Geografisch Besef in aardrijkskundig onderwijs

Een valideringsonderzoek naar de Gea-toets

Marchje van der Veen 3219364

Geo-communicatie

Masterthesis mei 2012

Begeleider Tine Béneker (UU) en Henk Notté (Cito Instituut)

Abstract [in English, thesis is written in Dutch]

Henk Notté, geographer and test constructor at Cito Institute, has developed a test (the gea-test) in order to measure geographical awareness amongst scholars aged 10 to 18 years old. The test is expected to distinguish scholars with a strongly developed geographical awareness from those with a less developed geographical awareness and to translate these outcomes into a score.

Being a part of its validity research this thesis revolves around the question: 'on what type of geographical knowledge scholars base their answers, when answering questions of the gea-test and what does that imply for the validity of the test? '. This research takes a closer look to the reasoning behind a good or wrong answered question. Why can a scholar answer the question correctly? Do certain misconceptions exist within geographical education?

The research method used in this thesis are 'thinking-aloud'-protocols, a research method often perceived as being extremely time-consuming and therefore rarely used. Because this method has never been applied on larger scales this research offers a unique insight in what goes on in the scholar's mind when asked to answer a geographical question.

In the first part of the research scholars were asked to make the gea-test consisting out of 35 randomly selected questions. Afterwards, scholars with either exceptionally high or low scores or scores equal to the class average were selected for the 'thinking-aloud'-protocols, a total of 24 scholars have participated in the protocols. During the 'thinking-aloud'-protocols scholars were asked to make each six preliminary selected exercises while thinking and reasoning aloud. This allows the researcher to clarify why a pupil has chosen a certain answer. The concept of geographical awareness comprises out of nine different skills and abilities: photo-interpretation, map and graph reading, scales, comparisons, factual knowledge, thematic mental map, topographical mental map, conceptual knowledge and principal knowledge. These skills were all evaluated and quantified during the 'thinking-aloud-protocols' and made a comparison possible between the outcomes of the gea-test and the 'thinking-aloud'-protocols could have been made by computer program TiaPlus.

Based on the results we can say that when answering questions of the gea-test scholars aged 11 to 12 mainly use the skills: scale, topographical mental map and conceptual knowledge while scholars aged 14 to 15 mostly have used the skills: thematic mental map, topographical mental map and conceptual knowledge and principal knowledge.

These results have led to the conclusion that the gea-test does make a rightful distinction between scholars with a strongly and scholars with a lesser developed geographical awareness.

This conclusion indicates that the gea-test is a valid method to evaluate the scholars capability regarding geographical skills such as the use of scales, topographical and thematic mental maps, conceptual knowledge and principal knowledge.

Indeling thesis

- Inhoudsopgave

1 Inleiding…… 4
1.1 Introductie
1.2 Doelstelling
1.3 Relevantie
1.4 Aanpak
1.5 Leeswijzer

2 Gea-toets……7
2.1 Tot stand koming gea-toets
2.3 Geografische besef in de gea-toets

3 Toetsconstructie, validiteit en betrouwbaarheid……20
3.1 Toets constructie
3.2 Toetsontwerp
3.3 Validiteit en betrouwbaarheid

4 Onderzoeksmethode en methodologie……25
4.1 Motivatie onderzoeksmethode en plaats van dit valideringsonderzoek binnen het valideringsproces van de gea-toets
4.2 Onderzoeksopzet
4.3 Conceptueel model en hypothese
4.4 Operationalisering van geografisch besef en 'hard-op-denk'-protocollen
Motivatie keuze voor vraagstukken die aan kinderen worden voorgelegd
4.5 Motivatie keuze voor scholen en groepen leerlingen

5 Resultaten……33
5.1 Resultaten uit kwantitatief onderzoek
5.2 Resultaten uit kwalitatief onderzoek
5.3 Analyse resultaten van de 'hard-op-denk'-protocollen

6 Conclusie……58
6.1 Beantwoording hoofdvraag
6.2 Beperkingen van het onderzoek
6.3 Suggesties toekomstig onderzoek

- Nawoord……63

- Bronnenlijst……64

- Bijlagen ……65

1 Inleiding

1.1 Introductie

Wie herinnert zich niet de entreetoets, de eindtoets in groep 8 of de eindexamens? De toets productie van het Cito loopt als een rode draad door de schoolcarrières van menig leerling. Deze toetsen worden echter afgenomen op vaststaande momenten (zoals in februari groep 8 bij de Cito-eindtoets) en kunnen daardoor niet gebruikt worden door docenten om het kennisniveau van hun leerlingen in de loop der jaren bij te houden. Het vak aardrijkskunde kent twee belangrijke toetsmomenten; als een onderdeel in 'wereldoriëntatie' bij de eindtoets in groep 8 en bij het eindexamen aardrijkskunde. Tot nu toe was het zeer lastig tussen deze toetsmomenten de geografische kennis van leerlingen langs een objectieve, landelijke, meetlat te leggen. Aardrijkskundedocenten stellen zelf hun toetsen samen en kunnen zodoende weinig vergelijken met andere klassen of scholen. Henk Notté, geograaf en toetsdeskundige van het Cito, ontwikkelde een toets (Gea-toets) om geografisch besef te kunnen meten. Deze toets zou te allen tijde gebruikt kunnen worden door docenten om het niveau van hun leerlingen van groep 6 lagere school tot en met 5/6 havo/vwo te testen (meer informatie over Gea-toets staat vermeld op: https://sites.google.com/site/geatoets/).

Toen ik het Cito benaderde voor een stage waaraan ik mijn scriptie aan kon verbinden werd mij voorgesteld een valideringsonderzoek te doen naar de werking van de Gea-toets. Op dit verzoek wilde ik graag ingaan vanwege de interessante manier van onderzoeken, namelijk door middel van 'hard-op-denk'-protocollen, waarbij aan de leerlingen gevraagd wordt hardop te redeneren bij het maken van een opgave. Op deze wijze kan achterhaald worden op basis van welke argumenten leerlingen een keuze maken bij het oplossen van aardrijkskundige vraagstukken. Aangezien deze onderzoeksopzet vaak als zeer tijdrovend wordt beschouwd, wordt deze methode weinig toegepast waardoor er ook niet veel bekend is over deze manier van onderzoeken. Kortom: Dit onderzoek biedt dus een uniek inzicht in de aardrijkskundige kennis en vaardigheden van scholieren op de basisschool en middelbare school.

1.2 Doelstelling

Op dit moment is de Gea-toets (huidige versie 4) bijna volledig klaar, maar voordat zij in gebruik kan worden genomen moet men eerst zeker weten of de niveaus die de toets hanteert goed zijn ingeschat. Dit zal op een aantal manieren onderzocht worden en dit onderzoek maakt daar een onderdeel van uit. Dit onderzoek geldt als een eerste try-out waarna een grootschalig proefonderzoek met 3000 leerlingen zal volgen.

Daarnaast moet onderzocht worden of de Gea-toets meet wat zij hoort te meten, oftewel de validiteit van de toets. De toets wordt geacht een onderscheid te maken tussen leerlingen met een sterk geografisch besef en leerlingen met een zwak geografisch besef en hieraan een score te verbinden. Het is belangrijk dat wordt nagegaan of de leerlingen de denkstappen zetten die door de toetsconstructoren in de vraagstelling is uitgezet. Beredeneren de leerlingen inderdaad het antwoord op de manier waarop dat van hen verwacht wordt? Met welke soort geografische benaderingswijze (feitenkennis, kaartlezen etc.) komen zij op hun antwoord?

Beide vraagstukken vormen de kern van dit onderzoek.
Dit leidt tot de volgende hoofdvraag:

Op welke soort geografische kennis doen leerlingen beroep bij het beantwoorden van de vragen van de Gea-toets en wat zegt dit over de validiteit van de toets?

Deze vraag kan worden opgesplitst in twee te onderzoeken deelvragen. Namelijk: ‘Op welke soort kennis doen leerlingen beroep bij het beantwoorden van de vragen?’ en 'wat zegt dit over de validiteit van de toets?'.

Om deze hoofdvraag te kunnen beantwoorden is het onderzoek opgesplitst in twee delen, een kwantitatief en een kwalitatief gedeelte.

Het kwantitatieve gedeelte van het onderzoek:
Om het niveau van de leerlingen vast te stellen zullen aan een aantal klassen van het leerjaar groep 8 van de lagere school en derde klas middelbare school worden gevraagd de Gea-toets op de computer af te nemen. De antwoorden zullen kwantitatief worden geanalyseerd en zodoende zal worden vastgesteld of de leerling een lage, gemiddelde of hoge score heeft behaald. Op basis van deze uitslag worden later leerlingen uitgekozen om verdere opgaven af te laten nemen. Het aantal leerlingen is niet representatief om te kunnen beoordelen of de toets betrouwbaar is (dit zal later bij uit proeftoetsing blijken), maar kan wel een indicatie geven of de niveaus die bij de Gea-toets gebruikt worden corresponderen met de leerjaren waarop de indeling is gebaseerd.

Het kwalitatieve gedeelte van het onderzoek:
De hoofdvraag zal uiteindelijk worden beantwoord door middel van een kwalitatieve onderzoeksmethode. Aan de hand van de scores die uit het kwantitatieve onderzoek naar voren zullen komen, zullen leerlingen van uiteenlopende niveaus (een aantal met hoge scores, lage scores en gemiddelde scores) later worden benaderd. Samen zullen we een zestal opgaven maken. Bij deze opgaven wordt aan de leerling gevraagd hardop te denken (deze manier van onderzoeken wordt dan ook een ‘hard-op-denk’-protocol genoemd), zodoende kan worden nagegaan welke afwegingen een leerling maakt bij het beantwoorden van een vraag. Deze gesprekken met de leerlingen zullen verlopen volgens een vast protocol, van te voren wordt in overleg met Henk Notté besproken welke opgaven een goede afspiegeling zijn van de verschillende vaardigheden die samen het geografisch besef vormen en naar welke begrippen verder zal worden doorgevraagd. De resultaten van dit deelonderzoek zullen worden geanalyseerd om zodoende te achterhalen waarom een leerling kiest voor een bepaald antwoord. Kan de leerling het juiste antwoord geven met feitenkennis (bijvoorbeeld topografie)? Of door gebruik te maken van geografische vaardigheden zoals kaartlezen en gebruik van windrichtingen? Kiest een twijfelende leerling toch steeds het juiste antwoord gebaseerd op ‘een gevoel’?

De resultaten uit het eerste en tweede deelonderzoek leveren interessante informatie op waarmee tijdens de verdere ontwikkeling van de Gea-toets rekening gehouden kan worden. Daarnaast verkrijgt men bij het tweede deelonderzoek inzicht in het keuzeproces van een leerling bij het maken van een geografische toets, iets waar nog weinig onderzoek naar is gedaan.

1.3 Relevantie

Het Cito heeft al veel onderzoek gedaan naar de resultaten van verschillende toetsen, maar deze onderzoeken hebben zich voornamelijk geconcentreerd op de inhoudelijke geografische kennis bij leerlingen. Zoals het percentage leerlingen die het juiste antwoord kunnen geven bij een bepaald vraagstuk. Dit onderzoek gaat dieper in op de achterliggende gedachtegang achter een goed- of fout gegeven antwoord. Waarom kan een leerling een vraag goed beantwoorden? Kan hij dit beredeneren vanuit een geografisch besef, een context waarin hij het vraagstuk plaatst? Of is de leerling dermate bekend met de leerstof dat het vraagstuk met behulp van algemene parate feitenkennis kan worden opgelost? Bestaan er bepaalde misconcepties? En in het geval van een fout gegeven antwoord, waarom hebben de leerlingen voor dit antwoord gekozen? Deze vragen zijn in het belang van de realisatie van de Gea-toets aangezien de toets verschillende vormen van geografische kennis bij leerlingen wil aanspreken. Uit gesprekken met toets deskundigen van het Cito bleek dat er tot nu toe heel weinig onderzoek gedaan is naar de beredenatie van leerlingen om voor

een bepaald antwoord te kiezen. Dit komt omdat deze manier van onderzoeken, met behulp van hardop-denk-protocollen, als erg tijdrovend en intensief wordt beschouwd. Hierdoor bestaat er ook weinig wetenschappelijke literatuur over deze manier van onderzoeken. Aangezien er wel behoefte bestaat aan onderzoek naar keuzeprocessen van leerlingen bij het maken van toetsen, maakt dat dit onderzoek van grote wetenschappelijke relevantie is.

Naast een wetenschappelijke relevantie heeft dit onderzoek ook een maatschappelijke relevantie. Dit onderzoek is namelijk een stap in het gehele valideringsproces van de Gea-toets. Een algemene toets voor geografisch besef kan een belangrijke bijdrage leveren aan de verbetering van het aardrijkskundeonderwijs, omdat ze als een spiegel dient voor leerlingen, leerkrachten, ontwikkelaars en beleidsmakers. Er zijn voor de schoolaardrijkskunde allerlei toetsen ontwikkeld om de leerresultaten te meten, maar iedere toets heeft haar eigen beperkingen. Ze is methodeafhankelijk; ze is niet of gebrekkig gekoppeld aan referentieniveaus; de validiteit en betrouwbaarheid zijn niet of onvoldoende vastgesteld; ze geeft onvoldoende inzicht in het bereikte niveau of de positie op een leerlijn voor aardrijkskunde of ze is niet algemeen toegankelijk (Universiteit Twente research center voor examinering en certificering, 2011). Hierdoor is het aannemelijk dat de samenleving gebaat zou zijn bij de ontwikkeling van een algemene onafhankelijke toets waardoor zowel voor de leerlingen als voor de leerkrachten een reëel beeld wordt geschetst van hoe hun geografische kennis ervoor staat.

1.4 Aanpak

Bij het beantwoorden van de hoofdvraag wordt er eerst een literatuuronderzoek verricht. Hierbij wordt er een overzicht gegeven van alle informatie die al beschikbaar is over toetsconstructie, validiteit en vormen van kennis en vaardigheden die kunnen worden gerekend tot geografisch besef. Deze informatie is afkomstig uit bronnen waarvan een aantal door het Cito beschikbaar zijn gesteld. Aan de hand van deze informatie worden de vragen voor het eigen kwantitatieve/kwalitatieve onderzoek samengesteld. Bij dit onderzoek worden leerlingen gevraagd hardop te redeneren bij het beantwoorden van een geografisch vraagstuk. Op deze manier wordt onderzocht of de leerlingen op het juiste antwoord komen via de denkstappen die de vraagmakers voor ogen hadden. In de conclusie wordt de hoofdvraag beantwoord met behulp van de informatie uit het theoretische kader en de uitkomsten van de interviews.

1.5 Leeswijzer

Deze thesis bestaat uit twee delen, namelijk uit een overzicht van bestaande informatie over toetsconstructie en het eigen onderzoek. Hoofdstuk 2 geeft een verdere toelichting op de ontwikkeling van de gea-toets en op welke wijze de toets aanspraak doet op het geografisch besef van leerlingen. In hoofdstuk 3 wordt informatie weergegeven wat bestaat uit resultaten uit eerdere onderzoeken die met toetsconstructies, validiteit of betrouwbaarheid te maken hebben. Hoofdstuk 4 geeft een beschrijving van de gebruikte methoden bij het onderzoek en hoe de informatie uit het theoretisch kader is gebruikt om het kwantitatieve en kwalitatieve onderzoek op te zetten. Daarnaast wordt hier aan de hand van het toetsingsmodel aangegeven welke plaats dit onderzoek inneemt in de gehele validering van de gea-toets en hoe het ‘hard-op-denk’-protocol werkt. Verder zal worden uitgelegd op welke manier de hypothese en het conceptueel model van het onderzoek zijn geconstrueerd. In hoofdstuk 5 volgen de resultaten uit het onderzoek waarbij de resultaten uit het kwantitatieve en het kwalitatieve gedeelte apart behandeld zullen worden. Hoofdstuk 6 betreft de conclusie waarin we de hoofdvraag beantwoorden met behulp van de informatie uit het theoretische kader en de resultaten uit het onderzoek. Hierop volgt het nawoord, de literatuurlijst en de bijlagen waarin de resultaten en het ‘hard-op-denk’-protocol terug te vinden zijn.

2 Gea-toets

In dit hoofdstuk wordt de gea-toets geïntroduceerd en toegelicht. Eerst zal in paragraaf 2.1 de nadruk worden gelegd op de inhoud van de toets. In paragraaf 2.2 zullen verschillende vormen van geografische vaardigheden en soorten kennis aan bod komen die samen als uitgangspunt dienen in deze scriptie bij het beoordelen van geografisch besef.

2.1 De gea-toets

Het idee voor de gea-toets is ontstaan bij de tentoonstelling 'Waar ben ik?' in het Onderwijsmuseum in Rotterdam waar een plek was ingericht waar men zijn geografische kennis kon testen. Voor dat laatste werd een interactieve toets ontwikkeld die je in tien minuten achter een computer in het museum kon maken. De toets kende vijf niveaus, want het publiek zou qua leeftijd en opleidingsniveau flink variëren. Ruim 2000 bezoekers hebben deelgenomen aan deze toets die was opgezet als een reis rond de wereld in negentien stappen van de fictieve reisgenoot Gea. Tijdens de virtuele rondreis kregen deelnemers vragen over het gebied waar ze doorheen reisden. Om de vragen goed te beantwoorden was geografische kennis nodig: praktische kennis over topografie, mensen en landschappen. Aan het eind van de toets kregen de reizigers meteen te horen hoe het stond met hun kennisniveau (Notte, H. & Schee, J. van der, 2008).

Aangezien de opgaven niet te gemakkelijk of te moeilijk mochten zijn, hield het programma rekening met het antwoord op de laatste vraag voordat een nieuwe vraag geselecteerd werd. Bij een fout antwoord belandde je op een lager niveau; gaf je het goede antwoord, dan werd de volgende vraag moeilijker. De inhoud werd specifieker, de context complexer, de nodige denkvaardigheid van een hogere orde, of een combinatie daarvan. Hierdoor werd de toets geschikt voor zowel volwassen als scholieren uit het basis- of middelbaar onderwijs.
Voor iedere etappe waren meerdere vragen beschikbaar. Het programma maakte daaruit een willekeurige keuze zodat je het spel nog een keer kon doen zonder dat je dezelfde vragen voorgeschoteld kreeg. De test maakte nieuwsgierig naar een vervolg aangezien er geen instrument bestond om de algemene basiskennis voor aardrijkskunde op verschillende niveaus te meten. Het uiteindelijke doel werd een gestandaardiseerde toets met duidelijk omschreven niveaus van aardrijkskundige basiskennis. Deze wens heeft geleid tot de ontwikkeling van de huidige gea-toets (Notte, H. & Schee, J. van der, 2008) .

Bij het maken van de gea-toets gaat de leerling denkbeeldig op wereldreis. Elke vraag gaat over een andere bestemming die een aantal graden verder naar het Oosten ligt totdat de leerling weer terug is in Nederland. De leerling wordt tijdens de gea-toets 32 vragen voorgelegd (uit een totale voorraad van 320 vragen). De gea-toets is een adaptieve toets, dit houdt in dat het niveau van de toets is afgestemd op de leerling, hierdoor blijft een leerling van niveau wisselen totdat deze op het juiste niveau zit. In totaal zijn er vijf verschillende niveaus waarvan niveau 1 correspondeert met het niveau van een groep 6 leerling, niveau 2 overeenkomt met een groep 8 leerling, niveau 3 met een leerling in de derde klas van het voortgezet onderwijs, niveau 4 op het niveau van een leerling in 5/6 havo/vwo en niveau 5 vanaf 5/6 vwo. De toets begint met een vraag op niveau 3 wat overeenkomt met de aardrijkskundige kennis van een leerling uit de derde klas van het voortgezet onderwijs. Bij het goed beantwoorden van een vraag gaat de leerling een niveau hoger (niveau 4), is de vraag verkeerd beantwoord dan gaat de leerling een niveau naar beneden (niveau 2). Uiteindelijk geeft het behaalde niveau en het aantal punten (die verkrijgbaar zijn voor elk goed antwoord) weer op welk niveau de leerling zit.

In de gea-toets komen verschillende vragen aan bod, dit zijn vragen over kaartlezen, topografie, landschap en bevolking. Bij vragen over kaartlezen dient een leerling bepaalde informatie uit een topografische of thematische kaart te halen en bij een topografische vraag moet de leerling een plaats of gebied op de kaart aanwijzen. Tot de categorie 'landschap' behoren de vragen die in het algemeen tot de fysische geografie worden gerekend zoals vragen over de natuurlijke omgeving,

vulkanen, oases, bergen, (land)ijs en klimaten. In de categorie 'bevolking' vallen de vragen die betrekking hebben op de sociale geografie zoals over welvaart, bevolkingsgroei, cultuur en de spreiding van talen en religies. Het is dan ook vanzelfsprekend dat deze categorieën een ongelijke spreiding kennen over de verschillende niveaus aangezien van leerlingen in groep 6 nog geen of weinig inzicht kan worden verwacht in meer complexere vraagstukken zoals de wens naar onafhankelijkheid van Tibetanen of voormalig koloniale relaties.
Daarom komen vragen in de categorie bevolking op niveau 1 niet voor. Andersom komen vragen in de categorieën kaartlezen en topografie op niveau 4 en 5 niet meer voor (en op niveau 3 nog maar zelden) aangezien dit niet aansluit bij de complexere verbanden waar op dit niveau naar gevraagd wordt (tabel 1).

Tabel 1: De verdeling van de categorieën van vragen die een leerling tegenkomt bij afname van de gea-toets per niveau.

Niveau	**kaartlezen**	**Topografie**	**Landschap**	**Bevolking**	**Diverse**	**Totaal**
1	17	8	5	0	0	30
2	6	3	11	9	1	30
3	1	1	13	14	1	30
4	0	0	14	14	2	30
5	0	0	14	15	1	30

Niet alleen zijn er verschillende aspecten van de geografie die bij de toets aan bod komen, ook zijn er diverse soorten vraagstellingen die aan bod komen (tabel 2). De meest voorkomende vraag is de meerkeuze vraag waarbij een leerling een keuze maakt tussen vier antwoordalternatieven genaamd optie A, B, C en D (eventueel ook een optie E of F). Een combo-box geeft meerdere opties weer, zoals bij de vraag 'Welke taal spreken ze in de Verenigde Staten?', een leerling krijgt hierbij 19 opties. Bij een hotspot vraag dient de leerling een pijltje op de juist plaats op een afbeelding te zetten. Vragen die hierbij horen zijn bijvoorbeeld 'Waar op deze afbeelding vind je de branding?' bij een afbeelding van de kust of 'Waar vaar je over de evenaar?' bij een afbeelding van Zuidoost Azië. Bij een open vraag kan een leerling niet kiezen uit een aantal opties, maar dienen zij zelfstandig het antwoord te formuleren. Dit zijn vaak vragen naar begrippen zoals 'Wat zie je op deze foto?'. Dan zijn er nog twee soorten vragen die heel weinig voorkomen, namelijk de meermeerkeuzevraag (een meerkeuzevraag waarbij meerdere antwoorden goed kunnen zijn) en de gatentekst-vraag (waarbij een leerling een zin compleet moet maken door een bepaald verschijnsel of verband in te vullen).

Tabel 2: Verhouding van de soorten vraagstellingen in de Gea-toets.

Soorten vragen	**Aantal**
Meerkeuze	242
Combo-box	28
Hotspot	27
Woord invullen/ open vraag	20
Meermeerkeuze	4
gatentekst	1
Totaal	**322**

2.2 Geografisch besef in de gea-toets.

In de gea-toets komen niet alleen verschillende onderwerpen en vraagstellingen aan bod, maar deze doen ook beroep op verschillende vaardigheden en inzichten van leerlingen.
Aangezien de gea-toets dient ter beoordeling van de geografische vaardigheden en het geografisch besef van leerlingen dient eerst te worden onderzocht wat onder geografisch besef wordt verstaan.

In deze paragraaf wordt eerst beschreven op wat voor soort kennis en bijbehorende vraagstukken het onderwijs beroep doet bij leerlingen. Vervolgens wordt beredeneerd welke vormen van kennis en (aardrijkskundige) vaardigheden tot geografisch besef gerekend kunnen worden en van al deze typen kennis en vaardigheden zal een voorbeeld worden gegeven van een vraag uit de Gea-toets waarin deze aan bod komt.

2.2.1 Soorten kennis

Er bestaan verschillende soorten kennis waar leerlingen gebruik van maken, Haladyna (2004, p.30) maakt een onderscheid tussen het reproduceren en het beheersen van vier soorten kennis (tabel 1). Bij het reproduceren van kennis wordt de leerling op een toets naar iets gevraagd op dezelfde manier en in precies dezelfde context als waarin de toetsstof eerder is uitgelegd of behandeld. Een voorbeeld hiervan is bijvoorbeeld een toets waarbij de topografische ligging wordt gevraagd van steden of landen die de leerlingen voor de toets kregen opgegeven. Het beheersen van bepaalde kennis is een complexer cognitief proces aangezien de leerling toont de toetsstof in uiteenlopende contexten te kunnen toepassen. Een leerling dient patronen of processen te kunnen herkennen en te benoemen waarmee hij of zij in eerdere lessen of instructies nog niet mee is geconfronteerd (Haladyna, 2004, p.30). In onderstaand schema worden er vier inhoudelijke typen kennis gegeven, namelijk: feiten, begrippen, principes en procedures.

Feiten: Een feit is iets waarvan zeker is dat het gebeurd is, waar is of daadwerkelijk bestaat. Ook al zijn feiten onderhevig aan een sociale context, het blijven gegevens waarover publiekelijke consensus bestaat en worden gezien als onontkenbaar en vaststaand. Voorbeelden van feiten zijn ‘de A is de eerste letter van het alfabet’, ‘Parijs is de hoofdstad van Frankrijk’, ‘Parijs ligt aan de Seine’ en ‘Nederland ligt op het Europese continent’. Vragen naar deze type kennis kunnen op directe wijze worden gesteld (‘Wat is de hoofdstad van Frankrijk?’) waarbij een leerling het juiste antwoord wel of niet weet. Feitenvragen zijn daardoor bijna altijd vragen die over reproduceerbare kennis gaan. Binnen het aardrijkskundig onderwijs valt de topografie in deze categorie van kennis.
Begrippen: Begrippen zijn objecten, verschijnselen of gebeurtenissen die herkenbaar zijn aan een vaststaande combinatie van kenmerken en daarmee omschreven kunnen worden. Geografische begrippen zijn: landbouw, vulkanen, woestijnen, branding, gebergte of aardbevingen. Begrippen kunnen concreet en abstract zijn, zo heeft een begrip als klimaat duidelijkere meetbare kenmerken dan segregatie. Toetsvragen over begrippen kunnen zowel reproductievragen (‘Wat is een vulkaan?’) als vragen zijn waarbij de leerling moet aantonen het verschijnsel te begrijpen (‘Leg uit waarom in Nederland geen vulkanen voorkomen en in Zuid-Italië wel’).
Principes: Met een principe wordt de relatie tussen twee of meerdere concepten bedoeld. Een principe neemt vaak de vorm ‘als…….dan…’ aan. Bijvoorbeeld ‘Naarmate men hoger komt wordt het kouder’. Er zijn twee vormen van principes: ten eerste als vaststaande gegeven (zoals ‘hete luchtstromen stijgen en koude luchtstromen dalen’) en ten tweede als waarschijnlijkheid (‘autorijden onder invloed van alcohol is waarschijnlijk gevaarlijker dan geheel nuchter’). Deze waarschijnlijkheden kunnen door middel van onderzoek worden bevestigd of ontkend (Haladyna, 2004, p.32). Het toetsen van principes kan door middel van reproductievragen (‘Naarmate men hoger komt wordt het dan kouder of warmer?’) maar ook door te kijken of een leerling de stof echt beheerst (‘wat zullen de gevolgen kunnen zijn van grootschalige houtkappingen in wintersportgebieden?’).
Procedures: Een procedure is een serie aan elkaar gerelateerde gebeurtenissen of (kennis)handelingen met een bepaald resultaat. Voorbeelden van mentale procedures zijn: het optellen en aftrekken van getallen, een tekst samenvatten en het berekenen van een gemiddelde. Daarnaast zijn er fysieke procedures die vaak als vaardigheden worden gezien: het slijpen van een potlood of het uitgummen van een tekst. Dit soort vaardigheden komen niet veel aan bod bij theoretische toetsingen maar zijn belangrijk bij examens die de praktijk nabootsen (Haladyna, 2004, p.29).

Niet alle aspecten in bovenstaande opsomming zijn terug te vinden in de opzet van de Gea-toets maar zowel de feitenkennis als de begripskennis en de principes zijn hierin duidelijk terug te vinden. Dit komt omdat op lagere niveaus van de gea-toets naar simpelere feitenkennis en bgripskennis wordt gevraagd terwijl op hogere niveaus wordt verwacht dat leerlingen meer complexere begrips- en principekennis tonen. De geografische vaardigheden die naar voren komen in de Gea-toets doen beroep op de reproduceerbare en beheersende kennis uit dit model.

2.2.2 Geografische visies en vaardigheden

Naast de soorten algemene kennis die Haladyna onderscheidt zijn er specifieke geografische vaardigheden en kennis waar de Gea-toets beroep op doet. De vele visies op deze geografische vaardigheden geven aan dat het niet eenvoudig is tot een eenduidige indeling hiervan te komen.
In het boek *Geographic Information Systems in inquiry-based secondary geography education* (2011, p.35-36) probeert Favier een beeld te schetsen van alle soorten vaardigheden en kennis die behoren tot het domein van geografisch besef. Dat de 'key concepts' van de geografie onderhevig zijn aan verschillende visies blijkt uit onderstaande tabel van begrippen die door diverse auteurs als 'geografisch' worden beschouwd.

Tabel 3: geografische kernbegrippen volgens verschillende auteurs over het geografisch vakgebied.

Author(s)	*Concepts*
Leat (1998): 'the big concepts of geography'	(1) Cause and Effect; (2) Classification; (3) Decision-making; (4) Development; (5) Inequality; (6) Location; (7) Planning; and (8) Systems
Kaufman (2004): 'the spatial-temporal primitives'	(1) Place; (2) Size; (3) Shape; (4) Distance; (5) Direction; (6) Connectivity; (7) Containment; (8) Pattern; (9) Duration; (10) Sequence; and (11) Frequency.
Holloway, Rice, and Valentine (2003): 'the key concepts'	(1) Space; (2) Time; (3) Place; (4) Scale; (5) Social formations; (6) Physical systems; and (7) Landscape and Environment
Jackson (2006): 'the key concepts'	(1) Space and Place; (2) Scale and Connection; (3) Proximity and Distance; and (4) Relational thinking
Qualifications and Curriculum Development Agency (2007): 'the key concepts'	(1) Place; (2) Space; (3) Scale; (4) Interdependence; (5) Physical and Human Processes; (6) Environmental interaction and Sustainable Development; and (7) Cultural Understanding and Diversity
Taylor (2008): 'the key concepts'	(1) Diversity; (2) Change; (3) Interaction; and (4) Perception and Representation
Janelle and Goodchild (2009): 'the foundation concepts for spatial thinking'	(1) Location; (2) Distance; (3) Network; (4) Neighbourhood and Region; (5) Scale; (6) Spatial heterogeneity; (7) Spatial difference; (8) Spatial dependence; and (9) Objects and Fields.
Colledge (2008): 'the geospatial concepts'	(1) Place-specific identity; (2) Location; (3) Magnitude; (4) Time; (5) Boundary and Distance; (6) Distribution or Arrangement; (7) Regions; (8) Frames of Reference; (9) Orientation and Direction; (10) Pattern; (11) Clustering and Dispersion; (12) Spatial Association; and (13) Density and Distance Decay
Gersmehl (2005): The 13 aspects of spatial thinking	(1) Location; (2) Condition; (3) Connection; (4) Comparison; (5) Aura; (6) Region; (7) Transition; (8) Analog; (9) Pattern; (10) Pattern comparison; (11) Rule and Exceptions; (12) Pattern change; and (13) Spatial model
Van der Schee (1987): 'the geographic core concepts'	(1) Location; (2) Distance; (3) Spatial distribution; (4) Spatial-property distribution; (5) Spatial association; (6) Spatial interaction; (7) Spatial system; and (8) Spatial diffusion
Jo and Bednarz (2009): the concepts of spatial thinking	(1) Magnitude; (2) Location; (3) Place-specific identity; (4) Enclosure; (5) Adjacency; (6) Arrangement; (7) Reference frame; (8) Shape; (9) Region; (10) Boundary; (11) Transition; (12) Movement; (13) Connection and Linkage; (14) Direction; (15) Distance; (16) Buffer; (17) Map projection; (18) Scale; (19) Relief; (20) Profile; (21) Gradient; (22) Layer; (23) Overlay; (24) Spatial association; (25) Hierarchy and Network; (26) Dominance; (27) Diffusion; (28) Density; (29) Dispersion and Clustering; (30) Pattern; and (31) Distribution

(Bron: Favier, 2011, p.35)

Een van de visies op geografie is die van Gersmehl (2008, p.101) die drie basis concepten, die de kern van de geografie vormen, onderscheidt: locatie, omstandigheden en connecties. Onder locatie valt

zowel topografische kennis als de manier om kaarten te gebruiken. Ook begrippen om afstand te beschrijven vallen hieronder: richting, aangrenzend of afgesloten. Het gaat dus om het herkennen van locaties en het doorgeven van deze informatie aan anderen. Onder de omstandigheden worden begrippen verstaan die worden gebruikt om de situatie van een specifieke plaats te omschrijven. Het gaat hierbij om het klimaat, de flora en fauna, de economie, de religie en de cultuur van een locatie en haar inwoners. Onder de connecties worden de manieren verstaan die de ene plaats aan de andere verbindt. Dit kan bijvoorbeeld een fysische verbinding zijn, zoals de wind die van de ene plaats naar de andere gaat, of economisch zoals woon-werkverkeer. Uit deze drie concepten komt zijn opvatting voort dat het vergelijken van en beschrijven van locaties een belangrijke geografische vaardigheid is (Gersmehl, 2008, p.105).

Rekening houdend met de verschillende visies op de concepten van Gersmehl en Favier die het geografisch besef vormen en de soorten kennis van Haladyna, zijn de vormen van geografische kennis zoals zij naar voren komen in de Gea-toets volgens de indeling van tabel 4 samengesteld.

Tabel 4: kennis en vaardigheden die voorkomen in de Gea-toets.

	Soorten vaardigheden en kennis	Vaardigheden en soorten kennis die voorkomen in de gea-toets
Vaardigheden	Informatie lezen	Teksten: - ***Begrijpend lezen***
		Beelden: - ***Natuurgetrouw (foto)*** - ***Abstract (kaart/grafiek)***
	Informatie analyseren	- ***Wisselen van analyseniveau en ruimtelijke schaal*** - ***Maken van vergelijkingen***
Kennis	Feitenkennis	- ***Feitenkennis*** - ***Thematisch kaartbeeld*** - ***Topografisch kaartbeeld***
	Begripskennis	- ***Begripskennis***
	Kennis van principes	- ***Kennis van principes***

Leerlingen die de Gea-toets maken worden dus op tien soorten geografische kennis en vaardigheden beoordeeld, namelijk: Begrijpend lezen, foto-interpretatie, kaart-en grafieklezen, wisselen van schaalniveau, maken van vergelijkingen, feitenkennis, thematisch kaartbeeld, topografisch kaartbeeld, begripskennis en kennis van principes. Hieronder wordt van elke vaardigheid en kennissoort een definitie en een voorbeeld gegeven.

Vaardigheden

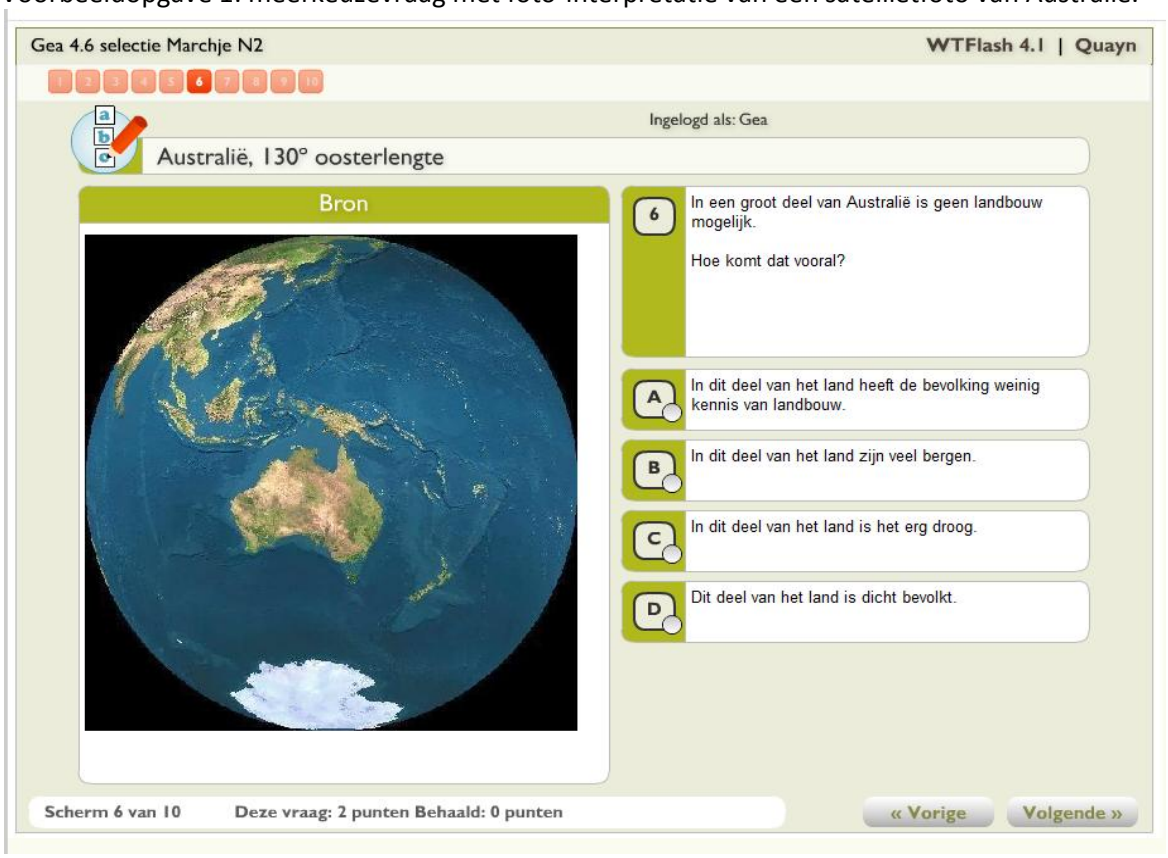
-Begrijpend lezen
De eerste vaardigheid die van belang is voor het maken van de gea-toets is niet van uitsluitend geografische aard maar des al niet te min van belang. Computervaardigheden en met name begrijpend lezen zijn noodzakelijk voor het beantwoorden van de vragen. Ondanks dat dit onderzoek niet dient ter beoordeling van deze vaardigheden zal begrijpend lezen wel worden opgenomen in de beoordelingscriteria van de 'hard-op-denk'-protocollen. Het is belangrijk erbij stil te staan gezien het feit dat sommige leerlingen moeite hebben met begrijpend lezen en/ of omgaan met computertaken

en hier zowel tijdens de kwantitatieve als de kwalitatieve data-verzameling tegenaan zouden kunnen lopen.

- Foto-interpretatie

Bij de geografische vaardigheid foto-interpretatie dienen leerlingen elementen die kenmerkend zijn voor bepaalde landen en gebieden aan de hand van een foto te herkennen als zodanig (Stichting Cito Instituut, 2012). De leerling stelt zichzelf hierbij de vraag: Welk verschijnsel zie ik hier?
Een vraag uit de Gea-toets waarbij dat naar voren komt is de vraag in voorbeeldopgave 1 waarbij een leerling uit een satellietfoto van Australië dient af te leiden waarom in een groot deel van Australië geen landbouw mogelijk is. Het juiste antwoord, 'in dit deel van Australië is het veel te droog', kan worden beredeneerd door de grote gele oppervlakten op de afbeelding van Australië juist te interpreteren. Daarnaast zijn er veel vragen in de gea-toets opgenomen waarbij de vaardigheid foto-interpretatie slechts indirect van belang is. Een voorbeeld hiervan is een foto-vraag waarbij een leerling moet aangegeven welke van de vijf foto's niet in bijvoorbeeld Frankrijk is genomen. Ondanks dat foto-interpretatie nodig is in deze vraagstelling om te begrijpen waarvan de foto een afbeelding is zijn het andere vaardigheden zoals het thematisch kaartbeeld waar de leerling beroep op moet doen om tot het juiste antwoord te komen.

Voorbeeldopgave 1: meerkeuzevraag met foto-interpretatie van een satellietfoto van Australië.



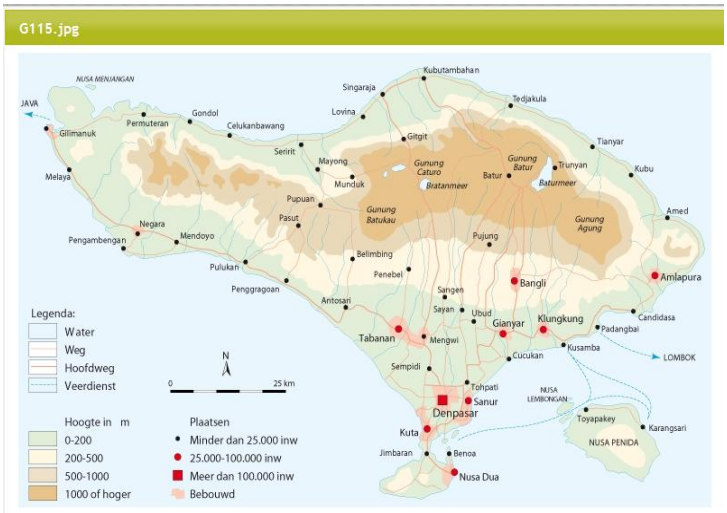
-Kaart- en grafiek lezen

Leerlingen kunnen verschillende kaarten interpreteren en kaartvaardigheden gebruiken met kaarten van verschillend schaalniveau met verschillende thema's (Stichting Cito Instituut, 2012). Voorbeeld opgave 2 is een vraag uit de Gea-toets waarbij de leerling wordt geacht een kaart te kunnen begrijpen en daar conclusies uit te kunnen trekken. Zo heeft de kaart een legenda, windrichtingen en geeft het hoogten weer door middel van verschillende kleuren. Dat de enige vaardigheid waar deze vraag op is gebaseerd kaartlezen is blijkt uit het feit dat er geen feitenkennis over Bali nodig is om de vraag te beantwoorden. Wanneer dezelfde vraag was gesteld over een ander land met een

bijpassende kaart had de leerling alsnog dezelfde denkhandelingen moeten verrichten om tot het juiste antwoord te komen.

Voorbeeld opgave 2: Meerkeuzevraag met kaartlezen van een kaart van Bali.

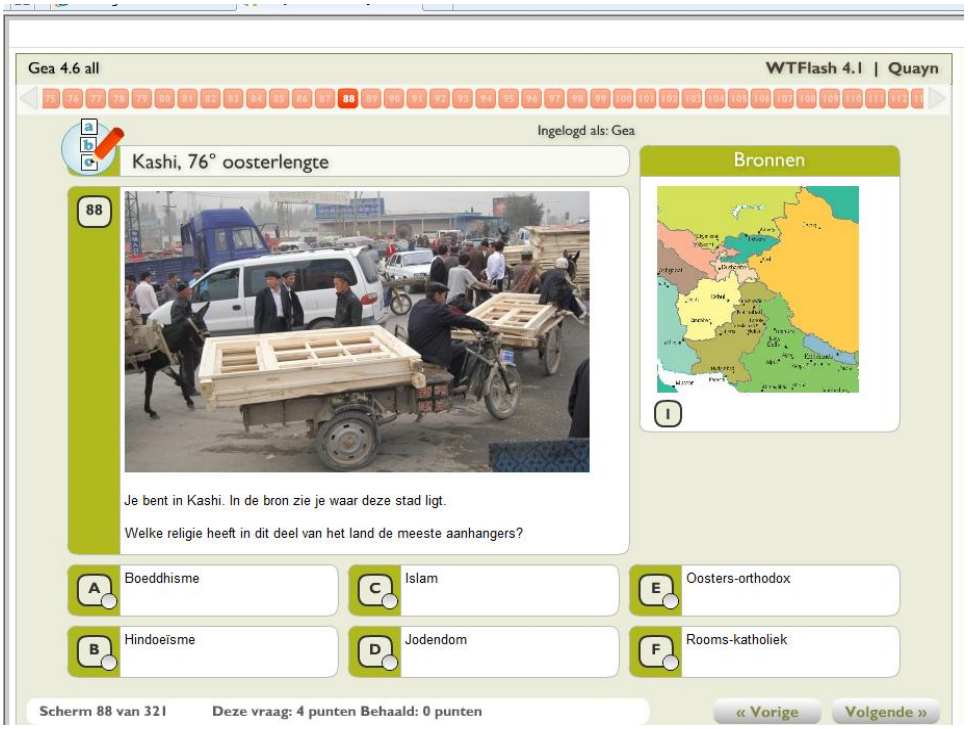
Afbeelding 1: De uitvergrote afbeelding van de kaart bij voorbeeldopgave 2.



Naast kaarten zijn er overigens ook enkele grafieken opgenomen in de gea-toets, deze geven de temperatuur en neerslag weer van een gebied waarbij de leerling kan aangeven tot welk klimaat het gebied behoort.

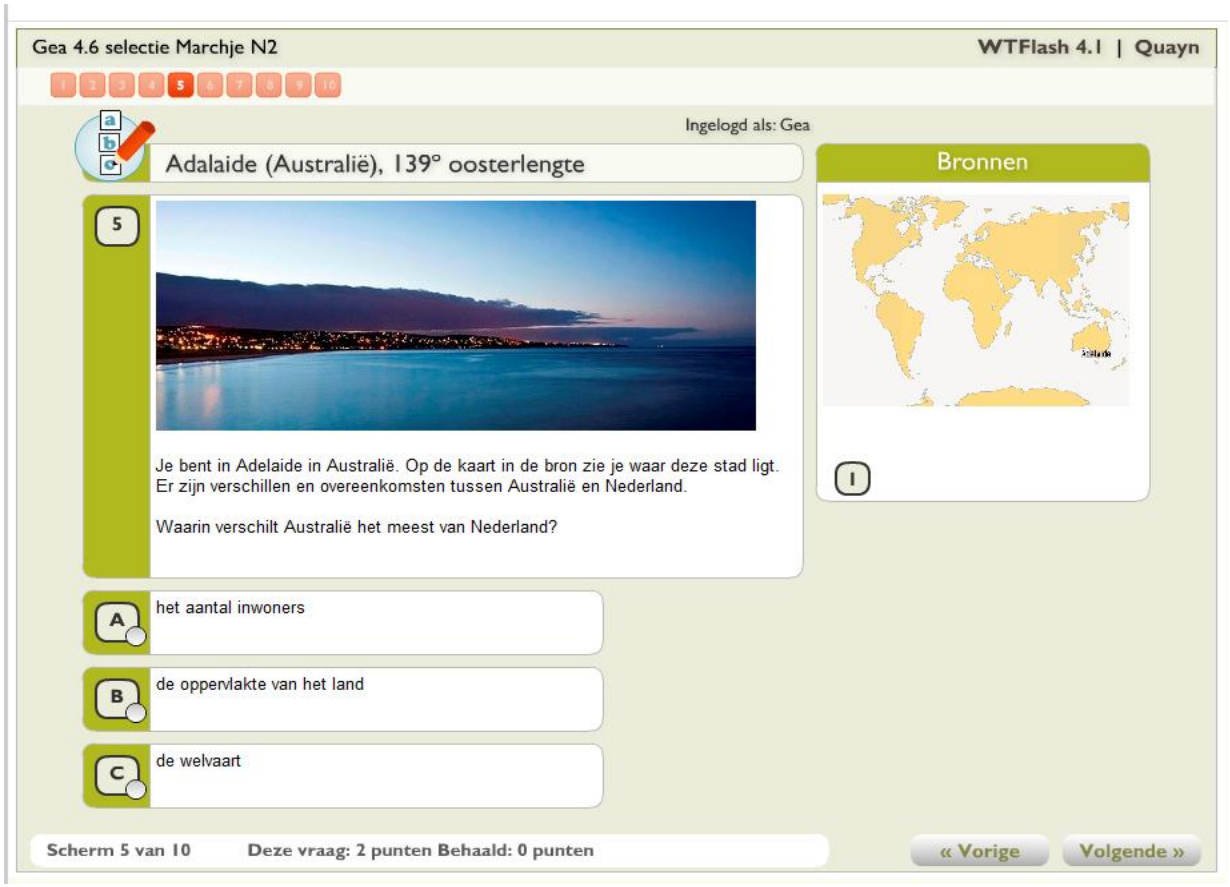
- Kennis van schaalniveau's
Kennis van schaalniveau's wordt gezien als een belangrijk onderdeel van geografisch besef. Uit het pabo-handboek 'Aardrijkskunde is overal' blijkt dat het kunnen wisselen van schaalniveau een vaardigheid is die op de basisschool al aan bod dient te komen (Notté, H. & J.K.Verheij, 2007, p.165). Het gebruik van verschillende schaalniveau's houdt in dat leerlingen weten waar het desbetreffende gebied ligt, dat je daar vanuit kunt uitzoomen en de kennis van belangrijke kenmerken van het grotere gebied kunt toepassen op het kleinere gebied.
Een voorbeeld van een opgave waarbij verschillende schaalniveau's worden gebruikt is voorbeeldopgave 3 over Kashi in Iran. Er wordt gevraagd welke religie in dit deel van het land de meeste aanhangers heeft. Allereerst dient de leerling door middel van het bijgevoegde kaartje in de bron erachter te komen in welk land en in welke regio deze stad ligt. Vervolgens dient de leerling mentaal 'uit te zoomen' naar een hoger schaalniveau (het Midden-Oosten) om daaruit af te leiden welke religie in deze regio veel voorkomend is.

Voorbeeldopgave 3: Meerkeuzevraag met schaalniveaus over Kashi (Iran).



- Vergelijking
Één van de kerndoelen van het aardrijkskundig onderwijs is dat leerlingen een vergelijking kunnen maken tussen de eigen omgeving en een andere omgeving, tussen binnen- en buitenland en vanuit verschillende perspectieven zoals landbouw, wonen, werken, bestuur, etcetera (Stichting Cito Instituut, 2012). Het draait bij het maken van de vergelijking om de relatie tussen de gebieden. Een voorbeeld van een vraag uit de Gea-toets waarbij het maken van een vergelijking wordt toegepast is in voorbeeldopgave 4 over Australië. Er wordt gevraagd waarin Nederland en Australië het meeste verschillen, het aantal inwoners, de oppervlakte of in welvaart. Het is de bedoeling dat leerlingen mentaal een vergelijking maken tussen Nederland en Australië met de begrippen 'aantal inwoners', 'oppervlakte' en 'welvaart', deze met elkaar vergelijken en daaruit concluderen dat het grootste verschil te maken heeft met het oppervlakte.

Voorbeeldopgave 4: Meerkeuzevraag met een vergelijking tussen Nederland en Australië.

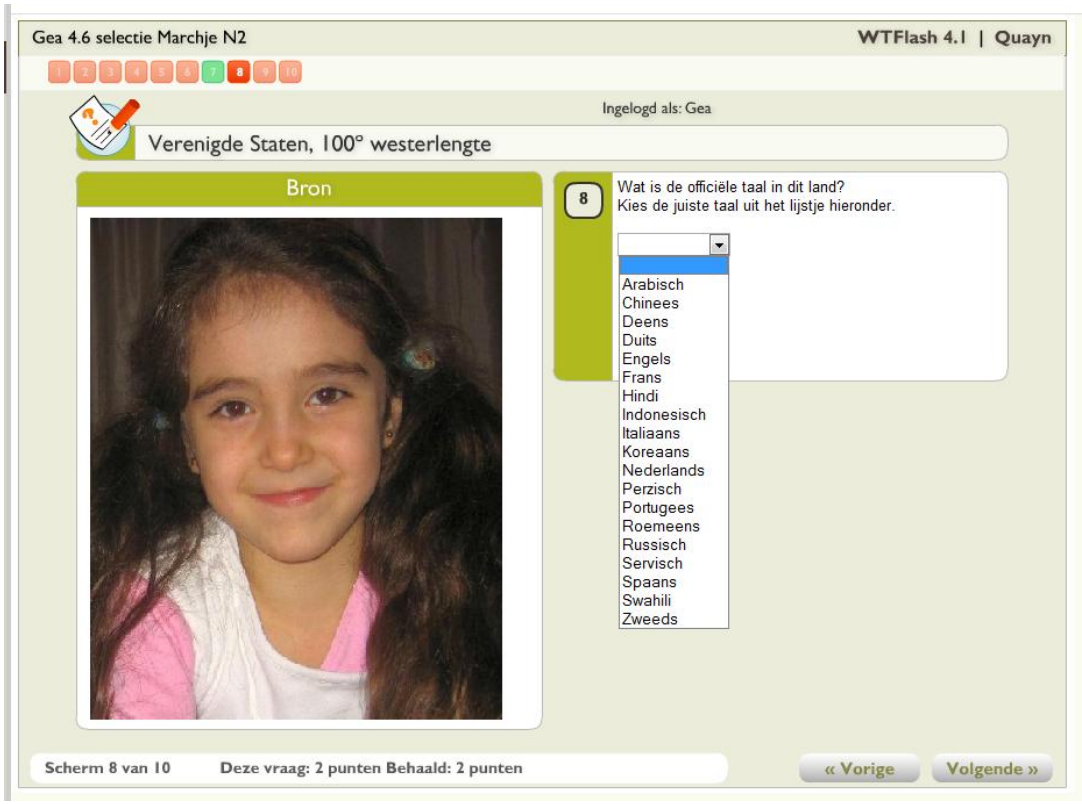


Kennis

- Feitenkennis

Een belangrijk onderdeel van geografische kennis is feitenkennis. Met feitenkennis alleen kan geen inzicht worden gegeven in de meer complexe levensomstandigheden van mensen die in bepaalde gebieden wonen, maar zij is wel een randvoorwaarde om meer complexere geografische begrippen te kunnen plaatsen. Voorbeeldopgave 5 is een feitenkennisvraag die in de Gea-toets voorkomt. Er wordt gevraagd naar de officiële taal in de Verenigde Staten. Dit zou door leerlingen beredeneerd kunnen worden wanneer zij een ‘mental map’ hebben van de spreiding van de talen over de wereld, maar het is waarschijnlijker dat dit voor de meeste leerlingen een feitenvraag is.

Voorbeeldopgave 5: Feitenkennisvraag door middel van een combo-box naar de taal van de VS.

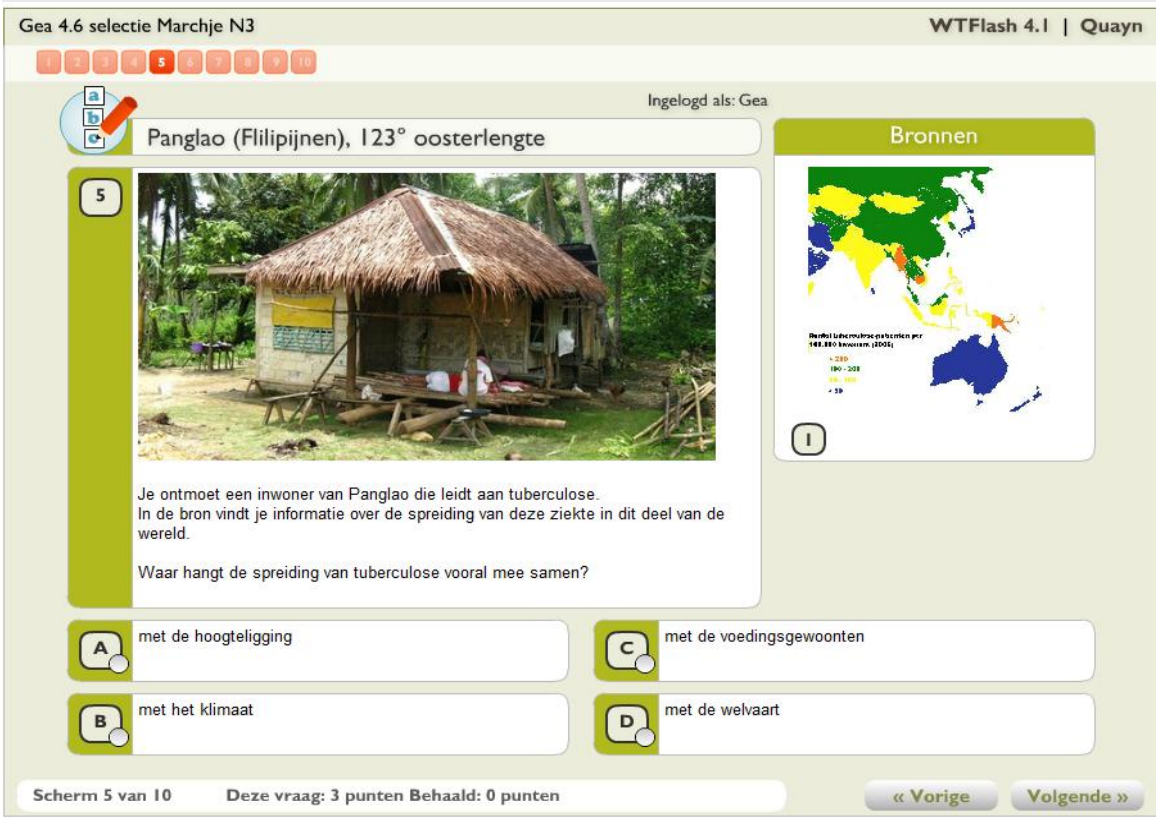


Naast de pure feitenvragen zoals in voorbeeldopgave 5 maakt de Gea-toets gebruik van twee soorten feitenkennis: het thematisch kaartbeeld en het topografisch kaartbeeld.

- Het thematisch kaartbeeld

Een Thematisch kaartbeeld houdt in dat leerlingen een ‘mental map’ hebben van de wereld waarbij zij van de ruimtelijke verdeling van verschillende thema’s bewust zijn. In een thematische ordening kan samenhang tussen verschijnselen beter worden beschreven en komen belangrijke begrippen van bijvoorbeeld de fysische geografie of economische geografie meer tot hun recht (Notté, 2002, p.17). Voorbeelden van een thematisch kaartbeeld zijn ‘zwakke/sterke economie’, ‘religie’ of ‘klimaat’. Leerlingen kunnen de ruimtelijke spreiding van verschijnselen aangeven en de samenhang tussen verschijnselen in een gebied begrijpen en verklaren (Stichting Cito Instituut, 2012). In de voorbeeldopgave wordt aan de leerling gevraagd waar de spreiding van tuberculose in Zuidoost Azië mee samenhangt, met de hoogteligging, het klimaat, voedingsgewoonten of de welvaart. Om deze vraag juist te kunnen beantwoorden moet de leerling zijn ‘mental map’ van het thema ‘welvaart’ aanspreken. Dan is het duidelijk dat de verschillen in welvaart overeenkomen met de verschillen in tuberculosespreiding.

Voorbeeldopgave 6: meerkeuzevraag naar het thematisch kaartbeeld over Tuberculose in Zuidoost Azië.

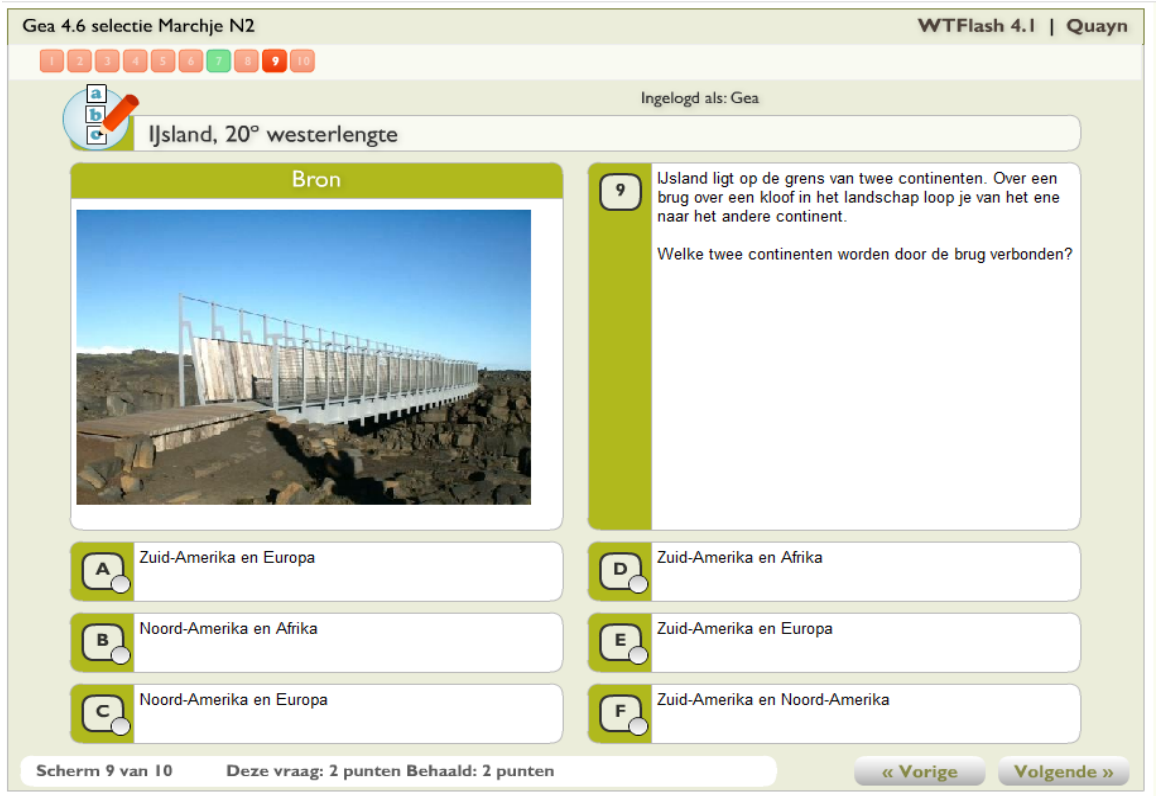


- Het topografisch kaartbeeld

Bij het topografisch kaartbeeld gaat het om het kunnen reproduceren van kennis over de ligging van landen, steden en natuurlijke verschijnselen zoals vulkanen. Leerlingen dienen de basistopografie van Nederland, Europa en de rest van de wereld te kennen en een topografisch wereldbeeld te ontwikkelen (Stichting Cito Instituut, 2012). Voorbeeldopgave 7 is een duidelijk voorbeeld van een opgave waarbij beroep wordt gedaan op topografische kennis van de leerling.

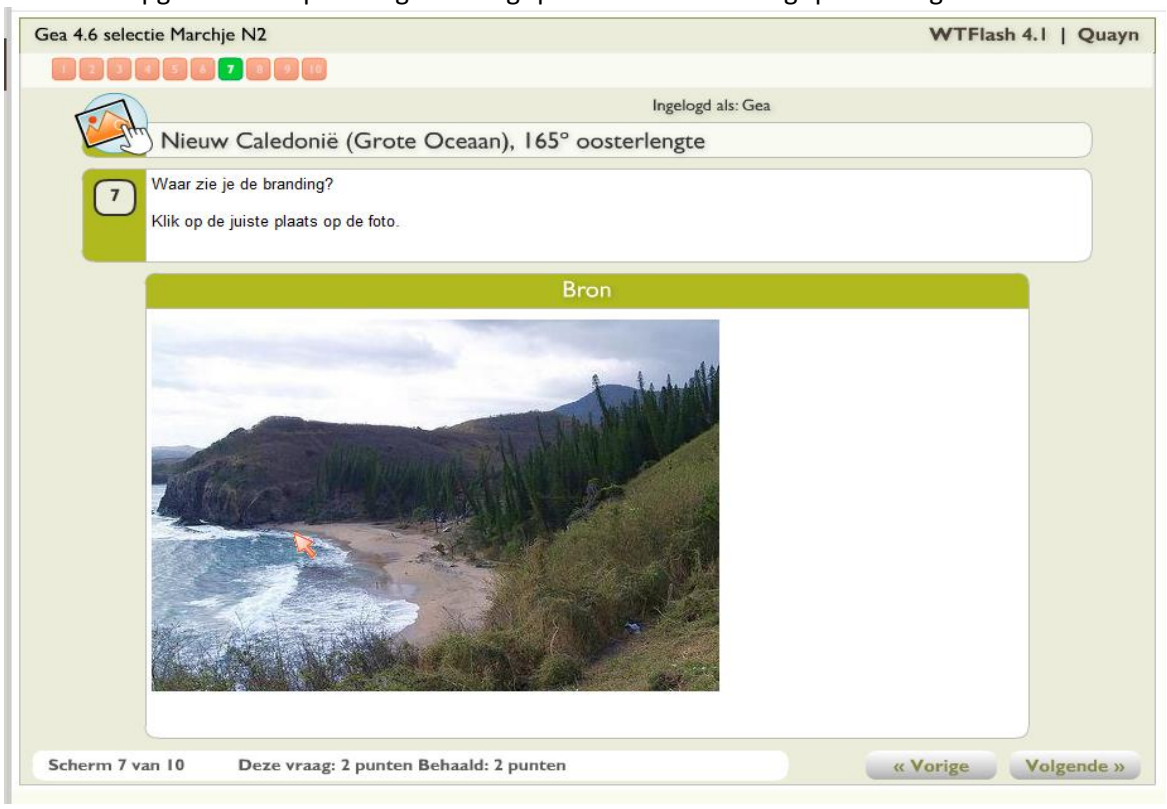
De vraag die hier wordt gesteld is op welke twee continenten het eiland IJsland ligt. Om de vraag te kunnen beantwoorden is het belangrijk dat leerlingen IJsland op de wereldkaart kunnen plaatsen, dat zij weten dat het land tussen Europa en Noord-Amerika ligt. Hiervoor dienen zij de namen van de continenten te kennen.

Voorbeeldopgave 7: Meerkeuzevraag naar topografisch kaartbeeld over de ligging van IJsland



- Begripskennis

Voorbeeldopgave 8: Hotspotvraag naar begripskennis over het begrip branding.



Begripskennis is een vorm van feitenkennis voor complexere begrippen. Het gaat hierbij om geografische termen zoals stuwwallen, uiterwaarden of vulkanen. Het aardrijkskundige schoolvak dient aandacht te besteden aan de begrippen waarmee kenmerken van de bevolking, evenals het landschap waarin zij wonen, worden beschreven en verklaard (Stichting Cito Instituut). Voorbeeldopgave 8 toont een vraag naar een geografisch begrip, de leerling wordt gevraagd met behulp van de muis aan te geven waar de branding op de foto wordt afgebeeld. Om deze vraag goed te kunnen beantwoorden dient een leerling te weten wat het begrip branding inhoudt.

- Kennis van principes.
Kennis van bepaalde principes zoals die bij tabel 4 zijn uitgelegd zijn nodig bij het beantwoorden van diverse vraagstukken van de gea-toets. Om bijvoorbeeld voorbeeld vraag 1 te kunnen beantwoorden dient men te weten dat droge gebieden een heel ongunstig klimaat voor landbouw is.

Kortom: geografisch besef is een verzamelnaam voor een aantal soorten van kennis en vaardigheden waarmee men (ruimtelijke) verschijnselen in een gebied mee kan verklaren en voorspellen. Bovenstaande indeling suggereert dat de gea-toets een strikt onderscheid maakt tussen de verschillende soorten van kennis en in welke vragen welke kennissoorten naar voren komen. In de praktijk, zoals al zichtbaar was bij kennis van principes, loopt dit meer door elkaar. Zo doet voorbeeldvraag 3 over Kashi in Iran niet alleen beroep op het vermogen van de leerling om uit te zoomen naar een hoger schaalniveau maar ook op het thematisch kaartbeeld van de leerling over de spreiding van religies over de aarde. Het kunnen combineren van verschillende soorten kennis en vaardigheden is daardoor essentieel voor het geografisch besef.

3 Toetsconstructie, validiteit en betrouwbaarheid

Dit hoofdstuk geeft een literatuuronderzoek weer waarin de begrippen die een belangrijke rol spelen in deze scriptie verder worden toegelicht. Allereerst wordt er in paragraaf 3.1 dieper ingegaan op de verschillende facetten van toetsconstructie zodat duidelijk wordt welke plaats het valideringsproces in neemt bij de ontwikkeling van de Gea-toets. In paragraaf 3.3 worden de begrippen validiteit en betrouwbaarheid behandeld. Deze begrippen spelen bij toetsconstructie een belangrijke rol aangezien de toets zowel valide als betrouwbaar dient te zijn.

3.1 Toetsconstructie

Het construeren van een toets verloopt volgens verschillende stappen. Het *handbook of test development* (Downing, S.M. & T.M. Haladyna e.a.,2006) kent twaalf stappen in dit proces die bij het Cito (Stichting Cito Instituut, 2011) doorgaans worden samengevat in de volgende acht stappen die ook bij de constructie van de gea-toets zijn aangehouden:

Stap 1: Doelspecificatie
De eerste stap vormt het overkoepelende plan voor de ontwikkeling van een toets. In deze fase worden de belangrijkste beslissingen genomen, zo wordt het doel en de doelgroep van de toets bepaald en welke toetsvorm hiervoor het meest geschikt is. Een andere belangrijke vraag is op welke manier de toets de capaciteiten van een leerling beoordeelt, dient de toets een selectie, predictie of certificatie te maken? Daarnaast worden alle essentiële taken, verantwoordelijkheden en fasen vastgelegd in een tijdsplanning (Downing, S.M. & T.M. Haladyna e.a.,2006, p.6).

Stap 2: Toetsspecificatie
Bij de stap van toetsspecificatie wordt bepaald hoeveel tijd het afnemen van de toets zou moeten innemen en hoe de toets afgenomen dient te worden. Daarnaast wordt de inhoud van de toets vastgesteld. Volgens het *handbook of test development* (Downing, S.M. & T.M. Haladyna e.a.,2006, p.7) is dit de belangrijkste stap in de eerste fase van toetsontwikkeling. De inhoud van de toets komt voort uit een analyse van (school)boeken en ander leermateriaal en wordt vaak samengebracht in een domeinbeschrijving. Zo staat er bijvoorbeeld in de domeinbeschrijving voor de Cito-eindtoets op de lagere school aangegeven welke topografische kennis de toetsconstructuren als bekend mogen beschouwen. Daarnaast wordt in een toetsmatrijs weergegeven hoe de opgaven, behorende bij bepaalde doelstellingen, worden verdeeld over tenminste twee dimensies: inhoudscategorieën en gedragscategorieën. De toetsmatrijs geldt als een blauwdruk dat een systematische constructie van een toets wil garanderen (Stichting Cito Instituut, 2011).

Stap 3: Itemconstructie
Dit is de stap waarin de items (de toetsopgaven) worden geconstrueerd. De inhoud van de items dienen inhoudelijk overeen te komen met de in de domeinbeschrijving en de toetsmatrijs beschreven leerstof. Een veel voorkomende format waarin een item verschijnt is de multiple choice vraag. Deze manier van toetsen wordt in het *handbook of test development* het werkpaard van de toetsingsindustrie genoemd vanwege de mogelijkheid deze veelzijdig in te zetten. Het kan worden gebruikt om alle verschillende niveau's van cognitieve denkprocessen te beoordelen en vanwege het beperkte aantal keuzemogelijkheden laten de antwoorden zich goed omzetten in een score (Downing, S.M. & T.M. Haladyna e.a.,2006, p.10).

Stap 4: Toetsafname (proefafname)
In deze fase zijn de items klaar om te worden voorgelegd aan de doelgroep. Allereerst dient er een try-out plaats te vinden waarbij vooral de nadruk zal komen te liggen op de werking van de toets en de eerste indruk van de leerlingen. Zijn er onderdelen onduidelijk? Of, zoals de vraag kan zijn bij een computertoets, doen alle functies van het programma wat ervan verwacht wordt?

Vervolgens volgt er een grotere proefafname waarbij er vooral op gelet wordt of de toets inhoudelijk aansluit bij wat men van de doelgroep kan verwachten. Een proefafname is bedoeld om een indruk te krijgen van hoe de items inhoudelijk en psychometrisch functioneren bij de leerlingen waarvoor de definitieve toets bedoeld is. Op basis van de resultaten van de proefafname zullen sommige items verwijderd of gereviseerd worden (Stichting Cito Instituut, 2011). Op dit moment bevindt de validering van de gea-toets zich in deze fase. De vragen zijn geconstrueerd en de doelgroepen en methoden afgebakend. De vraag is nu of de items daadwerkelijk het verwachte onderscheid maken tussen leerlingen met een zwak of sterk ontwikkeld geografisch besef.

Stap 5: Itemevaluatie
Nadat de toets is afgenomen in de steekproef dienen de daaruit voortgekomen data te worden geanalyseerd. Wat er in de toets- en itemanalyse naar voren komt zijn het aantal leerlingen die de toets hebben afgelegd en het aantal items in de toets. Daarnaast wordt er per item weergegeven: het percentage leerlingen die de vraag goed beantwoord hebben, de percentages van leerlingen die kozen voor de alternatieven, of er sprake is van ontbrekende waarnemingen, de correlatie met de rest van de toets, de betrouwbaarheid van het item en de score verdeling (Stichting Cito instituut, 2011).

Stap 6: Toetssamenstelling
Vervolgens worden de items die tijdens de proeftoetsing goed het onderscheid bleken te maken tussen vaardige en minder vaardige leerlingen geselecteerd waaruit een toets wordt opgesteld. Daarbij wordt erop gelet dat de vragen de verschillende onderwerpen van het toetsdomein aan bod laten komen (Stichting Cito instituut, 2011).

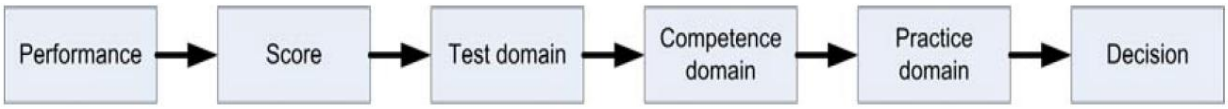
Stap 7: Referentiekader bepalen
In deze stap dient het referentiekader te worden bepaald en vindt er een normeringsonderzoek plaats. Het kan zijn dat de normering achteraf wordt aangepast afhankelijk van de uitslag (Stichting Cito instituut, 2011).

Stap 8: Handleiding/verantwoording
De laatste fase van het toetsconstructieproces bestaat uit het maken van een handleiding en instructies voor de diverse categorieën personen die bij de toetsing betrokken zijn. Ten behoeve van de opdrachtgever en het wetenschappelijk forum wordt er een verantwoording geschreven (Stichting Cito instituut, 2011).

3.2 Toetsontwerp

Het toetsontwerp laat zien welke aspecten invloed hebben op het tot stand komen van een toets en wat hun functies zijn. In het model van toetsontwerp (figuur 1) staan de verschillende toetsonderdelen schematisch weergegeven waaruit de toets uiteindelijk bestaat.

Figuur 1: Model van toetsontwerp



(Bron: Wools, S., T.Eggen & P.Sanders, 2009, p.4)

Performance staat hierin voor de prestatie van een leerling op een toets, deze wordt uitgedrukt in een score. Deze score moet worden geplaatst in het perspectief van het toetsdomein waarin staat welke toetsvragen en welke onderwerpen aan bod laten komen tijdens de toets. Zijn de vragen van de toets representatief voor alle vragen waar de toets over beschikt? Zijn alle onderdelen in de toets

vertegenwoordigd? Het toetsdomein is opgesteld aan de hand van het competentiedomein, hierin staat beschreven op welke competenties de toets een beroep op doet. Zijn alle belangrijke toetsdoelen vertegenwoordigd in het totesdomein? De competenties zijn gebaseerd op de practice domain, dit zijn de vaardigheden die iemand nodig heeft in het dagelijks leven aangaande het desbetreffende vakgebied. In het geval van de gea-toets zijn dit alle vaardigheden en kennis die worden gerekend tot geografisch besef? Uiteindelijk leidt dit tot een decision, een oordeel, over iemands kennis of vaardigheden. Wanneer dit schema wordt toepast op geografische kennis kun je zeggen dat het in de practice domain handig is dat mensen enige kennis hebben van het gebruiken van kaarten en weten waar belangrijke steden van Europa liggen. In de competentie domein zou dan kaartlezen en topografische kennis de doelstelling kunnen zijn. In het toets domein komt deze kennis naar voren in verschillende opgaven waarbij leerlingen demonstreren dat ze bepaalde verschijnselen met behulp van een kaart kunnen localiseren en laten zien dat ze beschikken over topografische kennis.

Momenteel bevindt de ontwikkeling van de Gea-toets zich in vierde stap van de toetsconstructie. De toets is klaar, alle items zijn geconstrueerd, nu moet nog onderzocht worden of de niveaus die in de gea-toets worden gehanteerd goed zijn ingeschat en of de toets daadwerkelijk een valide oordeel geeft over het geografisch besef van de doelgroep. In het model van toetsontwerp vindt dit onderzoek met name plaats in het toets domein en competentiedomein (verdere uitleg over positionering van dit onderzoek binnen het valideringsproces van de Gea-toets wordt in paragraaf 4.1 gegeven).

3.3 Validiteit en betrouwbaarheid

3.3.1 Validiteit

Dit onderzoek is een onderdeel van het valideringsproces van de Gea-toets. Om te weten welke plek dit onderzoek inneemt in dit proces zou eerst onderzocht moeten worden wat er wordt verstaan onder validering. Validiteit gaat over de interpretatie, het gebruik en de betekenis van toetsscores, oftewel, het is de bedoeling dat de scores op een toets iets zeggen over de vaardigheden van de leerling op het te toetsen onderwerp. Validiteit draait dus om de vraag of toetsscores de informatie bieden waarin we geïnteresseerd zijn en of op basis van deze scores bepaalde beslissingen genomen kunnen worden (Stichting Cito instituut, 2011, p.66). Valide waarnemingen geven (gemiddeld) de juiste uitkomsten, maar deze zijn onjuist wanneer zij systematisch een te hoge of te lage waarde hebben of wanneer zij een andere eigenschap meten dan bedoeld was. Bij onderzoek (zoals door middel van enquêtes) wordt validiteit bepaald door een juiste en concrete vraagstelling, bij het trekken van steekproeven wordt de validiteit bepaald aan de hand van een correcte manier van steekproeftrekking (de Vocht, 2009).
Het aantonen van validiteit kan worden gedaan door hiervoor verschillende bewijzen te vinden die gezamenlijk laten zien in hoeverre de toetsscores geschikt zijn om een bepaald doel of bepaalde vaardigheden te meten. Uiteindelijk zegt de combinatie van verschillende bewijzen uit diverse typen valideringstoetsen iets over de validiteit als geheel (Stichting Cito instituut, 2011, p.66). In de publicatie 'toetsen op school' wordt door het Cito zes verschillende type validiteitsbewijzen besproken.

- Inhoudsbewijzen: Inhoudsbewijzen laten zien dat de items die in de toets opgenomen zijn representatief zijn voor het leerstofdomein of vaardigheid waarover men uitspraken wilt doen (Stichting Cito instituut, 2011, p.67). Een inhoudsbewijs voor een toets zou een toetsmatrijs kunnen zijn waarin wordt aangegeven hoe de items over verschillende toetsonderdelen verdeeld zijn. Een ander inhoudsbewijs is een oordeel van docenten van het desbetreffende vak over het niveau en de relevantie van de items.
- Complexiteitsbewijzen: Een complexiteitsbewijs gaat dieper in op de kennishandelingen die van leerlingen wordt gevraagd bij het oplossen van een taak, leerlingen moeten de leerstof kunnen

oproepen, toepassen en analyseren. De complexiteit van de items van de toets dienen overeen te komen met de complexiteit van cognitieve processen bij de taken van het leerstofdomein. Bij het aantonen van de juiste complexiteit dienen de vaardigheden waarbij de kennishandelingen een rol spelen bij de onderzoekers bekend te zijn en moet bewezen worden dat leerlingen deze handelingen daadwerkelijk verrichten bij het oplossen van een taak. 'Hard-op-denk'-protocollen zouden hier als bewijs voor deze handelingen kunnen gelden. Hierbij wordt aan leerlingen gevraagd om hardop te beredeneren bij het maken van een aantal opgaven. Zodoende kan worden nagegaan of de juiste kennishandelingen worden gebruikt door de leerlingen (Stichting Cito instituut, 2011, p.68).
- Interne structuur bewijzen: Interne structuur duidt op de onderlinge verhouding tussen de toetsvragen en tussen de toetsvragen en de toets als geheel. De samenhang tussen de vragen zal groot zijn wanneer zij dezelfde soort vaardigheid toetst (bijvoorbeeld alle vragen gaan over kaartlezen). Anderzijds, wanneer er meerdere vaardigheden worden getoetst zal de samenhang minder zijn (bijvoorbeeld de samenhang tussen vragen over kaartlezen en vragen over taalvaardigheid). Deze samenhang wordt vaak kwantitatief getoetst en uitgedrukt in een correlatiecoëfficiënt (Stichting Cito instituut, 2011, p.69) .
- Externe structuur bewijzen: Met de externe structuur worden de resultaten van de leerlingen op de toets vergeleken met die van andere toetsen. De andere toetsscores en/of het oordeel van de docent heet dan een extern criterium. Een voorbeeld is een korte en een lange toets van een bepaalde vaardigheid. Wanneer leerlingen voor de korte toets een score van gelijke betekenis halen als bij een andere, lange toets dan betekent dit dat een korte toets volstaat om een betrouwbaar beeld te schetsen van de capaciteiten van een leerling (Stichting Cito instituut, 2011, p.69).
- Betrouwbaarheidsbewijzen: Betrouwbare toetsscores zijn een voorwaarde voor valide toetsscores (alhoewel het mogelijk is een betrouwbare toetsscore te hebben die niet valide is). Dit houdt in dat de vaardigheid op een juiste manier moet worden gemeten, maar dat het ook om de juiste vaardigheid gaat (Stichting Cito instituut, 2011, p.70). Meer hierover in paragraaf 3.2.2.
- Extrapoleerbaarheidsbewijzen: Extrapoleerbaarheidsbewijzen laten zien dat de geleverde prestaties tijdens een toets niet alleen een betekenis hebben in de context van de toets maar ook betekenis hebben in de beroepspraktijk. Hierbij is authenticiteit erg belangrijk, het vereiste van de toets dient zoveel mogelijk te lijken op de praktijk. Het praktijkexamen voor het behalen van het rijbewijs is hier een voorbeeld van (Stichting Cito instituut, 2011, p.71).

3.2.2 Betrouwbaarheid

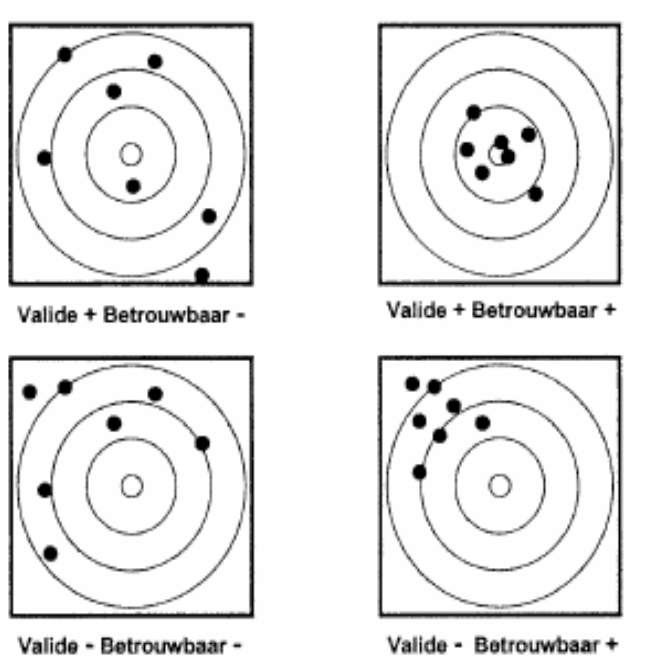
Betrouwbaar zijn waarnemingen die onder dezelfde omstandigheden herhaald, dezelfde uitkomst zullen opleveren. Betrouwbare uitkomsten hebben weinig spreiding, zo geeft een enquête een onbetrouwbaar beeld van opinies als de steekproef klein is en de meningen over het onderwerp sterk verdeeld (de Vocht, 2009).

Bussen rijden niet altijd op tijd, het weer kan ineens omslaan en ons humeur verschilt van dag tot dag. Kortom, iedere meting heeft te maken met toevallige omstandigheden. Hoe minder de meting afhankelijk is van toeval (variantie), des te betrouwbaarder is de meting. Variantie in toetsprestaties voortkomend uit toevallige factoren wordt beschouwd als construct-irrelevant variance (CIV) en ondermijnt de validiteit van de interpretatie van de toetsuitslagen (Downing, S.M. & T.M. Haladyna e.a., 2006, p.29). Om de betrouwbaarheid van de meting te vergroten dienen de omstandigheden waarin de metingen worden afgenomen zoveel mogelijk gelijk te zijn (Baarda & de Goede, 2009).

Onderstaand figuur geeft door middel van vier schietschijven de verschillende verhoudingen tussen validiteit en betrouwbaarheid weer. De schietschijf linksboven toetst de juiste vaardigheden, maar de variantie tussen de scores is hoog in tegenstelling tot de schietschijf rechtsonder. Hier is de variantie klein maar wordt de verkeerde vaardigheid gemeten. De schietschijf die linksonder is weergegeven is zowel niet valide als onbetrouwbaar vanwege de grote variantie in resultaten. De schietschijf rechtsboven geeft de gewenste verdeling van validiteit en betrouwbaarheid weer: de variantie tussen de resultaten is minimaal en de juiste vaardigheid/kennis wordt getoetst. Het

onderzoek dient dus zowel betrouwbaar (zo min mogelijk CIV) als valide (de juiste kennis wordt gemeten) te zijn.

Figuur 2: Validiteit en Betrouwbaarheid



(Bron: de Vocht, 2009)

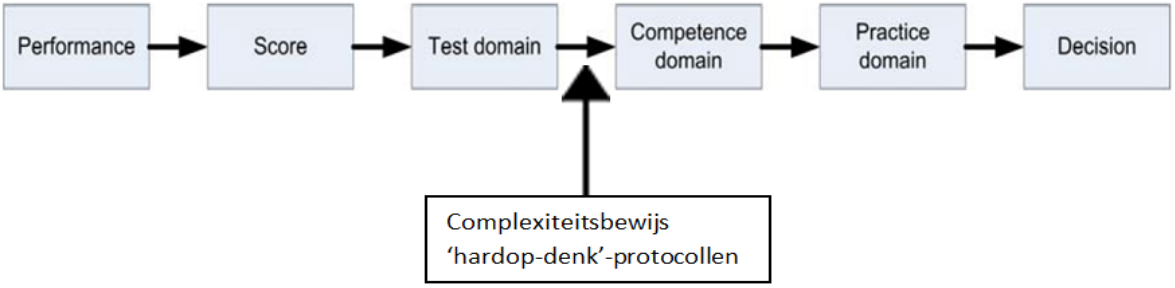
4 Onderzoeksmethode en methodologie

In dit hoofdstuk wordt behandeld hoe de informatie over toetsconstructie is geoperationaliseerd en terugkomt in de kwantitatieve en kwalitatieve dataverzameling van dit onderzoek. In paragraaf 4.1 wordt uitgelegd welke plaats dit onderzoek in neemt in het gehele valideringsproces van de Gea-toets. Paragraaf 4.2 laat zien hoe de hoofdvraag heeft geleid tot het formuleren van een conceptueel model en hypothese. Paragraaf 4.3 behandelt hoe het onderzoek is opgezet en een antwoord kan geven op de hoofdvraag. Paragraaf 4.4 toont de operationalisering van het begrip 'Geografisch besef' waarvan de betekenis zoals die in dit onderzoek wordt gehanteerd in het theoretisch kader is gemotiveerd. Daarnaast komt ook de opstelling van de 'hard-op-denk'-protocollen aan bod. Ten slotte wordt in paragraaf 4.5 wordt de keuze voor de scholen die aan het onderzoek hebben geparticipeerd verklaard, evenals hoe het motiveren van de leerlingen zal worden bevorderd.

4.1 Motivatie onderzoeksmethode en plaats van dit valideringsonderzoek binnen het valideringsproces van de gea-toets

Voor het valideren van toetsen is er een combinatie aan bewijzen nodig van zowel kwalitatieve als kwantitatieve aard (Stichting Cito Instituut, 2011, p. 67). De Gea-toets zal dus voordat zij door leerlingen en docenten in gebruik kan worden genomen onderworpen worden aan een aantal valideringstesten om te kunnen beoordelen of op basis van de Gea-toets een uitspraak zou kunnen worden gedaan over de vaardigheden van een leerling. Dit onderzoek is hier een onderdeel van en is gebaseerd op de hoofdvraag: ' ***Op welke soort geografische kennis doen leerlingen beroep bij het beantwoorden van de vragen van de Gea-toets en wat zegt dit over de validiteit van de toets?'***
Dit houdt in dat onderzocht wordt of de competenties die de opgave-makers voor oog hadden bij het construeren van de toets ook daadwerkelijk door de toetsvragen bij leerlingen worden aangesproken. Zo kan het zijn dat het de bedoeling is dat een leerling het juiste antwoord bij een bepaalde opgave beredeneert door middel van de vaardigheid kaartlezen, maar dat hij in werkelijkheid op het juiste antwoord terecht komt door feitenkennis. De vraag is dus: Zijn de vaardigheden die behoren tot het competentie domein op een juiste manier omgezet in toetsvragen?
Deze vraag kan worden beantwoord door middel van een complexiteitsbewijs aangezien men de kennishandelingen dient te achterhalen die van leerlingen gevraagd wordt bij het oplossen van een opgave. Het complexiteitsbewijs voor de Gea-toets wordt op twee manieren geleverd, enerzijds zal aan docenten uit het primaire en secundaire onderwijs worden gevraagd wat zij vinden van het niveau van de vragen en anderzijds zal er gebruik worden gemaakt van 'hard-op-denk'-protocollen die als bewijs voor de denkhandelingen van de leerlingen zou kunnen gelden. De 'hard-op-denk'-protocollen vormen de kern van dit onderzoek en op deze manier vormt dit onderzoek een schakel in het valideringsproces van de Gea-toets (figuur 3).

Figuur 3: De plaats van dit valideringsonderzoek door middel van 'hardop-denk'-protocollen van de Gea-toets in het gehele toetsontwerp.



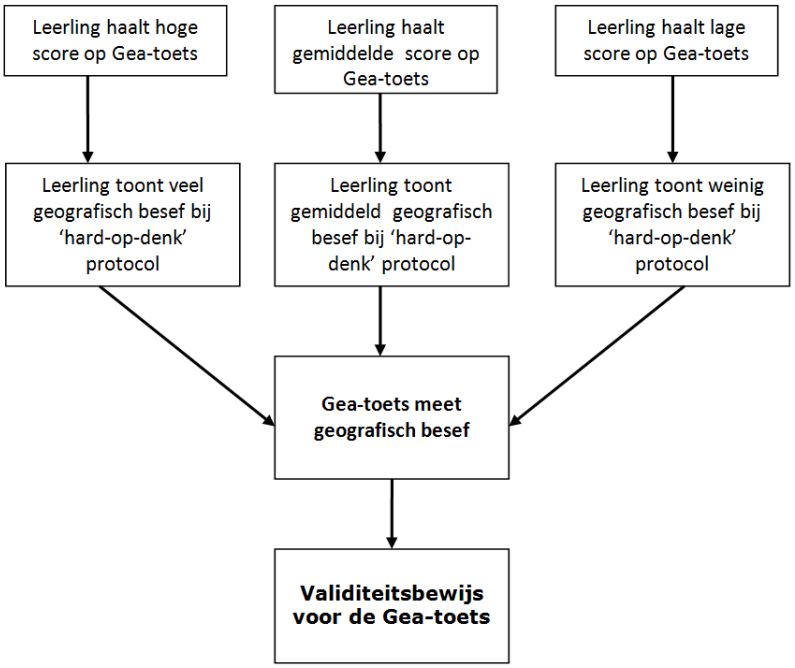
Bron: Wools, S., T.Eggen & P.Sanders, 2009, p.4

Met dit onderzoek zal getracht worden de hoofdvraag te beantwoorden aan de hand van de 'hardop-denk'-protocollen waarbij zes geselecteerde opgaves, die de verschillende geografische vaardigheden van de toets vertegenwoordigen, met leerlingen doorgenomen zullen worden.

4.2 Conceptueel Model en Hypothese

De verwachting is dat leerlingen met een hoge score op de toets ook meer geografische vaardigheden zullen tonen tijdens de 'hard-op-denk'protocollen dan leerlingen met een lage score. Het doel van de Gea-toets is door middel van een aantal opgaven leerlingen een score te kunnen geven die iets zegt over hun geografisch besef. Daarom is het belangrijk dat de leerlingen die een hoge score hadden op de toets tijdens het kwalitatieve deel blijk geven van een verder ontwikkeld geografisch besef dan leeftijdsgenoten met een lagere score. Daarnaast zouden zij moeten laten zien dat zij de denkstappen zetten bij het oplossen van een opgave die de toetsmakers voor ogen hadden. Wanneer het hun klasgenoten met lagere toetsscores niet lukt deze denkstappen te zetten tijdens de 'hard-op-denk'-protocollen zou dit een indicatie kunnen zijn dat de Gea-toets inderdaad geografisch besef meet. Dit zou betekenen dat de Gea-toets een valide toets is en dat men op basis van de toetsuitslagen een oordeel kan vellen over het werkelijke geografisch besef van een leerling.
Deze beredenering kan worden samengevat in het volgende conceptueel model:

Figuur 4: Conceptueel model



Uit het conceptueel model kan de volgende hypothese worden afgeleid:

Leerlingen met een hogere score op de Gea-toets zullen tijdens de 'hard-op-denk'-protocollen ook meer vaardigheden en denkprocessen tonen die toe te schrijven zijn aan geografisch besef dan leerlingen met een lagere score op de Gea-toets.

4.3 Onderzoeksopzet

In paragraaf 3.1 werd besproken op welke wijze het complexiteitsbewijs 'hardop-denk'-protocollen zal worden gebruikt om de hoofdvraag te beantwoorden. Het onderzoek bestaat uit twee delen aangezien het kwalitatieve onderzoek (de 'hardop-denk'-protocollen) een kwantitatief

vooronderzoek vereist. Het kwantitatieve onderzoek zal worden gebruikt om een indicatie te geven welke leerlingen geschikt zijn voor het kwalitatieve onderzoek.

Het kwantitatieve gedeelte van het onderzoek:
Om het niveau van de leerlingen vast te stellen zullen een aantal klassen (drie klassen groep 8 leerlingen van de lagere school en vier klassen in het voortgezet onderwijs) worden gevraagd de Gea-toets op de computer af te nemen. In werkelijkheid zal de Gea-toets een adaptieve toets worden met vijf verschillende niveaus waarop leerlingen kunnen uitkomen afhankelijk van hun geografisch besef, maar voor dit onderzoek zullen de groep 8 leerlingen 32 vragen worden voorgelegd op niveau 2 en de derde klas-leerlingen 35 vragen van niveau 3. De reden voor deze keuze is dat het niveau van de vragen voor niveau 2 is vastgesteld op basis van de domeinbeschrijving van het aardrijkskunde onderwijs op de lagere school. Hier wordt ook de Cito eind-toets op gebaseerd en het is waarschijnlijk dat dit niveau aansluit op de kennis van groep-8-leerlingen. De vragen voor de derde klassers zijn opgesteld aan de hand van het vereiste niveau voor het aardrijkskunde- eindexamen in de vierde klas van het VMBO. Leerlingen uit Havo 3 en VWO 3 zouden deze vragen moeten kunnen beantwoorden. De antwoorden zullen kwantitatief worden geanalyseerd en zodoende zal worden vastgesteld of de leerling een lage, gemiddelde of hoge score heeft behaald in vergelijking met zijn of haar klasgenoten. Op basis van deze uitslag worden later leerlingen uitgekozen om het 'hard-op-denk' protocol af te laten nemen.

Het kwalitatieve onderzoek:
Aan de hand van de scores die uit het kwantitatieve onderzoek naar voren komen zullen leerlingen met uiteenlopende scores (een aantal met hoge scores, lage scores en gemiddelde scores) later worden benaderd. Zij zullen zes opgaven gaan maken (zes niveau 2 opgaven voor groep8 leerlingen en zes niveau 3 opgaven voor derde klas leerlingen) die representatief zijn voor de vaardigheden die de toets vereist volgens de 'hard-op-denk' protocollen. Er wordt dus van de leerlingen verwacht dat zij hardop redeneren, waardoor het duidelijk wordt welke geografische vaardigheden worden gebruikt bij het beantwoorden van een vraag. Na afloop van het bespreken van iedere opgave wordt aan de leerling gevraagd of hij of zij het land van de opgave kan aankruisen op een blinde wereldkaart (wel met staatskundige grenzen). Dit is om een beeld te krijgen van het topografisch kaartbeeld van de leerling. Aardrijkskunde is veel meer dan alleen topografie, maar topografie speelt wel een grote ontwikkeling bij het geografisch besef van leerlingen. Wanneer een docent bepaalde thema's wil behandelen met de leerlingen zoals de spreiding van welvaart, religie, of klimaten wordt dit heel lastig als leerlingen een weinig ontwikkeld topografisch kaartbeeld hebben. In dat opzicht kan topografisch kennis als een basis van het aardrijkskundig onderwijs worden gezien die in de laatste klassen van het basisonderwijs wordt aangelegd. In het middelbaar onderwijs wordt alle nieuwe leerstof geplaatst in dit topografische referentiekader.
De verwachting is dat leerlingen die een hoge score hadden tijdens de afname van de kwantitatieve toets en de vragen van de 'hard-op-denk'-protocollen goed beantwoorden ook een completer topografisch kaartbeeld zullen hebben dan leerlingen die veel vragen onjuist beantwoorden. Daarnaast kan het ook zijn dat een leerling een vraag niet juist heeft beantwoord maar wel precies het land uit de opgave weet aan te wijzen. Op dat moment is het duidelijk dat het onjuist gegeven antwoord niet aan een gebrek aan topografische kennis van de leerling ligt, maar aan een onvolledig thematisch kaartbeeld.
Vervolgens wordt bij enkele opgaven nog gevraagd aan de leerling of hij/zij kan aangeven waar op de wereld een bepaald verschijnsel nog meer zou kunnen voorkomen (bijvoorbeeld bij een vraag over oases). Hieruit blijkt dat een leerling weet wat het begrip inhoudt en onder welke omstandigheden dit verschijnsel zich voordoet. Hieruit kan men opmaken in welke mate het thematisch kaartbeeld van de leerling is ontwikkeld. Aangezien de ontwikkeling van thematisch kaartbeeld zich voornamelijk afspeelt op de middelbare school, waarbij het topografisch kaartbeeld als uitgangspunt dient, is de verwachting dat met name leerlingen die een hoge score hadden op de Gea-toets op niveau drie beschikken over een goed ontwikkeld thematisch kaartbeeld.

Om onderzoeksgegevens zo betrouwbaar mogelijk te verkrijgen zullen de gesprekken verlopen volgens een vaste procedure die is vastgelegd in het protocol. Dit betekent dat bij iedere leerling wordt gestreefd het gesprek op een zo veel mogelijk gelijke manier te laten verlopen. De exacte inhoud van het protocol en het tot stand komen ervan zal in paragraaf 3.4 worden toegelicht. De gesprekken zullen worden opgenomen om te voorkomen dat informatie verloren gaat.

4.4 Operationalisering van Geografisch Besef en 'hard-op-denk'protocollen

Om het kwalitatieve onderzoek uit te voeren dienen eerst opgaven uit te worden gekozen die hierbij aan bod kunnen komen en de verschillende geografische vaardigheden die in paragraaf 2.3 van het theoretische kader benoemd zijn te vertegenwoordigen. Op die manier kan worden getest welke geografische vaardigheden de leerlingen beheersen.

Elke leerling die deelneemt aan het kwalitatieve onderzoek wordt zes vragen voorgelegd, bij groep 8 leerlingen zes vragen op niveau 2 en bij derde klas leerlingen zes vragen op niveau 3. In onderstaand schema staat per vraag genoteerd op welke soort geografische vaardigheden en kennis zij beroep doet wanneer een leerling deze probeert te beantwoorden. Zij zijn voor dit onderzoek uit een totaal van 69 vragen per niveau gekozen omdat deze vragen een goede afspiegeling zijn voor de verschillende soorten geografische vaardigheden en kennis in de Gea-toets. Uit de overige vragen zijn steekproefgewijs 32 vragen op niveau 2 en 35 vragen op niveau 3 geselecteerd voor de kwantitatieve afname. Een aantal van deze opgaven zijn al besproken in paragraaf 2.2.

Tabel 5: Vragen van de 'hard-op-denk'-protocollen ingedeeld op de soort geografische kennis en vaardigheden waarop zij gebaseerd zijn.

Vraag	Niveau	Onderwerp	BL	FI	KL/GR	KS	VG	FK	ThK	Tok	BK	PR
1	2	Brunico (Italië)	X	X	X	X			X			
2	2	Kuweit	X		X	X				X		
3	2	Garmeh (Iran)	X	X		X					X	
4	2	Bali	X		X							
5	2	Adalaide (Australië)	X				X	X				
6	2	Verenigde Staten	X					X				
1	3	Haditha (Irak)	X								X	X
2	3	Bangkok	X	X	X	X			X			
3	3	Panglao	X		X				X			
4	3	Rio de Janeiro	X		X						X	
5	3	Groenland	X				X		X			
6	3	Belfast	X					X		X		

Vaardigheden:
BL = Begrijpend Lezen
FI = Foto-interpretatie
KL/GR = Kaart- en Grafiek lezen
KS =Kennis van schaalniveaus
VG = Vergelijking

Kennis:
FK = Feitenkennis
ThK = Thematisch kaartbeeld
ToK =Topografisch kaartbeeld
BK =Begripskennis
Pr = Kennis van Principes

Nu er in totaal twaalf opgaven zijn geselecteerd die gedurende de 'hard-op-denk'-protocollen de leerlingen zullen worden voorgelegd dient het protocol te worden opgesteld. 'Hard-op-denk'-protocollen houden in dat een aantal leerlingen, in groepjes of individueel, wordt gevraagd opgaven hardop redenerend te maken. De onderzoeker legt de gesprekken vast met behulp van audio- of videotapes en stimuleert de leerlingen om hun gegeven antwoord te motiveren. Haladyna (2004,

p.200) geeft weer dat er twee mogelijkheden zijn om deze methode te gebruiken, ten eerste als een onderzoeksmethode en ten tweede als valideringsmiddel. Bij het gebruik van 'hard-op-denk'-protocollen als onderzoeksmethode is er eerder onderzocht op welke manier studenten een meerkeuzevraag beantwoorden. Zo bleek dat studenten meerkeuzevragen op vijf verschillende manieren kunnen benaderen, op 16 verschillende manieren de opties overwegen en vier cognitieve strategieën hebben om tot een juist antwoord te komen. Daarnaast kan het 'hard-op-denk'-protocol worden gebruikt om de inhoud van een toets en de achterliggende cognitieve processen te rechtvaardigen en beargumenteren. Verder kan op deze manier het juiste antwoord en de effectiviteit van de alternatieven worden beoordeeld. Uitkomsten van 'hard-op-denk-'protocollen vormen op die manier een validiteitsbewijs zowel voor de inhoud als voor het gewenste cognitieve gedrag (Haladyna, 2004, p.201).

Het protocol vormt de leidraad van de gesprekken met de leerlingen tijdens het kwalitatieve deel van het onderzoek en houdt in dat de gesprekken zoveel mogelijk op dezelfde manier zal worden vormgegeven. Hierdoor is de kans dat de onderzoeker de leerlingen (onbewust) naar een bepaald antwoord leidt kleiner, wat de betrouwbaarheid van het onderzoek vergroot. Het protocol zal bij iedere besproken opgave in elk gesprek erbij gehouden en ingevuld worden en bevat de volgende onderdelen:

Gegevens leerlingen

Allereerst dienen de gegevens van de leerlingen worden ingevuld, hierbij wordt de leerling gevraagd naar de naam en de klas. De onderzoeker vult zelf het niveau van de vraag in (voor groep 8 leerlingen niveau 2 en voor derde klas leerlingen niveau 3) en het nummer van de vraag.

Introductie van het onderzoek

Wanneer het de eerste vraag is die de leerling bij het 'hard-op-denk'-protocol dient te beantwoorden zal de onderzoeker zichzelf voorstellen en een korte toelichting op het onderzoek geven. Hierbij zal even worden teruggekomen op de Gea-toets die tijdens het kwantitatieve gedeelte van de toets in de klas is afgenomen. De leerling zal daarnaast worden uitgelegd wat van hem/haar verwacht wordt.

Leerling dient de vraag hardop voor te lezen

Wanneer de leerling een venster opent met een nieuwe vraag zal deze, evenals alle gegeven antwoorden, hardop voor worden gelezen. Op deze manier is het voor de onderzoeker duidelijk dat wanneer een leerling het antwoord op een vraag niet weet of een onjuist antwoord geeft dit niet wordt veroorzaakt doordat de leerling de vraag niet goed heeft kunnen lezen. Dit zou bijvoorbeeld in het geval van dyslectische leerlingen kunnen voorkomen.

De leerling geeft antwoord

Nadat de vraag opgelezen is dient de leerling de vraag te beantwoorden. De onderzoeker zal tot het moment dat de leerling het antwoord geeft geen uitspraken doen om te voorkomen dat de leerling in het beantwoorden van de vraag wordt gestuurd.

Motivatie keuze antwoord

Wanneer de leerling het antwoord heeft gegeven zal de onderzoeker de leerling vragen zijn keuze te motiveren. Waarom heeft de leerling gekozen voor antwoord C en niet voor antwoord A, B of D? Uit de beredenering van de leerling kan worden afgeleid van welke geografische vaardigheden een leerling gebruik maakt bij het beantwoorden van de vraag. Een leerling die een vraag juist beantwoordt, waarbij gebruik wordt gemaakt van een kaart, beschikt over de vaardigheid 'kaartlezen'. Een leerling kan zijn vaardigheden op twee manieren tonen, enerzijds door te beredeneren waarom bijvoorbeeld antwoord A juist is en anderzijds door alternatieven weg te strepen. Dus antwoord A moet juist zijn, want het is niet antwoord B,C of D. Beide manieren kunnen worden gebruikt om tot het juiste antwoord te komen aangezien ze beide laten blijken van geografische kennis.

In het schema van het 'hard-op-denk'-protocol kan worden aangekruist welke vaardigheden de leerling toont. Beheerst een leerling een vaardigheid goed, dan krijgt hij daar twee punten voor, beheerst hij een vaardigheid matig dan één punt en ontbreekt de vaardigheid volledig bij de motivatie van het antwoord dan krijgt een leerling geen punten. Uiteindelijk geeft het totaal aantal punten van de leerling bij de zes opgaven weer hoe vaardig de leerling is op het gebied van aardrijkskunde.

Bekend met het gebied.
In welke mate een leerling bekend is met een bepaald gebied kan invloed hebben op de beantwoording van de vraag. Een leerling die vaak op (wintersport-) vakantie is geweest naar Sankt Moritz (Zwitserland) weet uit ervaring misschien dat men in Zwitserland Duits en Frans spreekt en geen deel uitmaakt van de Euro-zone terwijl andere leerlingen dit kunnen beargumenteren doordat zij een 'mental map' hebben van de Euro-landen. Deze vraag doet bij de leerling die er regelmatig is geweest beroep op een andere soort kennis (namelijk op feitenkennis) dan bij de leerlingen die het juiste antwoord kunnen bedenken met behulp van hun 'mental maps' (behorende tot thematisch kaartbeeld). Om te achterhalen in hoeverre dit van invloed is wordt de leerlingen gevraagd of ze bekend zijn met het gebied. Er zijn op deze vraag vier categorieën van antwoorden, namelijk: 'nee', 'nee, wel bovenliggend', 'ja van horenzeggen' en 'ja, ik ben er geweest'.
In het geval dat een leerling volledig onbekend is met het bestaan van een land of gebied zal het antwoord hier 'nee' zijn. De categorie 'nee, maar wel bovenliggend' houdt in dat de leerling weinig of geen kennis heeft van het specifieke land maar wel bepaalde kenmerken van het gebied kan beredeneren vanuit kennis over het bovenliggende gebied. Een leerling die geen kennis heeft over bijvoorbeeld het land Bhutan, maar de vraag juist kan beantwoorden omdat zijn of haar kennis over Azië wel toereikend is, behoort tot deze categorie. De categorie 'Ja, van horenzeggen' houdt hier in dat de leerling wel eens van het specifiek gevraagde gebied of land heeft gehoord, bijvoorbeeld doordat hier op school een keer aandacht aan is besteed, de leerling heeft hierover een spreekbeurt gegeven of iets over het land op televisie gezien. De laatste categorie in welke mate een leerling bekend is met het gebied in de vraagstelling is 'ja, ik ben er geweest'. Deze leerling zou het juiste antwoord kunnen beargumenteren vanuit de feitenkennis die is opgedaan tijdens het bezoek aan dat gebied zoals de leerling die ieder jaar op wintersportvakantie gaat naar Sankt Moritz (Zwitserland).

Aanwijzen op de kaart
Vervolgens zal de leerling worden gevraagd of hij of zij het gebied uit de vraagstelling kan aanwijzen op een blinde kaart van de wereld. De reden hiervoor is om te achterhalen of de leerling een topografisch kaartbeeld van het gebied heeft. Op deze vraag zijn drie antwoorden mogelijk: 'precies', 'bovenliggend' en 'fout'. Wanneer een leerling het gebied of land exact kan omcirkelen weet een leerling 'precies' waar deze ligt. Bij een 'bovenliggend' gebiedskennis zal een leerling niet exact weten waar het gevraagd gebied ligt maar wel weten waartoe het gebied behoort. Hiertoe behoort bijvoorbeeld een leerling die niet precies Irak kan aanwijzen op de kaart, maar wel een cirkel kan zetten om het gebied waartoe Irak behoort (het Midden-Oosten). Bij 'fout' weet een leerling niet waar het gevraagde gebied ligt of de leerling plaatst het gebied in de geheel verkeerde regio.

Waar komt dit verschijnsel verder voor?
Uiteindelijk wordt nog aan de leerling, mits die de vraag juist heeft beantwoordt, of hij of zij kan aanwijzen waar het verschijnsel van de vraag verder op de wereld nog voorkomt. Zo wordt aan leerlingen gevraagd waar ter wereld men landen kunt vinden die: bijvoorbeeld oud-koloniën van Europese landen zijn, waar Engels wordt gesproken of waar de bevolking grotendeels boeddhistisch is. Hieruit blijkt of een leerling over een goed thematisch kaartbeeld aangaande dat onderwerp beschikt.

Het uiteindelijke protocol zoals die zal worden gebruikt tijdens de 'hard-op-denk'-protocollen is opgenomen in bijlage 1.

4.5 Motivatie keuze voor welke scholen en welke groepen leerlingen

Bij de selectie van scholen voor dit onderzoek heb ik mij grotendeels laten leiden door praktische overwegingen. Dit is mogelijk doordat het onderzoek geen specifieke persoonskenmerken vereist anders dan dat leerlingen in ofwel groep 8 lagere school ofwel derde klas van het voortgezet onderwijs zitten. In de onderzoeksopzet ben ik uitgegaan van in ieder geval twee klassen op de lagere school en twee klassen in het derde jaar van de middelbare school, maar aangezien dit onderzoek afhankelijk is van de beschikbaarheid van scholen en docenten is hiervan afgeweken. Bij het benaderen van scholen voor participering aan dit onderzoek werd in eerste instantie gedacht aan scholen waar nog een binding mee bestond. Vandaar dat de Koninklijke Scholengemeenschap Apeldoorn (KSG) in aanmerking kwam, zelf heb ik op deze middelbare school gezeten. Bij de KSG werd enthousiast op mijn verzoek gereageerd en kreeg ik de mogelijkheid de toets in drie klassen af te nemen. Vervolgens zijn de Koningin Wilhelmina School in Apeldoorn, de Vondelschool in Amersfoort en het Coornhert Gymnasium in Gouda bereid gevonden te participeren aan het onderzoek.

De Koningin Wilhelmina School

Op de Koningin Wilhelmina School waren twee groep 8 klassen beschikbaar voor de afname van de toets. De Gea-toets is in twee groepen 8 afgenomen op maandag 7 mei en dinsdag 8 mei om 09:00. De 'hard-op-denk'-protocollen zijn donderdag 10 mei afgenomen.

De Vondelschool

Op maandag 16 april is in groep 8a van de Vondelschool de gea-toets gemaakt. Op woensdag 18 april zijn de 'hard-op-denk'-protocollen afgenomen.

De KSG

Op de KSG waren drie klassen beschikbaar om de toets bij af te nemen, bij twee havo-3 klassen en bij een atheneum-3 klas. De Gea-toets is in klas H3a en H3c afgenomen op maandag 12 maart het vijfde en zesde lesuur, in klas A3a gedurende het zevende uur op vrijdag 16 maart. De 'hard-op-denk'-protocollen zijn afgenomen in klas A3a op vrijdag 13 en 20 april en in klas H3a en H3c op maandag 7 mei.

Het Coornhert Gymnasium

Op het Coornhert Gymnasium is in een derde klas op vrijdag 27 april de gea-toets afgenomen. Op vrijdag 11 mei vonden de 'hard-op-denk'-protocollen plaats.

Motivatie

Bij de uitvoering van het kwantitatieve gedeelte van de toets zou zich een probleem kunnen voordoen betreffende de motivatie van leerlingen aangezien het hier een Low-Stake toets betreft. Dit houdt in dat er bij het maken van de toets weinig op het spel staat voor de desbetreffende leerling, de uitslag telt immers niet mee voor het rapport of de overgang naar het volgende leerjaar. Als een leerling zich niet inzet voor het maken van de toets zullen de resultaten eronder lijden waardoor de uitslag van de toets de capaciteiten van de leerling onderschat. Het gevaar bestaat dat de variantie in toetsscores tussen leerlingen eerder veroorzaakt wordt door de motivatie waarmee ze de toets hebben afgelegd dan verschillen in wat zij daadwerkelijk weten en kunnen, wat de CIV zal doen toenemen (Wise, S.L., D.S. Bhola & Sheng-Ta Yang, 2006).

Belangrijk is dus dat leerlingen de toets serieus nemen ook al heeft het geen directe consequenties voor henzelf. In het *artikel Low Examinee effort in Low Stakes Assessments : Problems and Potential Solutions* (Wise, S.L.& C.E. de Mars, 2005) worden een aantal maatregelen genoemd die men kan nemen om de motivatie van leerlingen op peil te houden. Een aantal genoemde mogelijke oplossingen zijn: verhoog het belang van de toets, zorg voor een beloning bij een gewenste studiehouding en zorg voor feedback. Het verhogen van de waarde van de toets wordt hopelijk

bereikt door leerlingen voorafgaand aan de toets bewust te maken van het belang van het onderzoek en de mogelijkheid om iets bij te dragen aan de wetenschap. Daarnaast zal hen duidelijk verzocht worden de toets serieus èn individueel af te leggen en wordt hen een traktatie beloofd wanneer zij meewerken. Verder wordt benadrukt dat zij door het maken van de toets een goed overzicht zullen krijgen van hun eigen kennis (en kennishiaten) doordat op het einde van de toets de eigen resultaten uitgebreid zullen worden weergegeven. Tussendoor zal na elke beantwoorde vraag een uitleg verschijnen of de vraag juist of onjuist beantwoord is waardoor de leerling uitgebreid van feedback wordt voorzien.

5 Resultaten

In dit hoofdstuk worden de resultaten van het onderzoek gepresenteerd. Paragraaf 5.1 geeft de resultaten weer van het kwantitatieve onderzoek en laat zien hoe op basis van deze resultaten leerlingen voor het kwalitatieve onderzoek zijn geselecteerd. De resultaten van het kwalitatieve onderzoek worden in paragraaf 5.2 besproken per toetsvraag. Hieruit zal blijken met behulp van welke geografische vaardigheden de leerlingen tot een antwoord op de vraag komen. In paragraaf 5.3 worden de uitslagen van de 'hard-op-denk'-protocollen verder geanalyseerd ten opzichte van elkaar en de rest van de toets.

5.1 Resultaten uit kwantitatief onderzoek

Deze paragraaf bespreekt de resultaten uit het kwalitatieve onderzoek en hoe op basis van deze resultaten leerlingen zijn geselecteerd voor de 'hard-op-denk'-protocollen. Allereerst wordt behandeld hoe de toetsafname verlopen is en of er rekening gehouden dient te worden met bepaalde factoren die van invloed zijn geweest tijdens de toetsafname. Vervolgens wordt behandeld hoe de selectie van leerlingen voor de 'hard-op-denk'-protocollen is gemaakt.

5.1.1 Verloop van de toetsafname

De Gea-toets werd voor het eerst afgenomen maandag 12 maart in klas H3a en H3c van de KSG, de leerlingen werden allemaal dezelfde 35 vragen voorgelegd op niveau 3. Behalve dat een aantal computers van het computerlokaal een paar keer zijn vastgelopen is de toetsafname goed verlopen. Toch heeft zich een probleem voorgedaan dat de betrouwbaarheid van de bevindingen zou kunnen ondermijnen, tijdens de toetsafname bleek dat de gevolgen van de 'Low Stake'-mentaliteit, zoals die is beschreven in paragraaf 4.5, van grotere invloed is dan vooraf werd aangenomen. Aangezien de leerlingen alle vragen in dezelfde volgorde werden voorgelegd en deze van directe feedback werden voorzien kon niet voorkomen worden dat leerlingen de antwoorden van de toets aan elkaar doorgaven. Bij de toetsafname in klas A3a van de KSG op vrijdag 16 maart is de toets op twee punten aangepast om samenwerking te ontmoedigen: de directe feedback is eruit gehaald (de leerling krijgt pas na de beantwoording van de laatste vraag te zien welke vragen ze fout hebben beantwoord) en de vragen kwamen bij iedere leerling in een andere volgorde aan bod. De afname van de Gea-toets is vervolgens in A3a zeer goed verlopen, alle leerlingen zijn zelfstandig met de toets aan de slag gegaan en deze resultaten kunnen dus ook gebruikt worden om te bepalen met welke leerlingen uit die klas het 'hard-op-denk'-protocol kan worden afgenomen. Vanaf dat moment is besloten om ook bij kwantitatieve afname in de andere klassen de vragen voor elke leerling in een andere volgorde te doen evenals enkel feedback op het einde te geven. Vervolgens is er nog een kwantitatieve afname geweest in groep 8 van de Vondelschool in Amersfoort op 16 april bestaande uit 32 vragen op niveau 2. Op 27 april was er nogmaals een afname in de derde klas van het Coornhert Gymnasium in Gouda op niveau 3 en op 7 april een afname van twee groep 8 klassen van de Koningin Wilhelmina School in Apeldoorn op niveau 2.

5.1.2 Betrouwbaarheid gegevens

Zoals in 5.1.1 al naar voren is gekomen is de gea-toets onder wisselende omstandigheden door de leerlingen afgelegd wat de CIV van het onderzoek mogelijk heeft vergroot. Een bewijs daarvan kan worden gevonden in de Item Response Time. Er waren leerlingen bij die er gemiddeld 3 seconde per opgave over hebben gedaan, dit is onvoldoende tijd om de gehele opgave en de antwoordmogelijkheden te lezen en een overwogen keuze te maken. Daar het niet precies bekend is in welke mate leerlingen uit klas H3a en H3c antwoorden van elkaar hebben overgenomen, hoewel het zeker is dat er op grote schaal overleg tussen leerlingen heeft plaatsgevonden, is dan ook besloten de keuze voor leerlingen uit deze klassen voor het kwalitatieve onderzoek niet te baseren op basis van de resultaten van de toetsafname van de Gea-toets maar op de rapportcijfers voor het vak aardrijkskunde.

Verder kan de betrouwbaarheid van de resultaten van de kwantitatieve afname van de gea-toets zijn aangetast omdat bij sommige leerlingen de computer niet goed werkte.
Aangezien de kwalitatieve afname enkel diende als selectiemiddel voor de 'hard-op-denk'-protocollen heeft dit geen verstrekkende gevolgen voor de rest van het onderzoek. Bovendien worden de leerlingen niet geselecteerd op basis van hun absolute scores, maar op hun score ten opzichte van de rest van de klas. Wanneer een bepaalde vraag niet naar behoren functioneert, geldt dat voor de uitslag van iedere leerling waardoor er in de rangorde van de scores niets verandert.

5.1.3 Uitslag van de gea-toets en motivatie keuze voor leerlingen.
De keuze voor de leerlingen voor het 'hard-op-denk'-protocollen is gebaseerd op hun prestaties (op een schaal van 1 tot en met 10) tijdens de afnamen van de gea-toets. In iedere klas zijn de 'hard-op-denk'-protocollen afgenomen bij drie leerlingen: de leerling met de hoogste score, de leerling met de laagste score en de leerling met een score gelijk aan of rondom het klasse gemiddelde. In de klassen H3a en H3c hebben rapportcijfers een belangrijke rol gespeeld bij het selecteren van leerlingen. In totaal zijn 24 leerlingen geselecteerd, een overzicht van de kenmerken van deze leerlingen is in onderstaande tabel te vinden. De leerlingen uit klas H3a en H3c zijn geselecteerd op basis van rapportcijfers die vanwege privacy redenen niet verder bekend worden gemaakt. Aangezien de leerlingen van de KWS doorlopend achter de computer hebben plaatsgenomen is van deze leerlingen de exacte klas (groep 8 A, B of C) onbekend. Daardoor is de selectie van leerlingen van de KWS voor de 'hard-op-denk'-protocollen niet per klas maar vanuit de totale lijst van de scores genomen.

Tabel 6: overzicht van de kenmerken van de deelnemende leerlingen aan de 'hard-op-denk'-protocollen.

Code leerling	**geslacht**	**School (en eventueel klas)**	**Score Gea-toets/rapport**	**Label score**
B1	Jongen	Vondelschool 8a	8,3	Hoog
B2	Jongen	KWS	7,8	Hoog
B3	Meisje	KWS	7,8	Hoog
B4	Jongen	KWS	8,3	Hoog
B5	Jongen	Vondelschool 8a	6,1	Gemiddeld
B6	Jongen	KWS	6,7	Gemiddeld
B7	Meisje	KWS	6,4	Gemiddeld
B8	Meisje	KWS	6,4	Gemiddeld
B9	Meisje	Vondelschool 8a	4,7	Laag
B10	Jongen	KWS	5,6	Laag
B11	Jongen	KWS	5,3	Laag
B12	meisje	KWS	4,9	laag
M1	Meisje	KSG Atheneum 3	7,2	Hoog
M2	Meisje	KSG Havo 3a	Hoog	Hoog
M3	Meisje	KSG Havo 3c	Hoog	Hoog
M4	Jongen	Coornhert Gymnasium 3	8,5	Hoog
M5	Meisje	KSG Atheneum 3	5,4	Gemiddeld
M6	Jongen	KSG Havo 3c	Gemiddeld	Gemiddeld
M7	Meisje	KSG Havo 3a	Gemiddeld	Gemiddeld
M8	Jongen	Coornhert Gymnasium 3	5,4	Gemiddeld
M9	Jongen	KSG Atheneum 3	2,8	Laag
M10	Meisje	KSG Havo 3a	Laag	Laag
M11	Meisje	KSG Havo 3c	Laag	Laag
M12	Meisje	Coornhert Gymnasium 3	3,8	laag

5.2 Resultaten uit kwalitatief onderzoek

In deze paragraaf worden de resultaten van de 'hard-op-denk'-protocollen per vraag behandeld. Er wordt een vergelijking gemaakt tussen de vaardigheden en kennis die de itemconstructoren voor ogen hadden en de vaardigheden en kennis waar leerlingen daadwerkelijk beroep op doen bij het beredeneren van hun antwoord. Daarnaast wordt, in het geval van opvallende redeneringen, de motivatie waarmee leerlingen voor een bepaald antwoord kiezen met elkaar vergeleken. Bij elke vraag uit het 'hard-op-denk-'protocol staat in de tabel weergegeven hoeveel leerlingen de vraag juist beantwoord hadden, op welke geografische vaardigheden en kennis zij beroep hebben gedaan en of ze het gebied op de kaart hebben kunnen aanwijzen. In 5.2.1 worden de resultaten weergegeven van de opgaven op niveau 2 afgenomen in groep 8 en in 5.2.2 staan de resultaten van de opgaven op niveau 3 overeenkomstig met het niveau van de derde klas voortgezet onderwijs. Vanwege de privacy van de leerlingen worden zij niet met hun naam genoemd, maar is hen codes toebedeeld. B staat voor basisschoolscholier (de leerlingen uit groep 8 dus) en M staat voor middelbare scholier (leerlingen uit de derde klas). Overigens is besloten de vaardigheid begrijpend lezen uit de resultaten te halen. Tijdens de afname van de 'hard-op-denk'-protocollen is bij iedere vraag bijgehouden of de leerling de vraag zodanig goed heeft gelezen dat er geen onduidelijkheid kan bestaan over hetgeen aan hem gevraagd wordt. Alle 24 aan de protocollen deelnemende leerlingen hebben de vraag goed kunnen lezen en kregen ieder het maximaal aantal punten op dit onderdeel. Doordat iedereen dezelfde maximale score had levert het verder geen interessante bevindingen op dit te analyseren. Om inzicht te verschaffen op welke manier de vaardigheden van de 'hard-op-denk'-protocollen zijn beoordeeld zijn er bij sommige vragen kadertjes toegevoegd die de beredenering van de leerling laten zien.

5.2.1 Resultaten vragen niveau 2 (groep 8 lagere school).

- Vraag 1: Brunico (Italië)

De eerste vraag die de leerlingen werd gesteld was: 'Een van de 5 foto's is niet in Italië genomen. Welke?' (afbeelding 1). De eerste afbeelding is een kaart van Italië om de leerling een beetje op weg te helpen en vervolgens krijgt de leerling 5 verschillende foto's te zien. Een kaartje als hulpmiddel wordt bij dezelfde soort vergelijkende foto-vragen op niveau 3 niet meer afgebeeld, dan wordt er van de leerling voldoende topografische kennis verwacht om het land te kunnen plaatsen en heeft de leerling keuze uit zes foto's.

Afbeelding 1: vraag 1 van 'hard-op-denk'-protocol.

De leerling heeft de volgende vijf opties:

Bron 2:

Bron 3:

Bron 4:

Bron 5:

Bron 6:

De leerlingen dienen hierbij de woestijn te herkennen als klimaat dat in Italië niet voorkomt. Om deze vraag te beantwoorden zal een leerling gebruik maken van de vaardigheden foto-interpretatie, kaart- en grafieklezen, kennis van schaalniveaus en het thematisch kaartbeeld (met name die van klimaten). Van de twaalf leerlingen hebben twaalf de vraag juist beantwoord. Daarmee is deze vraag een van de best beantwoorde vragen van het 'hard-op-denk-protocol' op niveau 2.

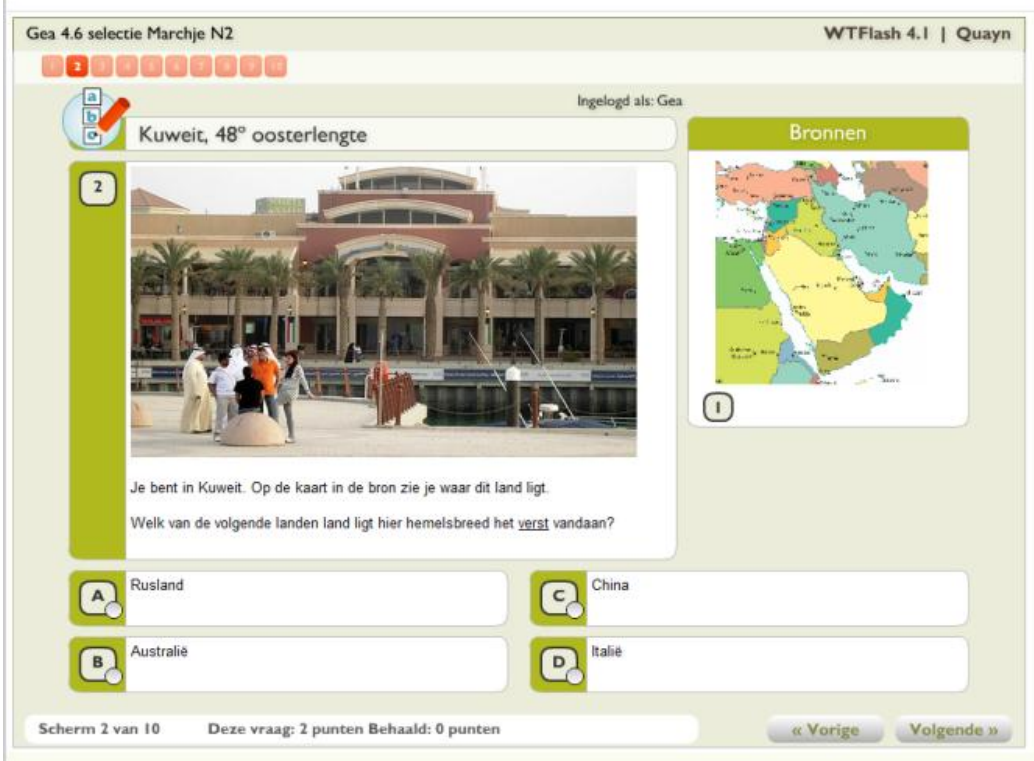
Tabel 7: Overzicht uitkomsten eerste vraag 'hard-op-denk'-protocol

Code leerling	**Score Gea-toets**	**Geografische kennis: FI**	**KL/GR**	**KS**	**ThK**	**Antwoord Juist of onjuist**	**Gebied aanwijzen op de kaart**
B1	Hoog	2	2	2	2	Juist	precies
B2	Hoog	2	2	2	2	Juist	precies
B3	Hoog	2	2	2	2	Juist	precies
B4	Hoog	2	2	2	2	Juist	precies
B5	Gemiddeld	2	2	2	2	Juist	precies
B6	Gemiddeld	2	2	2	2	Juist	precies
B7	Gemiddeld	2	2	2	2	Juist	precies
B8	Gemiddeld	2	2	2	2	Juist	precies
B9	Laag	2	2	2	2	Juist	Precies
B10	Laag	2	2	1	2	Juist	Precies
B11	Laag	2	2	2	2	Juist	Precies
B12	Laag	2	2	2	2	Juist	Precies

- Vraag 2: Kuweit

De tweede vraag die aan bod kwam bij de ‘hard-op-denk’protocollen is: ‘Je bent in Kuweit. Welke van de volgende landen ligt hier hemelsbreed het verst vandaan?’ (afbeelding 2). Bij deze vraag zit een kaartje van het gebied die de leerling kan uitvergroten, maar de vier antwoordmogelijkheden (Rusland, China, Australië en Italië) liggen allen buiten het kaartje.

Afbeelding 2: vraag 2 van ‘hard-op-denk’-protocol.



De leerling dient hierbij zijn topografische ‘mental map’ te kunnen aanspreken en de afstanden in te schatten en te vergelijken. Om deze vraag te beantwoorden zal een leerling gebruik maken van de vaardigheden kaart-en grafieklezen, kennis van schaalniveau’s en het topografisch kaartbeeld. Van de twaalf leerlingen hebben acht de vraag juist beantwoord. Daarmee is deze vraag een van de slechts beantwoorde vragen van het ‘hard-op-denk-protocol’ op niveau 2.

Tabel 8: Overzicht uitkomsten tweede vraag ‘hard-op-denk’-protocol

Code leerling	**Score Gea-toets**	**Geografische kennis: KL/GR**	**KS**	**ToK**	**Antwoord Juist of onjuist**	**Gebied aanwijzen op de kaart**
B1	Hoog	2	2	2	Juist	precies
B2	Hoog	2	2	2	Juist	precies
B3	Hoog	2	2	2	Juist	precies
B4	Hoog	2	2	2	Juist	bovenliggend
B5	gemiddeld	2	2	2	Juist	bovenliggend
B6	gemiddeld	2	2	2	Juist	precies
B7	gemiddeld	1	0	0	Onjuist	fout
B8	gemiddeld	2	0	1	Onjuist	bovenliggend
B9	Laag	2	2	2	Juist	bovenliggend

Code leerling	Score Gea-toets	Geografische kennis: KL/GR	KS	ToK	Antwoord Juist of onjuist	Gebied aanwijzen op de kaart
B10	Laag	2	0	0	Onjuist	Precies
B11	Laag	0	0	0	Onjuist	fout
B12	Laag	2	2	2	Juist	precies

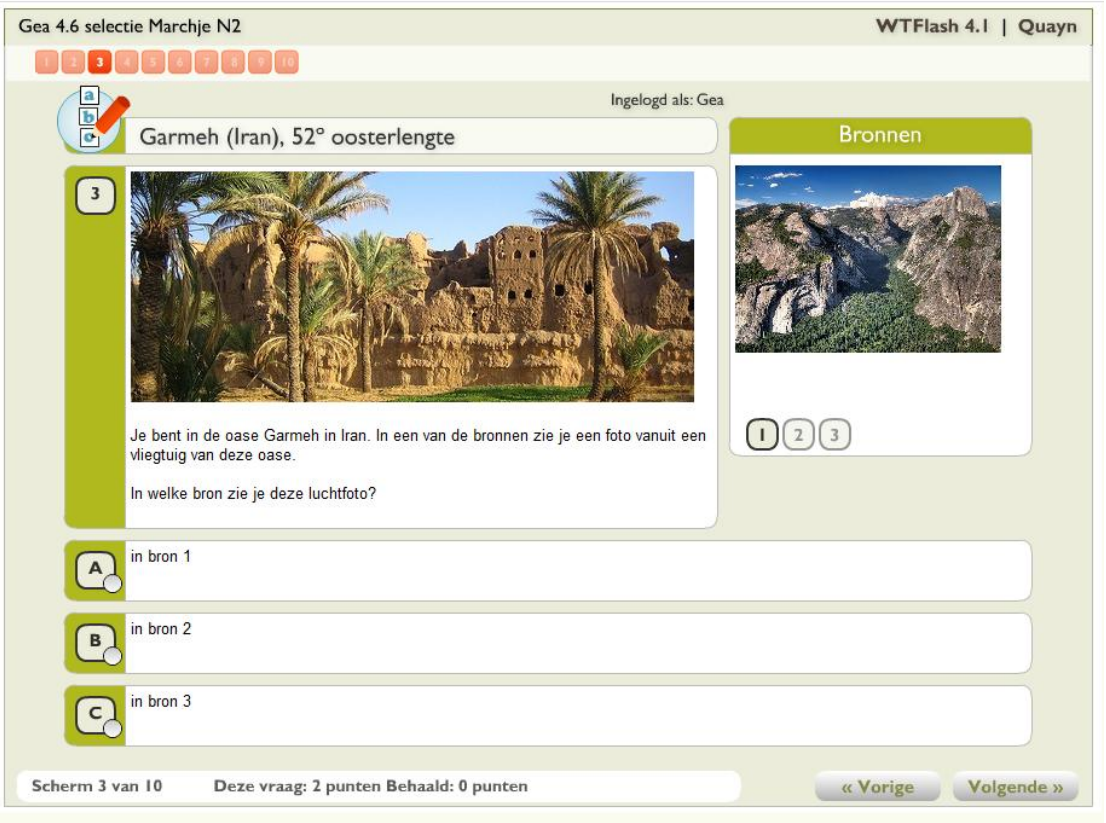
Kader 1: Antwoord van een leerling bij de tweede vraag over Kuweit

> Antwoord van leerling B11 bij 2:
>
> Leerling '.....Australië en Italië zijn het sowieso niet. Dus ik moet gokken tussen Rusland en China'
> Interviewer 'Hoezo weet je zeker dat het niet Australië en Italië zijn?'
> L 'Australië ligt sowieso niet naast Kuweit, en Italië ligt in Europa. Ik ga denk ik voor Rusland'.

- Vraag 3: Garmeh (Iran)

De derde vraag die aan bod kwam bij de 'hard-op-denk'protocollen is: 'Je bent in de oase Garmeh in Iran. In welke bron zie je een luchtfoto van deze oase?'(afbeelding 3). Bij deze vraag zitten drie afbeelding van een groen gebied die de leerling kan uitvergroten.

Afbeelding 3: vraag 3 van 'hard-op-denk'-protocol.



Wanneer de leerling op de 'Bronnen' klikt verschijnen de volgende afbeeldingen (bron 1 staat al weergegeven in de afbeelding).

Bron 2:

Bron 3:

De leerlingen zullen hierbij bron 2 herkennen als juiste antwoord aangezien het hier duidelijk gaat om een klein groen gebied te midden van de woestijn (zoals de leerling in kader 2). Om deze vraag te beantwoorden zal een leerling gebruik maken van de vaardigheid begripskennis aangezien hier impliciet wordt gevraagd wat een oase is. Van de twaalf leerlingen hebben tien de vraag juist beantwoord.

Tabel 9: Overzicht uitkomsten derde vraag 'hard-op-denk'-protocol

Code leerling	**Score Gea-toets**	**Geografische kennis:FI**	**KS**	**BK**	**Antwoord Juist of onjuist**	**Verder voorkomend?**	**Gebied aanwijzen op de kaart**
B1	Hoog	2	2	2	Juist	volledig	Precies
B2	Hoog	2	2	2	Juist	Niet volledig	Precies
B3	Hoog	2	2	1	Juist	fout	Precies
B4	Hoog	2	2	2	Juist	volledig	Precies
B5	gemiddeld	2	2	2	Juist	Niet volledig	Precies
B6	gemiddeld	2	2	2	Juist	-	Bovenliggend
B7	gemiddeld	2	2	1	Onjuist	-	Precies
B8	gemiddeld	2	2	1	Juist	-	Bovenliggend
B9	Laag	2	2	0	Juist	-	Bovenliggend
B10	Laag	2	1	2	Juist	-	Fout
B11	Laag	2	0	0	Onjuist	-	Fout
B12	Laag	2	2	1	Juist	-	Fout

Kader 2: Antwoord van een leerling bij de tweede vraag over Garmeh, Iran.

Antwoord van leerling B5 bij 3:

L '....Bron twee.'
I 'En hoe weet je dat zo zeker?'
L 'Nou, een oase is een plek waar bomen en water is in een woestijn. En dit is de enige foto waarop je dat ziet'.

- Vraag 4: Bali

De vierde vraag die bij de 'hard-op-denk'-protocollen naar voren komt is: 'Op de kaart van Bali staat Gunung Agung, Wat is dit waarschijnlijk?' (afbeelding 4). De leerling kan kiezen uit vier antwoorden en de kaart uitvergroten door daarop te klikken.

Afbeelding 4: vraag 4 van ‘hard-op-denk’-protocol.

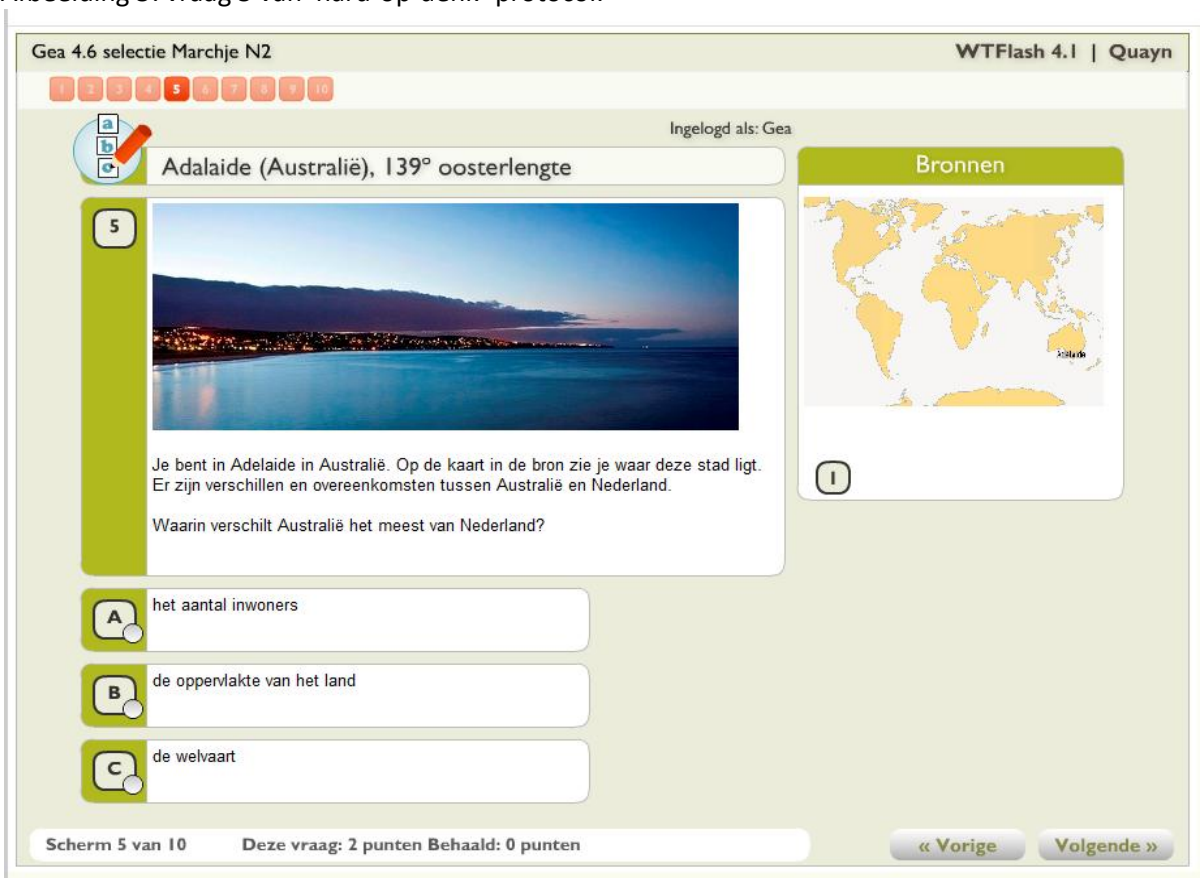
De leerling zal antwoord A herkennen als het juiste antwoord aangezien ‘Gunung Agung’ op een plek op de kaart staat waar het hoogste punt van het eiland zich bevindt. Om deze vraag te beantwoorden zal een leerling gebruik maken van de vaardigheden kaart-en grafieklezen. Van de twaalf leerlingen hebben twaalf de vraag juist beantwoord. Daarmee is deze vraag een van de best beantwoorde vragen van het ‘hard-op-denk-protocol’ op niveau 2.

Tabel 10: Overzicht uitkomsten vierde vraag ‘hard-op-denk’-protocol

Code leerling	**Score Gea-toets**	**Geografische kennis: KL/GR**	**Antwoord Juist of onjuist**	**Gebied aanwijzen op de kaart**
B1	Hoog	2	Juist	Bovenliggend
B2	Hoog	2	Juist	Fout
B3	Hoog	2	Juist	Bovenliggend
B4	Hoog	2	Juist	Bovenliggend
B5	Gemiddeld	2	Juist	Fout
B6	Gemiddeld	2	Juist	Bovenliggend
B7	Gemiddeld	2	Juist	Precies
B8	Gemiddeld	2	Juist	Fout
B9	Laag	2	Juist	Precies
B10	Laag	2	Juist	Bovenliggend
B11	Laag	2	Juist	Bovenliggend
B12	Laag	2	Juist	Bovenliggend

- Vraag 5: Adalaide (Australië)
De vijfde vraag die de leerlingen wordt voorgelegd is: ‘Waarin verschilt Australië het meeste van Nederland?’ (afbeelding 5). De leerling kan tussen drie antwoorden kiezen.

Afbeelding 5: vraag 5 van ‘hard-op-denk’-protocol.



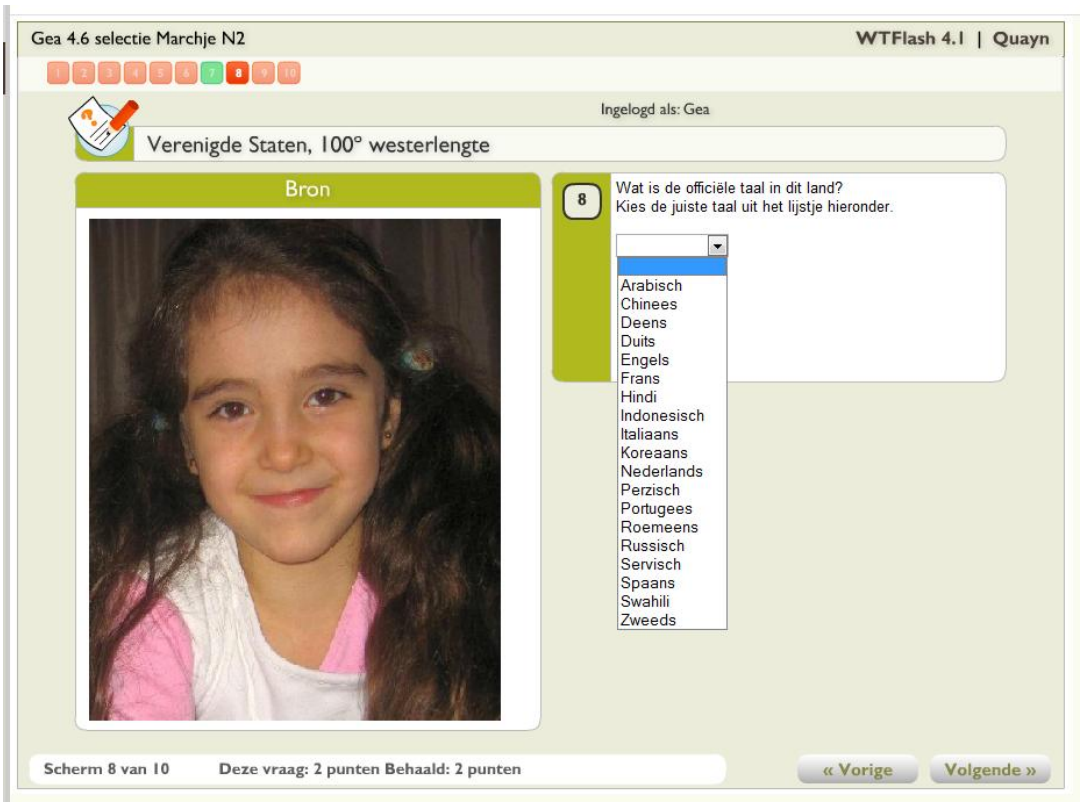
De leerlingen zal hierbij ‘de oppervlakte van het land’ herkennen als juiste antwoord. Om deze vraag te beantwoorden zal een leerling gebruik maken van de vaardigheden vergelijking van gebieden (Australië heeft een veel groter oppervlakte dan Nederland) en feitenkennis (‘Australië is een groot land’, ‘Australië is dunbevolkt’). Van de twaalf leerlingen hebben twaalf de vraag juist beantwoord.

Tabel 11: Overzicht uitkomsten vijfde vraag ‘hard-op-denk’-protocol

Code leerling	**Score Gea-toets**	**Geografische kennis: VG**	**FK**	**Antwoord Juist of onjuist**	**Gebied aanwijzen op de kaart**
B1	Hoog	2	2	Juist	precies
B2	Hoog	2	2	Juist	precies
B3	Hoog	2	2	Juist	precies
B4	Hoog	2	2	Juist	precies
B5	Gemiddeld	2	2	Juist	precies
B6	Gemiddeld	2	2	Juist	precies
B7	Gemiddeld	2	2	Juist	precies
B8	Gemiddeld	2	1	Juist	precies
B9	Laag	2	2	Juist	precies
B10	Laag	2	2	Juist	precies
B11	Laag	2	2	Juist	precies
B12	Laag	2	2	Juist	precies

- Vraag 6: Verenigde Staten
De laatste vraag die de leerlingen wordt gesteld is: ‘Wat is de officiële taal van de Verenigde Staten’? De leerling maakt een keuze uit 19 potentiële antwoorden.

Afbeelding 6: vraag 6 van ‘hard-op-denk’-protocol.



De leerlingen zal bij deze opgave ‘Engels’ aanwijzen als het juiste antwoord. Om deze vraag te beantwoorden zal een leerling gebruik maken van de vaardigheid feitenkennis over de verdeling van talen over de wereld. Waarschijnlijk zullen leerlingen vanwege films en televisieseries dermate bekend zijn met de Amerikaanse cultuur dat ze al kunnen bedenken dat daar Engels wordt gesproken voordat zij hun thematisch kaartbeeld van de talen op de wereld hoeven aan te spreken. Van de twaalf leerlingen hebben twaalf de vraag juist beantwoord.

Tabel 12: Overzicht uitkomsten zesde vraag ‘hard-op-denk’-protocol

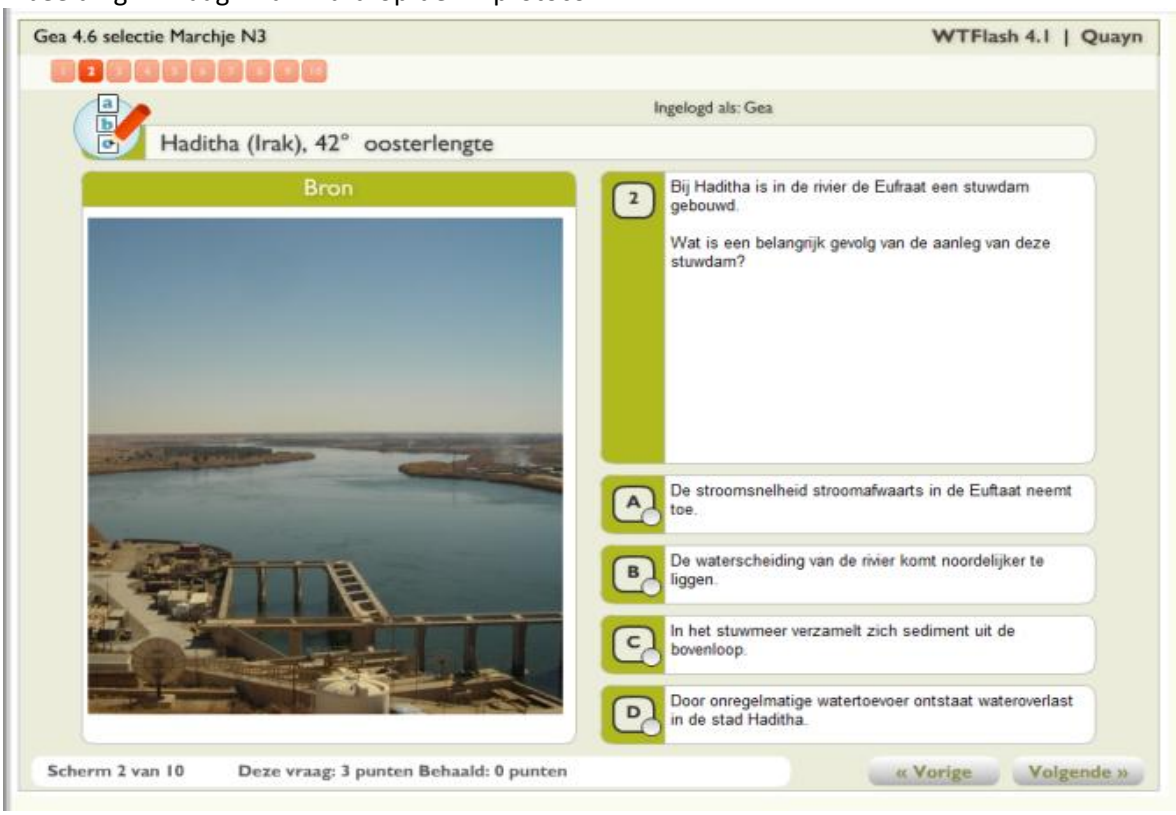
Code leerling	**Score Gea-toets**	**Geografische kennis: FK**	**Antwoord Juist of onjuist**	**Verder voorkomend?**	**Gebied aanwijzen op de kaart**
B1	Hoog	2	Juist	volledig	Precies
B2	Hoog	2	Juist	Niet volledig	Precies
B3	Hoog	2	Juist	Niet volledig	Precies
B4	Hoog	2	Juist	Niet volledig	Precies
B5	Gemiddeld	2	Juist	Niet volledig	Precies
B6	Gemiddeld	2	Juist	Niet volledig	Precies
B7	Gemiddeld	2	Juist	Niet volledig	Precies
B8	Gemiddeld	2	Juist	Niet volledig	Precies
B9	Laag	2	Juist	Niet volledig	Precies
B10	Laag	1	Onjuist	-	Precies

Code leerling	Score Gea-toets	Geografische kennis: FK	Antwoord Juist of onjuist	Verder voorkomend?	Gebied aanwijzen op de kaart
B11	Laag	2	Juist	Niet volledig	Precies
B12	Laag	2	Juist	-	Bovenliggend

5.2.2 Resultaten vragen niveau 3 (derde klas voortgezet onderwijs).

- Vraag 1: Haditha (Irak)
De eerste vraag die de leerlingen wordt gesteld is: 'Wat is een belangrijk gevolg van de aanleg van een stuwdam'? De leerling krijgt vier antwoordopties (afbeelding 7).

Afbeelding 7: vraag 1 van 'hard-op-denk'-protocol.



De leerlingen dienen optie C als het juiste antwoord te herkennen. Om deze vraag te beantwoorden zal een leerling waarschijnlijk gebruik maken van de vaardigheden begripskennis ('wat is de functie van een stuwdam ?') en kennis van principes ('naarmate het stuwmeer groter is verzamelt zich meer sediment in de bovenloop'). Van de twaalf leerlingen hebben vijf de vraag juist beantwoord. Opvallend was dat veel leerlingen niet wisten wat een waterscheiding is en dat de functie van een stuwdam ook niet bij iedereen bekend is (zoals blijkt uit kader 3).

Tabel 13: Overzicht uitkomsten eerste vraag 'hard-op-denk'-protocol

Code leerling	Score Gea-toets of rapportcijfer	Geografische kennis: BK	PR	Antwoord juist of onjuist	Gebied aanwijzen op de kaart
M1	Hoog	1	1	juist	Bovenliggend
M2	Hoog	1	1	Onjuist	Fout
M3	Hoog	1	1	juist	Bovenliggend
M4	Hoog	2	2	Juist	Bovenliggend
M5	gemiddeld	2	1	juist	Precies

Code leerling	Score Gea-toets of rapportcijfer	Geografische kennis: BK	PR	Antwoord Juist of onjuist	Gebied aanwijzen op de kaart
M6	gemiddeld	0	0	onjuist	Bovenliggend
M7	gemiddeld	1	0	Onjuist	Bovenliggend
M8	gemiddeld	1	1	Onjuist	Bovenliggend
M9	Laag	0	0	Onjuist	Fout
M10	Laag	1	0	Onjuist	Bovenliggend
M11	Laag	0	0	Juist	Fout
M12	Laag	1	1	Onjuist	Bovenliggend

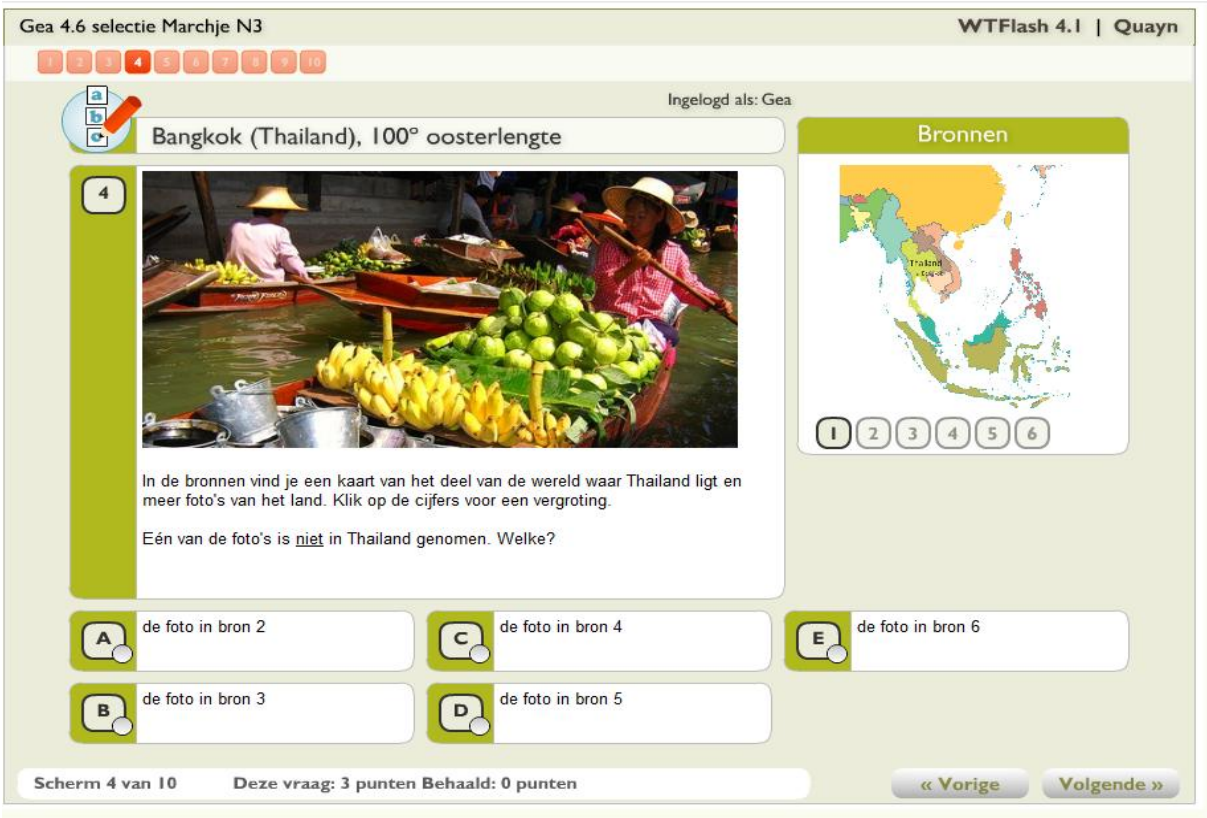
Kader 3: Antwoord van een leerling bij de eerste vraag over Haditha, Irak:

> Antwoord van leerling M9 bij 1:
>
> L 'uhhh, hier ben ik niet echt heel goed in. Ik weet niet echt wat ze met B bedoelen. De waterscheiding van de rivier komt noordelijker te liggen? Maar als ze bedoelen dat die stuwdam zelf de waterscheiding is dan..is dat misschien logisch. Bij C, ik weet niet echt wat sediment is....dus daar kan ik niet veel over zeggen. A klinkt niet, weet niet, niet heel erg logisch. Ja, ik twijfel tussen B en D, maar ik denk....ik ga denk ik voor B.

- Vraag 2: Bangkok

De tweede vraag die de leerlingen gesteld werd tijdens het 'hard-op-denk'-protocol was: 'Een van de vijf foto's is niet in Thailand genomen, welke?'. Er wordt de leerling een kaartje van de topografische ligging van Thailand gegeven en vijf verschillende foto's getoond.

Afbeelding 8: vraag 2 van 'hard-op-denk'-protocol.



Wanneer de leerling op de ‘Bronnen’ klikt verschijnen de volgende afbeeldingen.

Bron 2:

Bron 3:

Bron 4:

Bron 5:

Bron 6:

De leerlingen dienen hierbij de afbeelding in bron 3 van de Sultan Ahmet-moskee (internationaal bekend als de ‘blauwe moskee’) te herkennen als de foto die niet in Thailand gemaakt kan zijn. Wanneer een leerling weet dat de meerderheid van de Thaise bevolking boeddhistisch is zal het voor de leerling niet moeilijk zijn tot het juiste antwoord te komen. De verwachting is dat bron 5, waarop witte flatgebouwen zijn afgebeeld, sommige leerlingen aan het twijfelen zullen brengen aangezien voor sommigen dit onverenigbaar zou kunnen zijn met het beeld dat zij hebben van zuidoost Azië, namelijk als dat van een arme regio (zoals in kader 4). Vaardige leerlingen zullen hier waarschijnlijk snel door heen zien, want ook in arme landen heb je stadscentra met flats (grote verschillen tussen arm en rijk is juist kenmerkend voor armere landen). Om deze vraag te beantwoorden zal een leerling gebruik maken van de vaardigheden foto-interpretatie, kaart- en grafieklezen, kennis van schaalniveau’s en het thematisch kaartbeeld (met name die van religies). Van de twaalf leerlingen hebben zes de vraag juist beantwoord.

Tabel 14: Overzicht uitkomsten tweede vraag ‘hard-op-denk’-protocol

Code leerling	**Score Gea-toets of rapportcijfer**	**Geografische kennis: FI**	**KL/GR**	**KS**	**Thk**	**Antwoord juist of onjuist**	**Gebied aanwijzen op de kaart**
M1	Hoog	2	2	2	2	juist	Precies
M2	Hoog	2	2	2	2	Juist	Bovenliggend
M3	Hoog	2	2	2	2	juist	Bovenliggend
M4	Hoog	2	2	2	2	Juist	precies

Code leerling	Score Gea-toets of rapportcijfer	Geografische kennis: FI	KL/GR	KS	ThK	Antwoord Juist of onjuist	Gebied aanwijzen op de kaart
M5	gemiddeld	2	2	2	0	onjuist	bovenliggend
M6	gemiddeld	2	2	2	2	Juist	Precies
M7	gemiddeld	2	2	1	1	Onjuist	Bovenliggend
M8	gemiddeld	2	2	2	1	Onjuist	precies
M9	Laag	2	2	2	0	onjuist	fout
M10	Laag	2	2	1	0	Onjuist	Bovenliggend
M11	Laag	2	2	1	0	Onjuist	Fout
M12	Laag	2	2	1	1	Juist	precies

Kader 4: Antwoord van een leerling bij de tweede vraag over Thailand

> Antwoord van leerling M10 bij 2:
>
> L 'Ja...dan denk ik vijf, bron vijf. Want als ik aan Thailand denk dan denk ik niet aan hoge gebouwen. Zij wonen niet in flats. Dus dan is het vijf'.

- Vraag 3: Panglao (Filipijnen)

De derde vraag die aan de leerlingen gesteld wordt is: 'Waar hangt de spreiding van tuberculose in Zuidoost-Azië voornamelijk mee samen?' De leerling krijgt de keuze uit vier antwoorden en kan een kaartje met 'aantal tuberculosegevallen per land' uitvergroten .

Afbeelding 9: vraag 3 van 'hard-op-denk'-protocol.

Het juiste antwoord op de vraag is antwoord D, welvaart. In het kaartje staat het aantal tuberculose-gevallen weergegeven per land in Zuidoost Azië. Opvallend is dat Australië en Nieuw-Zeeland duidelijk het minste aantal tuberculosepatiënten hebben. Ook komt tuberculose in de armste landen van Zuidoost Azië meer voor dan in relatief rijke landen uit de regio. Hieruit kan een leerling opmaken dat de ruimtelijke spreiding van tuberculose samenhangt met welvaart.

Om deze vraag te beantwoorden zal een leerling gebruik maken van de vaardigheden foto-interpretatie, kaart-en grafieklezen, kennis van schaalniveau's en het thematisch kaartbeeld (met name die van welvaart).

Tabel 15: Overzicht uitkomsten derde vraag 'hard-op-denk'-protocol

Code leerling	**Score Gea-toets of rapportcijfer**	**Geografische kennis: KL/GR**	**ThK**	**Antwoord juist of onjuist**	**Verder voorkomend**	**Gebied aanwijzen op de kaart**
M1	Hoog	2	2	juist	volledig	Precies
M2	Hoog	2	2	Juist	-	Bovenliggend
M3	Hoog	2	2	Juist	volledig	precies
M4	Hoog	2	2	Juist	volledig	precies
M5	gemiddeld	2	0	Onjuist	-	Precies
M6	gemiddeld	2	2	Juist	volledig	Precies
M7	gemiddeld	2	2	Juist	-	bovenliggend
M8	gemiddeld	0	1	Onjuist	-	precies
M9	Laag	2	0	Onjuist	-	fout
M10	Laag	0	0	Onjuist	-	Bovenliggend
M11	Laag	1	0	Onjuist	-	Bovenliggend
M12	Laag	2	0	Onjuist	-	bovenliggend

Van de twaalf leerlingen hebben zes de vraag juist beantwoord. Opvallend is dat de leerlingen die de vraag verkeerd hebben beantwoord ervan overtuigd leken dat welvaart niet de reden kon zijn van de spreiding van tuberculose omdat ze zich niet konden voorstellen dat iets dat niet medisch was wel medische gevolgen kan hebben. Een aantal van hen dachten dat het aan het voedingspatroon lag en een enkeling linkte tuberculose aan het warme klimaat.

Kader 5: Antwoord van een leerling bij de derde vraag over Panglao, Filipijnen.

Antwoord van leerling M6 bij 3:

L 'Ik denk dat het gaat om welvaart, D'
I 'En waarom denk je dat?'
L 'Omdat....in Australië hebben ze wel veel welvaart, en hier zo hebben meer mensen dat en daar is minder welvaart'.
I 'Dus het ligt niet aan iets anders zoals de hoogteligging of het klimaat?'
L 'Dat denk...ik niet. Want hier en hier is het heel verschillend maar ze hebben hetzelfde klimaat.'

- Vraag 4: Rio de Janeiro (Brazilië)
De vierde vraag die de leerling werd gesteld was ‘De meeste steden in Brazilië zijn ontstaan aan de kust, wat is daarvoor een belangrijke reden?’. Bij de vraag kunnen leerlingen kiezen tussen vier antwoordmogelijkheden.

Afbeelding 10: vraag 4 van ‘hard-op-denk’-protocol.

De leerlingen dient hierbij ‘In de koloniale tijd waren de verbindingen met het moederland belangrijk’ (Optie A) als antwoord te geven. Om deze vraag te beantwoorden zal een leerling waarschijnlijk gebruik maken van de vaardigheden thematisch kaartbeeld (bijvoorbeeld van ‘de spreiding van koloniën over de wereld’, ‘verdeling van klimaat’, ‘hoogteligging’ en ‘beroepsbevolking’) en kaartlezen aangezien er in de bron een meer gedetailleerde kaart van de Braziliaanse kustlijn kan worden geopend.
Drie van de twaalf leerlingen hebben de vraag juist beantwoord. Daarmee is deze vraag een van de moeilijkste vragen van het ‘hard-op-denk-protocol’ op niveau 3.

Aan de twee leerlingen die de vraag juist hadden beantwoord en Brazilië op de kaart juist konden aanwijzen is gevraagd waar op de wereld men meer voormalig koloniale gebieden kan aantreffen waarvan de steden grotendeels aan de kust liggen. Leerling M4 heeft hierop Suriname, Indonesië en India op de kaart aangevinkt. Dit antwoord is gedeeltelijk goed gerekend want het gaat hierbij inderdaad om postkoloniale gebieden die verbindingen hadden via zeevaartroutes met het moederland, maar toch ontbraken de gebieden in Afrika en de rest van Amerika in dit antwoord. Leerling M6 gaf een completer beeld van koloniale territoria en gaf heel Zuid-Amerika, Afrika en Zuid-Oost Azië aan. Het is opvallend dat beide leerlingen geen eerste wereldlanden als koloniën hebben herkend, Australië, Nieuw-Zeeland, de Verenigde Staten en Canada werden hier niet genoemd.

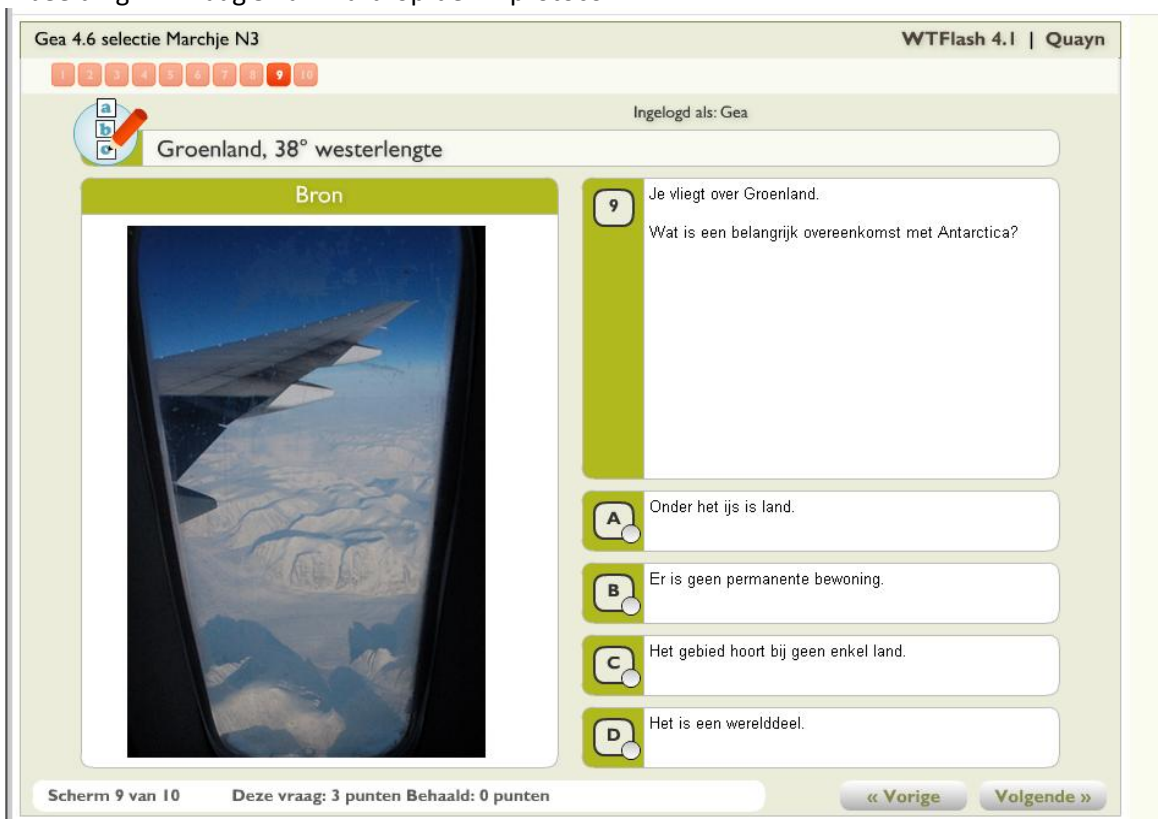
Tabel 16: Overzicht uitkomsten vierde vraag ‘hard-op-denk’-protocol

Code leerling	Score Gea-toets of rapportcijfer	Geografische kennis: KL/GR	ThK	Antwoord juist of onjuist	Verder voorkomend?	Gebied aanwijzen op de kaart
M1	Hoog	2	1	Onjuist	-	Precies
M2	Hoog	2	2	Juist	-	Bovenliggend
M3	Hoog	2	1	Onjuist	-	Precies
M4	Hoog	2	2	juist	Niet volledig	precies
M5	gemiddeld	2	1	onjuist	-	Precies
M6	gemiddeld	2	2	Juist	volledig	Precies
M7	gemiddeld	2	0	Onjuist	-	Bovenliggend
M8	gemiddeld	2	0	Onjuist	-	Precies
M9	Laag	2	0	Onjuist	-	fout
M10	Laag	1	0	Onjuist	-	precies
M11	Laag	2	0	Onjuist	-	bovenliggend
M12	Laag	2	0	Onjuist	-	precies

- Vraag 5: Groenland

De vijfde vraag die de leerlingen werd voorgelegd was: ‘Wat is een belangrijke overeenkomst tussen Groenland en Antarctica?’ (afbeelding 11). Op de vraag zijn vier antwoordalternatieven mogelijk.

Afbeelding 11: vraag 5 van ‘hard-op-denk’-protocol.



De bedoeling van de opgave is dat de leerling zijn vermogen om een vergelijking tussen twee gebieden te maken gebaseerd op informatie uit zijn Thematisch Kaartbeeld. Voor een compleet beredeneerd antwoord dient een leerling te weten dat:

1 onder Groenland land ligt en onder Antarctica ook.

2 Er permanente bewoning bestaat op Groenland.

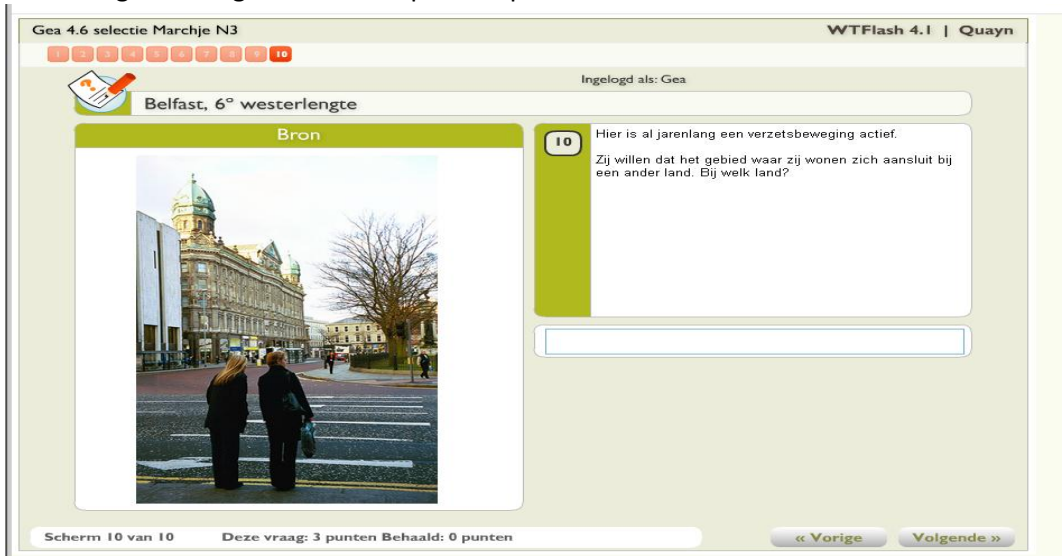
3 Dat Groenland behoort tot een ander land (Denemarken) en Antarctica niet politiek verdeeld is.
4 Dat Groenland geen werelddeel is.
Dit leidt tot het juiste antwoord, optie A.
Acht van de twaalf leerlingen hebben de vraag juist beantwoord, daarmee is deze vraag de best beantwoorde vragen van het 'hard-op-denk-protocol' op niveau 3. Veel leerlingen konden hier het juiste antwoord beredeneren door de foute antwoordalternatieven weg te strepen. Optie 'antwoord D' viel hierbij bij alle leerlingen het eerste weg. Leerlingen die wisten dat Groenland bij een ander land hoort (enkelen wisten Denemarken te noemen) konden hierdoor antwoord B en C schrappen want wanneer een gebied bij een land hoort zouden daar volgens hen ten minste ook een paar mensen moeten wonen. De leerlingen die het verkeerde antwoord gaven wisten vaak niet dat er onder het ijs van Antarctica land ligt.

Tabel 17: Overzicht uitkomsten vijfde vraag 'hard-op-denk'-protocol

Code leerling	**Score Gea-toets of rapportcijfer**	**Geografische kennis: VG**	**ThK**	**Antwoord juist of onjuist**	**Gebied aanwijzen op de kaart**
M1	Hoog	2	2	juist	Precies
M2	Hoog	1	1	Onjuist	Bovenliggend
M3	Hoog	2	2	Juist	Precies
M4	Hoog	2	2	Juist	precies
M5	Gemiddeld	2	2	Juist	Precies
M6	Gemiddeld	2	2	Juist	Precies
M7	Gemiddeld	2	2	Juist	Bovenliggend
M8	Gemiddeld	1	1	Onjuist	precies
M9	Laag	1	1	Onjuist	precies
M10	Laag	1	0	Onjuist	Precies
M11	Laag	2	2	Juist	Bovenliggend
M12	Laag	2	2	Juist	precies

- Vraag 6: Belfast (Noord-Ierland)
De laatste vraag die de leerlingen werd gesteld was: 'In Belfast is al jarenlang een verzetsbeweging actief. Zij willen dat het gebied waar zij wonen zich aansluit bij een ander land. Bij welk land?' (afbeelding 12).

Afbeelding 12: vraag 6 van 'hard-op-denk'-protocol.



De leerlingen dienen hierbij 'Ierland' als antwoord te geven. Om deze vraag te beantwoorden zal een leerling waarschijnlijk gebruik maken van de vaardigheden topografisch kaartbeeld ('Belfast ligt in Noord-Ierland') en feitenkennis ('In Noord-Ierland is de IRA actief, zij streven een Ierse staat na zonder banden met het Verenigd Koninkrijk').
Van de twaalf leerlingen hebben slechts twee leerlingen de vraag juist beantwoord. Daarmee is deze vraag de moeilijkste en slechtst beantwoorde vraag van het 'hard-op-denk-protocol' op niveau 3. Opvallend is dat veel leerlingen die de vraag juist hebben beantwoord dit grotendeels of geheel hebben gebaseerd op hun topografische kennis (zoals in kader 6). Zij weten waar Belfast ligt (Noord-Ierland) en beredeneren hieruit het juiste antwoord (Ierland) aangezien dit land hier het dichtste bij ligt. Wanneer hen wordt gevraagd naar de achterliggende reden waarom de verzetsbeweging zich wil aansluiten bij Ierland kunnen ze deze vraag vaak niet beantwoorden. Er was één leerling die een religieuze strijd noemde tussen katholieken en protestanten, maar doordat zij niet wist waar Belfast precies lag heeft ze alsnog de vraag onjuist beantwoord. Geen van de leerlingen wisten dat deze verzetsbeweging de IRA heette. De leerlingen die de vraag fout hebben beantwoord wisten zowel niets van de redenen achter het conflict (feitenkennis) als waar Belfast precies lag (topografische kennis).
Aan de leerlingen die de vraag juist hebben beantwoord is ook gevraagd of zij een gebied konden noemen waar eenzelfde soort conflict zich afspeelt. Leerling M1 kon deze vraag niet beantwoorden, leerling M4 beantwoordde de vraag met 'Al-Qaida'. Dit antwoord is gedeeltelijk goed gerekend aangezien het bij deze terroristische organisatie tevens gaat om een conflict waarbij religie centraal staat. Toch is het antwoord niet volledig aangezien verwacht werd dat de leerling een conflict zou benoemen waarin het voornamelijk om een sterk territoriale claim van een verzetsbeweging zou gaan. Vooraf werd verwacht dat de leerlingen de link tussen de IRA en de ETA zouden maken, maar aangezien geen van de leerlingen de IRA in hun antwoord hadden benoemd waren er ook geen leerlingen die een link naar de ETA wisten te maken.

Tabel 18: Overzicht uitkomsten zesde vraag 'hard-op-denk'-protocol

Code leerling	**Score Gea-toets of rapportcijfer**	**Geografische kennis: FK**	**ToK**	**Antwoord Juist of onjuist**	**Verder voorkomend?**	**Gebied aanwijzen op de kaart**
M1	Hoog	0	2	Juist	Fout	Precies
M2	Hoog	1	0	Onjuist	-	fout
M3	hoog	0	1	Onjuist	-	Bovenliggend
M4	hoog	1	2	Juist	Niet volledig	precies
M5	gemiddeld	0	1	onjuist	-	Fout
M6	gemiddeld	0	1	Onjuist	-	Bovenliggend
M7	gemiddeld	-	-	-	-	-
M8	gemiddeld	1	1	Onjuist	-	bovenliggend
M9	laag	0	0	Onjuist	-	Fout
M10	laag	-	1	-	-	Bovenliggend
M11	laag	-	-	-	-	-
M12	laag	0	1	Onjuist	-	precies

Kader 6: Antwoord van een leerling bij opgave zes over Noord-Ierland:

Antwoord van leerling M1 bij 6:

L 'Belfast, dat is in Noord-Ierland....Dus dan willen ze dat het bij Ierland komt.'
I 'En waarom geef je dit antwoord?'
L 'Ja Ierland. Daar liggen volgens mij geen andere landen aan vast, omheen. Dus lijkt het me logisch dat ze dat willen'
I 'Kun je verder nog een reden geven waarom zij dat graag zouden willen?'
L 'Misschien omdat het in Ierland niet zo goed gaat met de economie? Dus dan zou dat misschien handiger zijn?....Als ze zouden samenwerken?'

In onderstaande tabel is weergegeven wat de totale scores zijn van het 'hard-op-denk'-protocol per leerling (tabel 18).
Bij alle leerlingen is duidelijk te zien dat het totaal aantal goede antwoorden, precieze én verkeerde aanwijzingen op de kaart en de totaal score voor de vaardigheden samenhangt met de scores op de gea-toets of op het rapport. Zo gaven de leerlingen uit de derde klas met een hoge score gemiddeld 4,5 goede antwoorden op de zes vragen uit het protocol tegenover gemiddeld 2 goede antwoorden voor de leerlingen met gemiddelde scores en slechts 1 goed beantwoorde vraag bij de leerlingen die geselecteerd waren op basis van een slechte score. Bij de leerlingen in groep 8 is dit verband ook heel duidelijk te zien, maar het onderlinge verschil is wel minder groot. Dit zou kunnen doordat de landen in de vragen van het 'hard-op-denk'-protocol op niveau 2 vrij groot en herkenbaar zijn van vorm (zoals de Verenigde Staten en Australië).

Tabel 19: uitslag resultaten van de leerlingen met het 'hard-op-denk'-protocol

Code leerling	**Score Gea-toets/ rapportcijfer**	**Aantal juiste antwoorden**	**Aantal precieze aanwijzingen op de kaart**	**Aantal verkeerde aanwijzingen op de kaart**	**Totaal aantal punten**
B1	Hoog	6	5	0	28
B2	Hoog	6	5	1	28
B3	Hoog	6	5	0	27
B4	Hoog	6	4	0	28
Gemiddeld B1-4		**6**	**4,75**	**0,25**	**27,75**
B5	Gemiddeld	6	4	1	28
B6	Gemiddeld	6	4	0	28
B7	Gemiddeld	4	5	1	22
B8	Gemiddeld	5	2	1	23
Gemiddeld B5-8		**5,25**	**3,75**	**0,75**	**25,25**
B9	Laag	6	4	0	26
B10	Laag	4	4	1	21
B11	Laag	4	3	2	18
B12	Laag	6	3	1	27
Gemiddeld B9-12		**5**	**3,5**	**1**	**23**
M1	Hoog	5	5	0	23
M2	Hoog	3	0	2	21
M3	Hoog	4	3	0	22
M4	Hoog	6	5	0	27
Gemiddeld M1-4		**4,5**	**3,25**	**0,5**	**23,25**
M5	Gemiddeld	2	4	1	19
M6	Gemiddeld	4	4	0	21

Code leerling	Score Gea-toets/ rapportcijfer	Aantal juiste antwoorden	Aantal precieze aanwijzingen op de kaart	Aantal verkeerde aanwijzingen op de kaart	Totaal aantal punten
M7	Gemiddeld	2	0	1	17
M8	Gemiddeld	0	3	0	16
Gemiddeld M5-8		**2**	**2,75**	**0,5**	**18,25**
M9	Laag	0	1	5	12
M10	Laag	0	2	0	9
M11	Laag	2	0	3	12
M12	Laag	2	4	0	17
Gemiddeld M9-12		**1**	**1,75**	**0,5**	**12,5**

5.3 Analyse resultaten van de 'hard-op-denk'-protocollen

Zoals in de vorige paragraaf naar voren is gekomen zijn de vaardigheden uit het 'hard-op-denk'-protocol gekwantificeerd door een cijfer te verbinden aan de mate waarin een leerling deze beheerst (0= niet, 1=matig, 2=volledig). Op deze manier kunnen de scores voor deze vaardigheden met behulp van het computerprogramma TiaPlus worden vergeleken met de rest van de toetsscores van de leerlingen voortkomend uit de afname van de gea-toets. Hiervoor werden alle scores van de groep 8 leerlingen die deelnamen aan de kwalitatieve afname van de gea-toets én de resultaten uit de 'hard-op-denk'-protocollen ingevoerd. In tabel 20 staan de uitslagen van TiaPlus voor de 'hard-op-denk'-protocollen van de twaalf groep 8 leerlingen weergeven. Hierbij moet de kanttekening gemaakt worden dat het gebruik van dit computerprogramma nauwkeuriger en betrouwbaarder wordt naarmate er meer cases zijn. Aangezien slechts 24 leerlingen hebben deelgenomen aan de 'hard-op-denk'-protocollen, waarvan er twaalf er niveau worden vergeleken, kan uit de resultaten geen harde conclusies worden getrokken. Het programma geeft echter wel een belangrijke indicatie van het belang van verschillende vaardigheden die het geografisch besef vormen.

In de tabel wordt iedere vaardigheid die bij een vraag uit de 'hard-op-denk'-protocollen een bepaalde score oplevert als een aparte variabel ingevoerd. Zo dient een leerling bij het beantwoorden van de eerste vraag van het 'hard-op-denk'-protocol op niveau 2 een beroep te doen op de vaardigheden 'foto-interpretatie', 'Kaart lezen/Grafiek Lezen', 'Kennis van schaalniveaus' en 'Thematisch kaartbeeld'. Deze staan allemaal als aparte variabelen vermeld (tabel 20). In totaal gaat het om veertien variabelen van vaardigheden die nodig zijn om de zes vragen van de 'hard-op-denk'-protocollen juist te kunnen beantwoorden (variabel 1 tot en met 14). Na afloop van iedere vraag werd aan de leerling ook gevraagd om het land waar de vraag over ging aan te kruisen op een wereldkaart. Deze gegevens zijn ook gekwantificeerd (0=fout, 1=bovenliggend en 2=precies), en vervolgens aan de analyse toegevoegd (variabel 15 tot en met 20). Bij alle vaardigheden is het maximaal aantal te verkrijgen punten vastgesteld op 2. In de kolom daarnaast staat vermeld wat het gemiddeld aantal punten per vaardigheid is. Hoe hoger dit getal is, des te meer deze vaardigheid aanwezig is onder de leerlingen. Dit komt ook weer terug in de P-waarde die aangeeft hoeveel procent van de leerlingen over een vaardigheid beschikt.

In de rechterkolom staat de RIR-waarde weergegeven. Dit getal geeft de samenhang tussen de betreffende variabel en de resultaten uit de rest van de toets weer. Hoe hoger dit getal, des te meer de score van de variabel aansluit bij de toetsscores uit de kwantitatieve afname van de gea-toets én de andere resultaten uit de 'hard-op-denk'-protocollen. Kortom: hoe hoger de RIR-waarde, des te belangrijker is de variabel in het maken van een onderscheid tussen vaardige en minder vaardige leerlingen. Dit betekent dat een vaardigheid met een hoge RIR-waarde een essentiële vaardigheid is om tot het juiste antwoord op de vraag te komen. Aan de andere kant houdt dit ook in dat vaardigheden met een lage RIR-waarde of een RIR-waarde van 0 weinig onderscheid maken tussen leerlingen met veel of weinig geografische vaardigheden. Een variabel met een negatieve RIR-waarde geeft aan dat er een negatief verband bestaat tussen de betreffende variabel en de rest van

de toets. Bij deze variabel scoren leerlingen die een lage toetsscore hadden hoger dan de leerlingen met betere toetsscores.

Tabel 20: totaal uitslag resultaten groep 8 basisonderwijs

Nummer variabel	Code vaardigheid	Gemeten bij opgave	Maximum aantal punten	Gemiddeld aantal punten	P-waarde	RIR-waarde
1	Foto-interpretatie	1: Brunico, Italië	2	2,00	100	0
2	Kaart lezen/grafiek lezen	1: Brunico, Italië	2	2,00	100	0
3	Kennis van Schaalniveaus	1: Brunico, Italië	2	1,92	96	22
4	Thematisch Kaartbeeld	1: Brunico, Italië	2	2,00	100	0
5	Kaartlezen/ grafiek lezen	2: Kuweit	2	1,75	88	45
6	Kennis van Schaalniveaus	2: Kuweit	2	1,33	67	52
7	Topografisch kaartbeeld	2: Kuweit	2	1,42	71	53
8	Foto-interpretatie	3:Garmeh, Iran	2	2,00	100	0
9	Kennis van Schaalniveaus	3: Garmeh, Iran	2	1,75	88	54
10	Begripskennis	3: Garmeh, Iran	2	1,33	67	61
11	Kaart lezen/ Grafiek lezen	4: Bali	2	2,00	100	0
12	Vergelijking van Gebieden	5: Adalaide, Australië	2	1,92	96	15
13	Feitenkennis	5: Adalaide, Australië	2	1,83	92	12
14	Feitenkennis	6: Verenigde Staten	2	1,92	96	22
15	Topografisch Kaartbeeld	1: Waar ligt Italië?	2	2,00	100	0
16	Topografisch Kaartbeeld	2: Waar ligt Kuweit?	2	1,33	67	42
17	Topografisch Kaartbeeld	3: Waar ligt Iran?	2	1,25	63	71
18	Topografisch Kaartbeeld	4: Waar ligt Bali?	2	0,92	46	-28
19	Topografisch Kaartbeeld	5: Waar ligt Australië?	2	2,00	100	0
20	Topografisch Kaartbeeld	6: Waar liggen deVerenigde Staten?	2	1,92	96	19

Uit bovenstaande tabel blijkt dat zowel hoge als lage RIR-waarden voor zijn gekomen in de afname van de 'hard-op-denk'-protocollen onder de groep 8 leerlingen. Zes variabelen (variabel nummer 6, 7, 9, 10, 16 en 17) hadden een RIR-waarde boven de 50, zij spelen dus een belangrijke rol in de betreffende vraag van het 'hard-op-denk'-protocol bij het beoordelen van het geografisch besef van leerlingen. Dit betekent dat leerlingen die bij deze opgaven veel punten kregen, ook vaak hogere toetsscores hebben in totaal en dat leerlingen die hier weinig punten kregen hebben tijdens de gehele afname van de gea-toets ook meer fouten gemaakt. Twee variabelen (variabel 12 en 13) hadden een RIR-waarde tussen de 0 en 20, hierdoor komt deze variabel weinig overeen met de rest van de toetsscores en zijn deze variabelen weinig van invloed in het onderscheiden van leerlingen met verschillend geografisch besef. Zeven variabelen (variabel 1, 2, 4, 8, 11, 15 en 19)hadden een RIR-waarde van 0, dit houdt in dat zij geen belangrijke rol spelen in het onderscheiden van leerlingen

met veel of juist weinig geografische vaardigheden en kennis. Wanneer men naar het gemiddeld aantal verkregen punten kijkt wordt duidelijk waarom, alle leerlingen hebben op deze vragen 2 punten gescoord, gelijk aan het maximaal aantal punten dus. De variabelen met een lage RIR-waarde, of een RIR-waarde van 0 betreffen voornamelijk de vaardigheden 'foto-interpretatie' en 'kaart lezen/ grafiek lezen'. Dit is te verklaren doordat deze vaardigheden niet de vaardigheden zijn waar de vraag voornamelijk om draaide, maar eerder vaardigheden zijn die indirect een rol spelen bij het beantwoorden van de vraag. Bij het vergelijken van foto's ('Welke foto is niet in Italië genomen?') is foto-interpretatie (begrijpt de leerling wat op de foto staat?) belangrijk maar draait het erom dat hun thematisch kaartbeeld van Italië goed genoeg ontwikkeld is om de vraag te kunnen beantwoorden. Men kan ervan uitgaan dat leerlingen in regulier onderwijs begrijpen wat op een foto staat afgebeeld, bijvoorbeeld een gebouw, een strand of een woestijn. Deze variabel geldt daarom eerder als een check om zeker te weten dat wanneer leerlingen een vraag fout beantwoorden het niet wordt veroorzaakt doordat de afbeelding onduidelijk is of vaag is weergegeven. Één variabel (variabel 18) geeft een negatief verband aan, het betreft hier de vraag: 'Waar ligt Bali?', waarbij de leerling Bali op de wereldkaart moest aanwijzen. De achterliggende reden van dit negatieve verband is dat veel leerlingen met een hoge score op de gea-toets Bali 'bovenliggend' hebben kunnen aanwijzen. Dit betekent dat zij een kruisje hebben gezet in een buurland dat vlak bij Bali ligt, maar niet Bali of Indonesië hebben aangekruist. Hierdoor kregen zij op dit onderdeel slechts één van de twee te verkrijgen punten. Één van de leerlingen die een lage score op de gea-toets heeft gehaald kon Bali echter precies aanwijzen (een vriendinnetje van haar was daar op vakantie geweest en had daarover verteld) en kreeg zodoende twee punten op dit onderdeel. De negatieve RIR-score is dus grotendeels veroorzaakt door een toevalstreffer onder de leerlingen die, vanwege het geringe aantal deelnemende leerlingen aan de 'hard-op-denk'-protocollen, behoorlijk van invloed op de score is geweest.
Onderstaande tabel geeft de RIR-waarden weer ingedeeld per vaardigheid zoals die aan bod zijn gekomen tijdens de 'hard-op-denk'-protocollen in groep 8 van het basisonderwijs.
De vaardigheid 'topografisch kaartbeeld' kent de meeste hoge RIR-waarden evenals kennis van schaalniveaus en daarmee is er bij deze vaardigheden veel samenhang tussen het onderscheid dat deze vaardigheid maakt tussen leerlingen en de scores van de gea-toets. Daarnaast lijkt de gea-toets ook direct aanspraak te doen op de vaardigheden thematisch kaartbeeld, begripskennis en kennis van principes. De vaardigheden kennis van schaalniveaus, vergelijking van gebieden en feitenkennis lijken ook, weliswaar in mindere mate, te worden aangesproken. Het belang van foto-interpretatie kan op basis van deze resultaten niet worden aangetoond. De RIR-waarde van thematisch kaartbeeld en vergelijking van gebieden is laag, maar deze vaardigheden behoort op niveau 2 nog niet tot de direct getoetste vaardigheden in de gea-toets. Het thematisch kaartbeeld komt in de gea-toets vooral voor als vragen over de bevolking waarvan er op niveau 2 nog niet veel gesteld worden aangezien leerlingen vaak pas op de middelbare school aan de slag gaan met thema's als de spreiding van welvaart, religie of klimaten.

Tabel 21: RIR-waarden per vaardigheid in 'hard-op-denk'-protocollen groep 8

vaardigheid	**0-20**	**20-50**	**50-100**	**totaal**
Foto-interpretatie	2			2
Kaart lezen/ Grafiek lezen	2	1		3
Kennis van schaalniveaus		1	2	3
Vergelijking van Gebieden	1			1
Feitenkennis	1	1		2
Thematisch kaartbeeld	1			1
Topografisch kaartbeeld	2	1	3	7
Begripskennis			1	1
Totaal aantal gemeten	**9**	**4**	**6**	**20**

Onderstaande tabel geeft de resultaten weer van de afname van de 'hard-op-denk'-protocollen in de derde klas voortgezet onderwijs. Elf variabelen (variabel nummer 1, 2, 6, 8, 10, 14, 15, 16, 17, 18 en 20) hadden een RIR-waarde boven de 50, zij spelen dus een belangrijke rol in de betreffende vraag

van het 'hard-op-denk'-protocol bij het beoordelen van het geografisch besef van leerlingen. Met name de RIR-waarden van variabel 14 en 17 zijn erg hoog, respectievelijk 81 en 84. Beide variabelen zijn ingedeeld in de categorie topografisch kaartbeeld waarbij leerlingen konden aangeven waar de Filipijnen en Belfast ligt. Aangezien deze topografische eenheden niet algemeen bekend bleken te zijn onder alle leerlingen waren het uitsluitend de leerlingen met een hoge score op de gea-toets die op dit onderdeel veel punten kregen. Dezelfde soort verklaring zou ook de andere hoge RIR-waarden bij topografisch kaartbeeld kunnen verklaren. Verder heeft ook de vaardigheid kennis van principes een hoge RIR-waarde. Dit houdt in dat om de vraag 'wat is geen gevolg van de aanleg van een stuwdam?' juist te kunnen beantwoorden, een leerling de vraag 'wat is de functie van een stuwdam?/wat is het gevolg van de aanleg van een stuwdam?' kan beantwoorden.

Vijf variabelen (variabel 5, 11, 12, 13 en 19) hadden een RIR-score tussen 20 en 50 waardoor zij een redelijke rol spelen in de beoordeling van de geografische kennis van een leerling. Slechts twee variabelen (variabel 7 en 9) had een RIR-waarde tussen de 0 en 20, hierdoor komt deze variabel weinig overeen met de rest van de toetsscores. Twee variabelen (variabel 3 en 4) hadden een RIR-waarde van 0, dit houdt in dat zij geen rol spelen in het onderscheiden van leerlingen met veel of juist weinig geografische vaardigheden en kennis. Het betreft hier wederom de vaardigheden foto-interpretatie en kaart lezen/ grafiek lezen die werden toegepast bij de foto-vergelijkingsvraag over Thailand ('Welke foto is niet in Thailand genomen?'). Dit houdt in dat alle leerlingen begrepen hebben wat er op de foto's was afgebeeld en het kaartje van Thailand als zodanig hebben erkend. Maar verder werd op de vaardigheden foto-vergelijking en kaart lezen/grafiek lezen niet direct beroep gedaan. Uit de uitslag blijkt dat leerlingen bij deze vraag bij het juiste antwoord komen door met name beroep te doen op hun thematisch kaartbeeld.

Tabel 23: totaal uitslag resultaten derde klas voortgezet onderwijs

Nummer variabel	**Code vaardigheid**	**Gemeten bij opgave**	**Maximum aantal punten**	**Gemiddeld aantal punten**	**P-waarde**	**RIR-waarde**
1	Begripskennis	1: Haditha, Irak	2	0,92	46	57
2	Kennis van Principes	1: Haditha, Irak	2	0,67	33	71
3	Foto-Interpretatie	2: Bangkok, Thailand	2	2,00	100	0
4	Kaart lezen/grafiek lezen	2: Bangkok, Thailand	2	2,00	100	0
5	Kennis van Schaalniveaus	2: Bangkok, Thailand	2	1,67	83	40
6	Thematisch Kaartbeeld	2: Bangkok, Thailand	2	1,08	54	70
7	Kaart lezen/ Grafiek lezen	3: Panglao, Filipijnen	2	1,58	79	18
8	Thematisch Kaartbeeld	3: Panglao, Filipijnen	2	1,08	54	61
9	Kaartlezen/ Grafiek lezen	4: Rio de Janeiro	2	1,92	96	17
10	Begrips kennis	4: Rio de Janeiro	2	0,75	38	65
11	Vergelijking van Gebieden	5: Groenland	2	1,67	83	45
12	Thematisch Kaartbeeld	5: Groenland	2	1,58	79	39
13	Feitenkennis	6: Belfast	2	0,25	13	34
14	Topografisch kaartbeeld	6: Belfast	2	0,83	42	81

Nummer variabel	Code vaardigheid	Gemeten bij opgave	Maximum aantal punten	Gemiddeld aantal punten	P-waarde	RIR-waarde
15	Topografisch kaartbeeld	1: Waar ligt Irak?	2	0,92	46	57
16	Topografisch kaartbeeld	2: Waar ligt Thailand?	2	1,25	63	72
17	Topografisch kaartbeeld	3: Waar liggen de Filipijnen?	2	1,42	71	84
18	Topografisch kaartbeeld	4: Waar ligt Brazilië?	2	1,58	79	70
19	Topografisch kaartbeeld	5: Waar ligt Groenland?	2	1,75	88	27
20	Topografisch kaartbeeld	6: Waar ligt Belfast?	2	0,83	42	61

Onderstaande tabel geeft de RIR-waarden weer ingedeeld per vaardigheid zoals die aan bod zijn gekomen tijdens de 'hard-op-denk'-protocollen in de derde klas voortgezet onderwijs.
Opvallend is dat met name de vaardigheid 'topografisch kaartbeeld' hoge RIR-waarden kent en daarmee is er veel samenhang tussen het onderscheid dat deze vaardigheid maakt tussen leerlingen en de scores van de gea-toets. Dit betekent dat de beschikking hebben over een goed topografisch kaartbeeld belangrijk is voor het succesvol afleggen van de gea-toets. Daarnaast lijkt de gea-toets ook direct aanspraak te doen op de vaardigheden thematisch kaartbeeld, begripskennis en kennis van principes. De vaardigheden kennis van schaalniveaus, vergelijking van gebieden en feitenkennis lijken ook, weliswaar in mindere mate, te worden aangesproken. Het belang van foto-interpretatie kan op basis van deze resultaten niet worden aangetoond. De RIR-waarden van kaart lezen/ grafiek lezen vallen ook laag uit, maar deze vaardigheid behoort op niveau 3 ook niet meer tot de direct getoetste vaardigheden in de gea-toets (tabel 1).

Tabel 23: RIR-waarden per vaardigheid in 'hard-op-denk'-protocollen derde klas voortgezet onderwijs.

Vaardigheid	0-20	20-50	50-100	Totaal
Foto-interpretatie	1			1
Kaart lezen/ Grafiek lezen	3			3
Kennis van schaalniveaus		1		1
Vergelijking van Gebieden		1		1
Feitenkennis		1		1
Thematisch kaartbeeld		1	2	3
Topografisch kaartbeeld		1	6	7
Begripskennis			2	2
Kennis van principes			1	1
Totaal aantal gemeten	**4**	**5**	**11**	**20**

6 conclusie

In dit hoofdstuk worden de resultaten van de 'hard-op-denk'-protocollen in verband gebracht met de gegevens uit het theoretisch kader en de methodologie. Eerst wordt de hoofdvraag van het onderzoek beantwoord en de hypothese verworpen dan wel aangenomen, vervolgens wordt er nog gekeken naar de beperkingen van het onderzoek en suggesties gedaan voor toekomstig onderzoek.

6.1 Beantwoording hoofdvraag, aanname hypothese

6.1.1 Beantwoording van de hoofdvraag:

In deze paragraaf worden conclusies uit de resultaten van het onderzoek getrokken en uiteindelijk het antwoord gegeven op de hoofdvraag:

' ***Op welke soort geografische kennis doen leerlingen beroep bij het beantwoorden van de vragen van de Gea-toets en wat zegt dit over de validiteit van de toets?***

De meest opvallende uitkomsten en het antwoord op de hoofdvraag van het onderzoek kunnen in een aantal punten worden samengevat:

- *Denkfouten en misconcepties*

Uit de 'hard-op-denk'-protocollen is gebleken dat er sprake is van het bestaan van enkele misconcepties bij aardrijkskundige thema's, maar dat veel leerlingen (met name diegenen geselecteerd op basis van hoge scores op de gea-toets) daar wel doorheen prikken. Dit werd duidelijk bij de tweede vraag van het 'hart-op-denk'-protocol op niveau 3 over Thailand (Welke foto is niet in Thailand gemaakt?). Leerlingen met een lage score raakten in de war bij het zien van een afbeelding van luxe appartementen-flats terwijl leerlingen met hogere scores vaak wisten dat ook in armere landen men luxe gebouwen kan aantreffen. Verder leverde de vraag over de spreiding van tuberculose op de Filipijnen en Zuidoost Azië ook enige verwarring op bij sommige leerlingen aangezien veel leerlingen niet het verband konden leggen tussen welvaart en de spreiding van een ziekte. Zij vonden het logischer wanneer dit te maken had met eetgewoonten of klimaat.

- *Opvallende relevantie topografische kennis in de derde klas*

Uit de hoge RIR-waarden bij de vaardigheid topografisch kaartbeeld onder de leerlingen uit de derde klas voortgezet onderwijs lijkt het erop dat topografische kennis een goede indicator is voor de totale score op de gea-toets. RIR-waarden geven immers de samenhang van een variabel met de rest van de toetsscores weer. Dit geeft aan dat topografische kennis als een basis kan worden beschouwd van waaruit het geografisch besef van leerlingen zich verder kan ontwikkelen. Wanneer deze kennis ontbreekt wordt het lastig voor leerlingen om de vraag goed te beantwoorden wanneer deze bijvoorbeeld beroep doet op het thematisch kaartbeeld van leerlingen. Zo wordt een vraag over het klimaat of de economische situatie van bijvoorbeeld Brazilië ten opzichte van de andere landen in de regio erg ingewikkeld wanneer een leerling denkt dat Brazilië in Noord-Afrika ligt. Opvallend is dat leerlingen van groep 8, niet enkel de leerlingen met hoge scores op de gea-toets maar collectief, veel beter scoorden op de vragen om een land aan te wijzen dan de leerlingen in de derde klas. Een verklaring hiervoor kan worden gevonden in het feit dat in groep 8 naar 'makkelijkere/ herkenbaardere' landen en gebieden werd gevraagd zoals de VS en Australië dan in de derde klas waar werd gevraagd naar de Filipijnen en Belfast. Daarnaast behoort het verwerven van topografische kennis over gebieden buiten Europa tot het curriculum van het laatste jaar van de lagere school waardoor de leerlingen in groep 8 voor het beantwoorden van deze vragen uit recentelijk opgedane kennis konden putten. Hierdoor zou het kunnen dat leerlingen in de derde klas deze stof ook ooit hebben beheerst maar dat deze kennis verwaterd is. Leerlingen uit de derde klas die hoge scores behaalden zowel op de gea-toets als tijdens de 'hard-op-denk'-protocollen zullen deze kennis waarschijnlijk beter onderhouden hebben.

Kortom: een goed topografisch kaartbeeld leidt niet automatisch tot een hoge score op de gea-toets, daarvoor zijn andere vaardigheden zoals het thematisch kaartbeeld en begripskennis ook voor nodig,

maar een goed ontwikkeld topografisch kaartbeeld lijkt wel een voorwaarde te zijn voor de ontwikkeling van geografisch besef.

- Persoonlijke binding met onderwerp/land/streek
Een ander aspect dat van invloed is geweest op de afname van de 'hard-op-denk'-protocollen, maar waarnaar verder geen specifiek onderzoek is gedaan, is de rol die de persoonlijke binding met een onderwerp of gebied speelt. Opvallend was dat sommige leerlingen over bepaalde onderwerpen veel wisten of bepaalde vragen konden beantwoorden omdat zij daar toevallig veel vanaf weten. Dit is de reden dat er uit de 'hard-op-denk'-protocollen bij één variabel sprake was van een negatief verband. Terwijl de meeste leerlingen Bali op de wereldkaart enkel ongeveer konden aankruisen was er een leerlinge die was geselecteerd op basis van haar lage score tijdens de gea-toets die het eiland precies kon aanwijzen. Dit kwam omdat een vriendinnetje van haar net op vakantie in Bali was geweest en daarover verteld heeft. Tijdens het verloop van de 'hard-op-denk'-protocollen is het nog een aantal keer voorgekomen dat leerlingen een antwoord wisten te geven op een vraag gebaseerd op toevallige informatie uit de omgeving zoals een klasgenoot die uit Iran kwam en daarover had verteld of dat een leerling die toevallig zijn spreekbeurt of verslag over dat onderwerp heeft geschreven. Hoewel dit onderzoek niet voldoende gegevens heeft om hierover algemeen geldende uitspraken te kunnen doen doet dit het vermoeden rijzen dat leerlingen informatie waar ze op een bepaalde manier meer affiniteit mee hebben of persoonlijk bij betrokken zijn beter tot zich nemen of kunnen onthouden.

- Actualiteit
Opvallend in de uitkomsten van het onderzoek is de rol die de actualiteit speelt bij het beantwoorden van geografische vragen. Wanneer bepaalde conflicten of situaties actueel zijn worden zij veelvuldig genoemd in de media. Zo gaven leerlingen bijna allemaal aan van Iran of Irak gehoord te hebben op televisie vanwege de oorlogen en vredesmissies waarover al een aantal jaar wordt gerapporteerd. Aan de andere kant zijn conflicten die de laatste jaren op de achtergrond zijn geraakt de leerlingen relatief onbekend. Een voorbeeld hiervan is het territoriale conflict tussen Ierland en Groot-Brittannië over Noord-Ierland waarnaar gevraagd werd in de zesde vraag van de 'hard-op-denk'-protocollen. De leerlingen die het juiste antwoord konden geven op deze vraag deden hierbij beroep op hun topografische kennis. De achterliggende reden van het conflict konden zij niet precies geven, evenals de naam van de organisatie benoemen.

- vergelijking basisschool met middelbare school
De uitslag van de 'hard-op-denk'-protocollen op niveau 3 (derde klas voortgezet onderwijs) komen redelijk overeen met de bevindingen van deze protocollen op niveau 2 (groep 8 lagere school). Aangezien de leerlingen in groep 8 andere vragen kregen voorgelegd dan de leerlingen in de derde klas voortgezet onderwijs zijn de resultaten inhoudelijk niet te vergelijken. Opvallend is wel dat leerlingen uit groep 8 veel meer landen en gebieden precies hebben kunnen aanwijzen dan leerlingen op de middelbare school, daarnaast ligt het aantal juist beantwoorde vragen veel hoger in groep 8 dan in de derde klas. Of hierbij sprake is van een causaal verband of dat het hier eerder twee kenmerken betreft van leerlingen met een grote interesse in aardrijkskunde is moeilijk te zeggen. Een combinatie van deze twee verklaringen zou ook mogelijk zijn. Alhoewel de mogelijkheid van het bestaan van een causaal verband ook zou verklaren waarom leerlingen met hoge scores op de middelbare school meer landen kunnen aanwijzen dan leerlingen met een lage score.

De leerlingen op de lagere school hebben beduidend hogere scores behaald op de 'hard-op-denk'-protocollen dan de leerlingen op de middelbare school. Hierdoor rijst het vermoeden dat de gekozen vraagstellingen op niveau 2 wellicht aan de makkelijke kant zijn geweest. Aangezien de leerlingen van groep 8 wel zeer uiteenlopende, ook lage, scores hebben behaald op de klassikale afname van de gea-toets kan het zijn dat de moeilijkheidsgraad van de opgaven van het 'hard-op-denk'-protocol niet geheel representief is voor de rest van de toets. Het is dus mogelijk dat met een andere selectie voor

de ‘hard-op-denk’protocollen de leerlingen de vragen minder goed hadden gemaakt en dat de scores meer overeen waren gekomen met de scores van de middelbare scholieren. Het onderzoek heeft bij de niveau 3-vragen ook meer inzichten opgeleverd, aangezien er drie vragen waren op niveau 2 tijdens de ‘hard-op-denk’-protocollen (vraag 1 over Italië, vraag 4 over Bali en vraag 5 over Australië) die alle leerlingen goed hebben beantwoord. Deze vragen zijn tijdens de ‘hard-op-denk’-protocollen dus niet heel nuttig gebleken in het onderscheiden van verschillende niveaus onder de leerlingen. Aangezien het niet is voorgekomen dat vragen op niveau 3 door iedere leerling goed waren beantwoord voldeed het aantal goed en slecht beantwoorde vragen op dit niveau meer aan de verwachtingen dan de uitslagen van de ‘hard-op-denk’-protocollen op niveau 2.

- Beantwoording hoofdvraag: Op welke soort geografische kennis doen leerlingen beroep bij het beantwoorden van de vragen van de gea-toets?
Uit de ‘hard-op-denk’-protocollen is gebleken dat de leerlingen op niveau 2 (groep 8) met name gebruik maken van de vaardigheden begripskennis, topografische kaartbeeld en kennis van schaalniveaus. Op niveau 3 (derde klas voortgezet onderwijs) komen de vaardigheden thematisch kaartbeeld, topografisch kaartbeeld, begripskennis en kennis van principes uit de ‘hard-op-denk’-protocollen als meest gebruikte vaardigheden bij het oplossen van de vragen. Hiermee is het belang van goed ontwikkelde begripskennis, thematisch kaartbeeld (vanaf de middelbare school), topografisch kaartbeeld, kennis van schaalniveaus (op de lagere school) en kennis van principes (vanaf de middelbare school) voor het afleggen van de Gea-toets aangetoond. Dit betekent voor de Gea-toets dat deze vaardigheden die behoren tot het competentie domein op een juiste manier zijn omgezet in toetsvragen.
De andere vaardigheden waarvan het belang in dit onderzoek niet is aangetoond kunnen toch een rol spelen bij de afname van de gea-toets aangezien de vragen uit de ‘hard-op-denk’-protocollen niet specifiek waren gericht op het toetsen van die vaardigheid. Vaardigheden zoals foto-interpretatie zijn hierin factoren die indirect een rol spelen maar waar niet direct een beroep op werd gedaan. Aangezien slechts zes vragen waren opgenomen in het ‘hard-op-denk’-protocol zou bij een andere selectie van vragen wellicht andere uitkomsten gemeten worden. Misschien kan het belang van de vaardigheden die in dit onderzoek lage RIR-waarden opleverden dan wel worden aangetoond.

De resultaten die de ‘hard-op-denk’-protocollen opleverden geven een indicatie dat de gea-toets een valide wijze is om de hierboven genoemde geografische vaardigheden met hoge RIR-waarden te meten. Het belang van de andere vaardigheden is in dit onderzoek nog niet aangetoond.

6.1.2 Aanname hypothese
In paragraaf 4.3 werd de volgende hypothese geformuleerd:

‘Leerlingen met een hogere score op de Gea-toets zullen tijdens de ‘hard-op-denk’-protocollen ook meer vaardigheden en denkprocessen tonen die toe te schrijven zijn aan geografisch besef dan leerlingen met een lagere score op de Gea-toets’.

Wanneer deze hypothese wordt aangenomen betekent dat dat de gea-toets inderdaad geografisch besef meet en een onderscheid maakt tussen vaardige en minder vaardige leerlingen. Uit tabel 19 blijkt dat leerlingen met een hoge score op de gea-toets ook vaak een hogere score hadden voor de vaardigheden die bij de ‘hard-op-denk’-protocollen gemeten waren dan de leerlingen die op basis van een lage score waren geselecteerd. Daarnaast konden zij ook vaker het land dat in de vraag centraal stond aanwijzen op een wereldkaart. Uit tabel 20 en 22 is gebleken dat een groot aantal variabelen uit de ‘hard-op-denk’-protocollen correspondeerden met de uitslagen uit de afname van de gea-toets. Dit betekent dat er een samenhang is tussen de hogere scores in de gea-toets en de hoge scores gemeten met behulp van protocollen, evenals een verband tussen de lage scores in de gea-toets en de lage scores van de protocollen. Gebaseerd op deze resultaten kan men zeggen dat leerlingen met een hogere score op de gea-toets inderdaad tijdens de ‘hard-op-denk’-protocollen

meer vaardigheden en denkprocessen hebben getoond die toe te schrijven zijn aan geografisch besef dan de leerlingen met een lage score op de gea-toets. Op basis van deze bevinding kan de hypothese worden aangenomen.

6.2 Beperkingen van het onderzoek

Bij het uitvoeren van het onderzoek zijn er een aantal beperkingen die in acht genomen moeten worden. Ten eerste zijn de uitkomsten uit de 'hard-op-denk'-protocollen niet statistisch significant dus kunnen er geen generaliserende, algemeen geldende uitspraken worden gedaan. De conclusies die zojuist getrokken zijn, zijn dan ook enkel gebaseerd op de uitslagen van de 'hard-op-denk'-protocollen die met slechts 24 leerlingen zijn gevoerd, het kan zijn dat de keuze voor andere leerlingen een andere uitkomst zou hebben opgeleverd. Doordat het onderzoek bij een beperkt aantal leerlingen is uitgevoerd zou een per toeval gekozen uitzonderlijk vaardige, of juist onvaardige, leerling invloed kunnen hebben gehad op de resultaten. Alhoewel dit onderzoek een goede indicatie aangeeft van het geografisch besef van leerlingen in groep 8 of derde klas van het voortgezet onderwijs zal uitgebreider onderzoek een vereiste zijn om statistisch significante gegevens bij dit onderwerp te genereren.

Een tweede factor die van invloed is geweest op het verloop van het onderzoek is dat het niet makkelijk is geweest scholen te vinden die tijd en ruimte in hun programma hadden om te kunnen participeren aan dit onderzoek. Daarnaast lijkt het aardrijkskundige onderwijs op veel bassischolen vaak voornamelijk topografie in te houden en ligt de nadruk in groep 8 logischerwijs op van onderdelen die tijdens de Cito-eindtoets voorbij zullen komen. Bij veel scholen leek er een angst te bestaan om negatief beoordeeld te worden ondanks dat in het verzoek tot deelname aan het onderzoek is benadrukt dat de uitslagen enkel zouden worden gebruikt voor dit onderzoek en niet ter evaluatie van de individuele leerling of het aardrijkskundig onderwijs van de school zouden dienen. Voor een aantal scholen is dit toch een argument geweest om af te zien van participatie aan het onderzoek, zij vonden het niveau van hun leerlingen bij voorbaat niet optimaal en spraken bijvoorbeeld van 'hiaten in kennis' en 'verouderde lesmethoden'. Hierdoor bestaat het vermoeden dat scholen, met name de basisscholen, die wel participeerden aan het onderzoek zich geen zorgen maakten over de uitkomsten en waarschijnlijk meer tijd besteden aan aardrijkskunde dan veel andere scholen. Daarnaast staan de participerende scholen als goede scholen bekend. De Cito-scores van de Vondelschool liggen al jaren boven landelijk gemiddelde en uit een gesprek met de directeur van de Koningin Wilhelmina School bleek dat leerlingen van deze school 'ook veel door hun ouders werden gestimuleerd'. Het is zeer waarschijnlijk dat wanneer hetzelfde onderzoek was afgenomen bij scholen die bijvoorbeeld als 'zwak' bekend staan de resultaten anders waren geweest. Dit geldt met name voor de basisscholen aangezien aardrijkskunde op de middelbare school een vast onderdeel van het curriculum betreft, en dus altijd dient te worden gegeven in de derde klas HAVO/VWO, terwijl basisscholen zich waarschijnlijk in groep 8 zullen richten op vaststaande onderdelen van de Cito-eindtoets (aardrijkskunde is opgenomen in het onderdeel wereldoriëntatie, deze is echter niet verplicht).

6.3 Suggesties toekomstig onderzoek

Zoals al is aangegeven in paragraaf 6.2 zou aanvullend onderzoek een completer beeld kunnen opleveren van de ontwikkeling van het geografisch besef van leerlingen. Daarnaast dient ook vermeld te worden dat ondanks dat dit onderzoek enkel gericht is op geografisch besef er weldegelijk ook beroep gedaan wordt van de kennis en vaardigheden van leerlingen op andere vakgebieden. Om de vragen te beantwoorden is ook begrijpend lezen, een analytisch denkvermogen en ruimtelijk inzicht een vereiste. Wanneer een leerling hier grote moeite mee bleek te hebben en bijna geen vragen juist heeft beantwoord is het aannemelijk dat het niet puur een gebrek is aan aardrijkskundig besef maar dat de leerling waarschijnlijk ook moeite heeft met andere vakken die dezelfde soort denkwijzen vereisen zoals geschiedenis of maatschappijleer. Het zou dan kunnen dat een leerling

niet zozeer beschikt over gebrekkige aardrijkskundige kennis maar in het geheel een zwakke leerling is.

Bij verder onderzoek is het dan ook aan te raden dat dit gegeven in de analyse van 'hard-op-denk'-protocollen wordt meegenomen. Het uitvoeren van 'hard-op-denk'-protocollen bij andere schoolvakken zouden ook zeer interessante uitkomsten kunnen hebben. Hieruit zou een profiel kunnen ontstaan van niet alleen denkfouten, misconcepties en het gebruik van vaardigheden en kennis die kenmerkend zijn voor leerlingen met een weinig ontwikkeld geografisch besef, maar een meer algemeen beeld van denkfouten bij zwakkere leerlingen die wellicht vakoverstijgend zijn.

- Nawoord

Met deze thesis hoop ik mijn master Geo-Communicatie van de opleiding Sociale Geografie aan de Universiteit Utrecht af te sluiten.

Nu ik mijn thesis aan het afronden ben is mijn interesse in aardrijkskundig onderwijs zeker gegroeid. Ik vond het erg interessant om te werken met 'hard-op-denk'-protocollen aangezien daar weinig gebruik van wordt gemaakt. Enerzijds gaf het gebrek aan informatie over de werking en gebruik van deze onderzoeksmethode veel vrijheid, anderzijds was het soms ook juist lastig om hier invulling aan te geven. Naarmate mijn thesis vorderde werd ik steeds meer benieuwd naar de uitkomsten van het onderzoek en uiteindelijk werd ik door een aantal bevindingen zelfs oprecht verrast. Tijdens de afname van de 'hard-op-denk-'protocollen werd ik er voortdurend op gewezen hoe complex en divers de gedachtegangen van leerlingen zijn wanneer hen gevraagd wordt een aardrijkskundige opgave te maken. Er zijn zoveel manieren waarop leerlingen een vraagstuk kunnen benaderen dat het lang niet altijd makkelijk was om te beslissen in welke vaardigheden- of kenniscategorie een redenatie viel.
Ik hoop dat dat me uiteindelijk toch gelukt is. Daarnaast hoop ik dat deze thesis zou kunnen bijdragen aan de validering van de Gea-toets, hierdoor zou de thesis niet alleen van wetenschappelijke maar ook van maatschappelijke relevantie zijn. Tijdens de 'hard-op-denk'-protocollen bleek dat de niveaus van leerlingen onderling erg verschilden, vanwege de adaptieve aard van de Gea-toets zou zij daarom ook zeer geschikt zijn om in het vervolg door docenten te worden gebruikt om de aardrijkskundige kennis van hun leerlingen te kunnen vergelijken.

Bij deze wil ik graag mijn begeleider vanuit de Universiteit, mevr. Beneker, en mijn begeleider vanuit het Cito, dhr. Notté, bedanken voor hun adviezen. Ook mijn medestudenten hebben vaak suggesties aangedragen waar ik weer mee aan de slag kon als het onderzoek op een bepaald punt minder voorspoedig verliep. Daarnaast wil ik de docenten van de Koningin Wilhelmina School, de Vondelschool, de Koninklijke Scholengemeenschap en het Coornhert Gymnasium bedanken voor hun bereidheid mee te werken aan het onderzoek en het enthousiasme waarmee ze mij hebben ontvangen. En ten slotte natuurlijk de leerlingen die hebben meegewerkt aan het kwantitatieve en kwalitatieve onderzoek, want zonder hen waren er geen data en dus ook geen resultaten.

-Bronnenlijst

Baarda, D.B. & M.P.M. de Goede (2006), Basisboek Methoden en Technieken, handleiding voor het opzetten en uitvoeren van kwantitatief onderzoek. Groningen/Houten: Wolters-Noordhoff.

Downing, S.M. & T.M. Haladyna e.a. (2006), Handbook of Test development. New York-London: Routledge.

Favier, T. (2011), Geographic Information Systems in inquiry-based secondary geography education: theory and practice. Amsterdam: Ipskamp.

Gersmehl, P. (2008), Teaching Geography. New-York-London: the Guilford Press.

Haladyna, T.M. (2004), Developing and Validating Multiple-Choice Test Items. London: Lawrence Erlbaum Associates Publishers.

Notté, H. (red.) (2002), Aardrijkskunde voor de basisschool: een domeinbeschrijving als resultaat van een cultuurpedagogische discussie. Arnhem: Citogroep Arnhem.

Notté, H. & J.K.Verheij (2007), Aardrijkskunde is overal: een praktisch handboek voor het aardrijkskundeonderwijs op de basisschool. Enschede: uitgeverij Esstede.

Notte, H. & Schee, J. van der (2008). De wereld rond met Gea. Hoe staat het met onze geografische kennis? Geografie, jaargang 17 (2), 40-41.

Universiteit Twente research center voor examinering en certificering (2011), Een meetlat voor geografisch besef [online]
http://www.rcec.nl/onderzoeksprojecten/Notte.docx/ (geciteerd 15 februari 2012)

Stichting Cito Instituut (2011), Toetsen op School [online]
http://toetswijzer.kennisnet.nl/html/toetsenopschool/toetsenopschool.shtm (geciteerd 16 februari 2012)

Stichting Cito Instituut (28 februari 2012), Concept toetswijzer aardrijkskunde.

Vocht, A. de (2009), Validiteit en betrouwbaarheid. Reader Universiteit Utrecht Onderzoeksmethode II.

Wise, S.L., D.S. Bhola & Sheng-Ta Yang (2006), Taking the time to improve the validity of Low-Stakes Tests: the Effort-Monitoring CBT. Conference of the National Counsil of Measurement in Education in San Francicso.

Wise, S.L.& C.E. de Mars (2005), Low Examinee effort in Low Stakes Assessments : Problems and Potential Solutions. Educational Assessment 10, p.1-17.

Wools, S., T.Eggen & P.Sanders (2009), Evaluation of validity and validation by means of the argumentbased approach. Arnhem.

- Bijlage 1: het 'Hard-op-denk'-protocol

Naam:…… klas…… N…….

Opgave nr…….

(1,2,3 Alleen bij eerste opgave)

1 introductie van onderzoek, wat onderzoeken?
2 terugkomen op de toets
3 eventueel voorbeeldopgave

	BL	FI	KL/GR	KS	VG	FK	ThK	ToK	BK	Pr
4 lees hardop voor										
5 geef antwoord: 0 juist 0 onjuist										
6 Waarom is dit antwoord goed?										

---

7 Ben je bekend met het gebied?

0 nee
0 nee, wel bovenliggend
0 ja, van horenzeggen
0 ja, is er geweest

8 Kun je het gebied aanwijzen op de kaart?

0 precies
0 bovenliggend
0 fout

9 Waar zou dit verschijnsel verder kunnen voorkomen?

0 precies
0 bovenliggend
0 fout

- Bijlage 2: Uitslag Tia-Plus 'hard-op-denk'-protocollen groep 8

Toets en Item Analyse

Item Label	Item nr.	Gew.	Key	A	B	C	D	E	F	% O/D	# Mis-sing	Max	Mean	P	Sd	RSK	Rit	Rir	AR
				------ P- en A- waarden ------								---------- Gewogen ----------							
	1	1								0	0	1	0,75	75	0,43	0,43	39	35	85
	2	1								0	0	1	0,08	8	0,28	0,28	35	32	85
	3	1								0	0	1	0,67	67	0,47	0,47	52	48	85
	4	1								0	0	1	0,50	50	0,50	0,50	28	22	85
	5	1								0	0	1	0,67	67	0,47	0,47	22	17	85
	7	1								0	0	1	1,00	100	0,00	0,00	0	0	85
	8	1								0	0	1	0,83	83	0,37	0,37	45	41	85
	9	1								0	0	1	0,92	92	0,28	0,28	34	31	85
	10	1								0	0	1	0,92	92	0,28	0,28	23	20	85
	11	1								0	0	1	0,42	42	0,49	0,49	43	38	85
	12	1								0	0	1	0,83	83	0,37	0,37	53	50	85
	13	1								0	0	1	0,83	83	0,37	0,37	10	5	85
	14	1								0	0	1	0,75	75	0,43	0,43	37	33	85
	15	1								0	0	1	0,67	67	0,47	0,47	67	63	84
	16	1								0	0	1	0,58	58	0,49	0,49	42	37	85
	20	1								0	0	1	0,42	42	0,49	0,49	76	73	84
	22	1								0	0	1	0,75	75	0,43	0,43	23	18	85
	23	1								0	0	1	0,50	50	0,50	0,50	56	52	84
	24	1								0	0	1	0,42	42	0,49	0,49	64	60	84
	25	1								0	0	1	0,58	58	0,49	0,49	26	20	85
	26	1								0	0	1	0,67	67	0,47	0,47	43	39	85
	28	1								0	0	1	0,25	25	0,43	0,43	63	59	84
	29	1								0	0	1	0,92	92	0,28	0,28	23	20	85
	30	1								0	0	1	0,92	92	0,28	0,28	34	31	85
	31	1								0	0	1	0,58	58	0,49	0,49	22	16	85
	32	1								0	0	1	0,83	83	0,37	0,37	-6	-11	86
	33	1								0	0	2	2,00	100	0,00	0,00	0	0	85
	34	1								0	0	2	2,00	100	0,00	0,00	0	0	85
	35	1								0	0	2	1,92	96	0,28	0,14	27	24	85
	36	1								0	0	2	2,00	100	0,00	0,00	0	0	85
	37	1								0	0	2	1,75	88	0,60	0,30	49	43	85
	38	1								0	0	2	1,33	67	0,94	0,47	60	52	84
	39	1								0	0	2	1,42	71	0,86	0,43	60	53	84
	40	1								0	0	2	2,00	100	0,00	0,00	0	0	85
	41	1								0	0	2	1,75	88	0,60	0,30	57	52	84
	42	1								0	0	2	1,33	67	0,75	0,37	67	62	84
	43	1								0	0	2	2,00	100	0,00	0,00	0	0	85
	44	1								0	0	2	1,92	96	0,28	0,14	19	16	85
	45	1								0	0	2	1,83	92	0,55	0,28	19	13	85
	46	1								0	0	2	1,92	96	0,28	0,14	27	24	85
	47	1								0	0	2	2,00	100	0,00	0,00	0	0	85
	48	1								0	0	2	1,33	67	0,75	0,37	48	41	85
	49	1								0	0	2	1,25	63	0,83	0,41	76	72	84
	50	1								0	0	2	0,92	46	0,64	0,32	-23	-30	86
	51	1								0	0	2	2,00	100	0,00	0,00	0	0	85
	52	1								0	0	2	1,92	96	0,28	0,14	23	20	85
	53	1								0	0	3	2,25	75	0,60	0,20	59	54	84
	54	1								0	0	3	1,50	50	0,87	0,29	33	23	85
	55	1								0	0	3	1,83	61	0,37	0,12	37	33	85
	56	1								0	0	3	1,83	61	0,55	0,18	5	-2	86
	57	1								0	0	3	2,00	67	0,00	0,00	0	0	85
	58	1								0	0	3	2,08	69	0,28	0,09	-27	-30	86

SubGroepnummer : 0 SubTest nummer : 0
Aantal personen in toets : 12 Aantal geselecteerde items: 52
Minimum toetsscore : 0 Maximum toetsscore : 84

- Bijlage 3: Uitslag Tia-Plus ‘hard-op-denk’-protocollen derde klas voortgezet onderwijs

Toets en Item Analyse

Item Label	Item nr.	Gew.	Key	A	B	C	D	E	F	% O/D	# Missing	Max	Mean	P	Sd	RSK	Rit	Rir	AR
				------ P- en A- waarden ------								Gewogen							
	1	1								0	0	1	0,33	33	0,47	0,47	51	48	91
	2	1								0	0	1	0,08	8	0,28	0,28	40	38	91
	3	1								0	0	1	0,33	33	0,47	0,47	17	13	91
	4	1								0	0	1	0,42	42	0,49	0,49	50	46	91
	5	1								0	0	1	0,08	8	0,28	0,28	60	59	91
	6	1								0	0	1	0,33	33	0,47	0,47	63	61	90
	7	1								0	0	1	0,75	75	0,43	0,43	40	37	91
	8	1								0	0	1	0,50	50	0,50	0,50	11	7	91
	9	1								0	0	1	0,58	58	0,49	0,49	-5	-9	91
	10	1								0	0	1	0,08	8	0,28	0,28	-19	-21	91
	11	1								0	0	1	0,75	75	0,43	0,43	22	19	91
	12	1								0	0	1	0,42	42	0,49	0,49	54	51	91
	13	1								0	0	1	0,75	75	0,43	0,43	48	45	91
	14	1								0	0	1	0,83	83	0,37	0,37	11	8	91
	15	1								0	0	1	0,42	42	0,49	0,49	37	33	91
	16	1								0	0	1	0,50	50	0,50	0,50	24	20	91
	17	1								0	0	1	0,58	58	0,49	0,49	47	43	91
	18	1								0	0	1	0,58	58	0,49	0,49	47	43	91
	19	1								0	0	1	0,50	50	0,50	0,50	41	37	91
	20	1								0	0	1	0,92	92	0,28	0,28	29	27	91
	21	1								0	0	1	0,58	58	0,49	0,49	55	52	91
	22	1								0	0	1	0,50	50	0,50	0,50	-30	-34	91
	24	1								0	0	1	0,83	83	0,37	0,37	8	4	91
	25	1								0	0	1	0,58	58	0,49	0,49	4	-1	91
	26	1								0	0	1	0,50	50	0,50	0,50	48	45	91
	27	1								0	0	1	0,25	25	0,43	0,43	30	27	91
	28	1								0	0	1	0,33	33	0,47	0,47	26	22	91
	29	1								0	0	1	0,83	83	0,37	0,37	59	57	91
	30	1								0	0	1	0,58	58	0,49	0,49	61	58	90
	31	1								0	0	1	0,17	17	0,37	0,37	-19	-22	91
	32	1								0	0	1	0,50	50	0,50	0,50	27	23	91
	33	1								0	0	1	0,17	17	0,37	0,37	34	31	91
	35	1								0	0	1	0,50	50	0,50	0,50	24	20	91
	36	1								0	0	2	0,92	46	0,64	0,32	60	57	90
	37	1								0	0	2	0,67	33	0,62	0,31	74	71	90
	38	1								0	0	2	2,00	100	0,00	0,00	0	0	91
	39	1								0	0	2	2,00	100	0,00	0,00	0	0	91
	40	1								0	0	2	1,67	83	0,47	0,24	44	40	91
	41	1								0	0	2	1,08	54	0,86	0,43	73	70	90
	42	1								0	0	2	1,58	79	0,76	0,38	24	18	91
	43	1								0	0	2	1,08	54	0,95	0,48	66	61	90
	44	1								0	0	2	1,92	96	0,28	0,14	19	17	91
	45	1								0	0	2	0,75	38	0,83	0,41	69	65	90
	46	1								0	0	2	1,67	83	0,47	0,24	48	45	91
	47	1								0	0	2	1,58	79	0,64	0,32	44	39	91
	48	1								0	0	2	0,25	13	0,43	0,22	37	34	91
	49	1								0	0	2	0,83	42	0,69	0,34	83	81	90
	50	1								0	0	2	0,92	46	0,64	0,32	60	57	90
	51	1								0	0	2	1,25	63	0,72	0,36	75	72	90
	52	1								0	0	2	1,42	71	0,64	0,32	86	84	90
	53	1								0	0	2	1,58	79	0,64	0,32	72	70	90
	54	1								0	0	2	1,75	88	0,43	0,22	30	27	91
	55	1								0	0	2	0,83	42	0,80	0,40	65	61	90
	56	1								0	0	3	2,00	67	0,00	0,00	0	0	91
	57	1								0	0	3	2,00	67	0,00	0,00	0	0	91
	58	1								0	0	3	2,00	67	0,00	0,00	0	0	91
	59	1								0	0	3	2,00	67	0,00	0,00	0	0	91
	60	1								0	0	3	2,00	67	0,00	0,00	0	0	91
	61	1								0	0	3	1,67	56	0,62	0,21	64	60	90

SubGroepnummer : 0 SubTest nummer : 0
Aantal personen in toets : 12 Aantal geselecteerde items: 59
Minimum toetsscore : 0 Maximum toetsscore : 91