student heeft de vragen goed begrepen.	5-puntschaal: 1 = helemaal oneens; 5 = helemaal eens
	volledig antwoord gegeven	9c) De student heeft de vragen volledig beantwoord.	5-puntschaal: 1 = helemaal oneens; 5 = helemaal eens
	alle vragen ingevuld	9d) De student heeft alle vragen ingevuld.	5-puntschaal: 1 = helemaal oneens; 5 = helemaal eens
	voldoende tijd gehad	9e) De student heeft voldoende tijd gehad.	5-puntschaal: 1 = helemaal oneens; 5 = helemaal eens
	cijfer ingeschat	10) De student schat in dat hij het volgende cijfer krijgt,	open vraag; 10-puntschaal