

WEVER



Op Weg naar Effectieve VERkeerseducatie

WEVER

op Weg naar Effectieve VERkeerseducatie

Uitgevoerd door:

SWOV: Divera Twisk, Jessica Witvoet, Simone Wesseling

Royal HaskoningDHV: Niki Hukker, Jan Vissers, Geertje Hegeman, Jasper Karsten, Wenkuo Chen

CITO: Erik Roelofs, Meggy Wijnen

CROW-KpVV: Wilma Slinger

In opdracht van het Interprovinciaal Overleg (IPO)

Foppe Koen

Amersfoort, april 2017

Inhoudsopgave

Inhoudsopgave	2
1. Inleiding	4
2. Ontwikkeling en kenmerken FietsmeetlatBO	6
3. Onderzoeksvragen en methode	8
3.1 Onderzoeksvragen	8
3.2 Methode	8
3.2.1 Deelnemers	9
3.2.2 Inhoud en afname FietmeetlatBO 3.0	10
3.2.3 Vragenlijsten	10
3.2.4 Procedure	10
3.2.5 Scoring en Data handling	10
3.2.6 Statistische analyses	12
4. Resultaten	14
4.1 Gebruikservaringen	14
4.2 Psychometrische eigenschappen en betrouwbaarheid	15
4.2.1 Situatiebewustzijn	16
4.2.2 Beslissen in complexe situaties	19
4.2.3 Sociaal en moreel handelen	27
4.2.4 Gevaarherkenning	29
4.2.5 Samenhang deeltolsten	30
4.3 Beïnvloeding testcores	31
4.3.1 Motivatie	31
4.3.2 Fietservaring	32
4.3.3 Cito-scores	33
4.3.4 Verkeerseducatie	34
4.3.5 Stedelijkheid	35
5. Discussie en Conclusie	37
5.1 De bruikbaarheid van de FietsmeetlatBO	37
5.1.1 Hoe functioneert FietsmeetlatBO 3.0 in de praktijk?	37
5.1.2 Hoeveel tijd hebben de leerlingen nodig om de toets te maken?	37
5.1.3 Blijven de leerlingen voldoende gemotiveerd gedurende de toets?	37
5.1.4 Investering in FietsmeetlatBO als basis voor toekomstige meetlatten	38
5.2 Psychometrische eigenschappen, betrouwbaarheid en beïnvloeders	38
5.2.1 Situatiebewustzijn	38
5.2.2 Sociaal en moreel handelen	39
5.2.3 Gevaarherkenning	39
5.2.4 Beslissen in complexe situaties	39
5.2.5 Worden met de toetsonderdelen verschillende competenties gemeten?	40
5.3 Conclusie	41
6. Aanbevelingen voor vervolg	42
Literatuur	45
Bijlagen	46
Bijlage 1 Informatiebrief scholen	46
Bijlage 2 Informatiebrief ouders	50
Bijlage 3 Onderzoekprotocol afname WEVER 3.0	52
Bijlage 4 Vragenlijst Verkeerseducatie voor docent	53
Bijlage 5 Vragenlijst Evaluatie FietsmeetlatBO	55
Bijlage 6 Vragenlijst Fietservaring	56
Bijlage 7 Vragenlijst Ervaringen in het verkeer	58

Voorwoord

Dit betreft een document met de resultaten uit de test met de FietsmeetlatBO3.0 van het WEVER project gericht op het Basis Onderwijs. Het doel van deze test was om bruikbaarheid, betrouwbaarheid, aantrekkelijkheid en technische randvoorwaarden van het instrument vast te stellen. Deze resultaten worden gebruikt om in vervolgonderzoek de Meetlat te verbeteren. WEVER betreft een samenwerking tussen Royal HaskoningDHV, SWOV, CITO en CROW. Het geteste prototype van de Meetlat is samengesteld uit testen die door Royal HaskoningDHV en SWOV zijn ontwikkeld, met advies van CITO. De hier gerapporteerde test met de FietsmeetlatBO 3.0 is door SWOV en RHDHV uitgevoerd, gerapporteerd en geanalyseerd in opdracht van Interprovinciaal Overleg (IPO).

1. Inleiding

Het doel van het WEVER project is te komen tot een structurele effectmeting van verkeerseducatie in de periode 2016-2020. Dit rapport is de uitkomst van het tweede van de vijf jaren (2016). Het bevat de resultaten van de toepassing van de nieuwe versie van de 'Fietsmeetlat BO'. Deze versie is doorontwikkeld op basis van de resultaten van het prototype (FietsmeetlatBO 1.0), en bij ongeveer 300 basisschoolleerlingen uit groep 8 afgenomen. Doel van deze afname is inzicht te krijgen in de kwaliteit van de meetlat, in de verkeerscompetenties van leerlingen aan het begin van groep 8 en in de invloed van externe factoren op het toetsresultaat. Resultaat daarvan is een beproefde fietstoets om de veiligheidsrelevante competenties van jonge fietsers in groep 8 vast te kunnen stellen. Voor een volledig overzicht van de achtergronden van WEVER, de onderlinge samenhang van rapport met de overige WEVER rapportages, en voor een doorkijkje naar het vervolgonderzoek wordt verwezen naar het samenvattende voortgangsverslag: "WEVER: 'op Weg naar Effectieve VERkeerseducatie' (Twisk et al., 2016). Op die punten beperkt dit rapport zich tot de hoofdlijnen.

In het project "op Weg naar Effectieve VERkeerseducatie" (WEVER) werken provincies en kennisinstituten samen aan het verhogen van de effectiviteit van verkeerseducatie. In aanloop tot dit project is opdracht verleend voor een onderzoek naar de haalbaarheid van het ontwikkelen van een *meetlat voor verkeerscompetenties*. Het eerste deel van deze opdracht betrof een theoretische analyse hiervan (Vissers et al., 2015). Daarin is vastgesteld dat een eerste versie van deze meetlat (het prototype) zich zou moeten richten op fietsende leerlingen uit groep 8 van de basisschool en de volgende competenties: Sitatiebewustzijn, Beslissen in complexe situaties, Sociaal en moreel handelen en Gevaarherkenning. Het tweede deel van het onderzoek betrof de ontwikkeling en toepassing van de prototype meetlat, de FietsmeetlatBO 1.0. Voor het derde deel van het onderzoek, waarvan hier de resultaten worden gerapporteerd, is de meetlat op basis van de resultaten verder ontwikkeld en getest. Via een tussenversie (FietsmeetlatBO 2.0) zijn we uitgekomen op de meest recente versie van de meetlat, vanaf hier te noemen de FietsmeetlatBO 3.0.

De ontwikkeling van de FietsmeetlatBO valt in het bredere kader van het streven naar systematische evaluaties van verkeerseducatieprogramma's (zie voor gedetailleerde beschrijving en achtergrond Slinger et al., 2016). Een inventarisatie uit 2015 liet zien dat van het grootste gedeelte van verkeerseducatieprogramma's geen effecten bekend zijn, dat in de meeste effectmetingen gebruik is gemaakt van zelfrapportagegegevens (n=9), dat in slechts vier studies objectieve metingen, zoals gevaarherkenningstoetsen, zijn gebruikt en dat in vier studies het gedrag van verkeersdeelnemers in het verkeer werd geobserveerd (Slinger et al., 2016). Naar aanleiding van deze resultaten is het rapport '10 gouden regels effectmeting' (2016) opgesteld, dat handvatten biedt voor het goed uitvoeren van effectmetingen (Vissers & Hegeman, 2015).

Hoewel het doel van elk verkeerseducatieprogramma is dat deelnemers zich veiliger gaan/kunnen gedragen, en er wel degelijk is geïnvesteerd in de verbetering van het verkeersonderwijs via de toolkit en de checklist, is de effectiviteit van de meeste programma's niet bekend (Hegeman et al., 2016). Financiers en ontwikkelaars weten dus niet of het uiteindelijke doel ook feitelijk bereikt wordt. Dit geeft vooral een (ethisch) probleem voor educatieve programma's die in scholen gebruikt worden. Het betekent immers dat leerlingen - zonder dat zij of hun ouders kunnen kiezen - blootgesteld worden aan programma's die mogelijk niet effectief zijn, die mogelijk leiden tot zelfoverschatting of zelfs kunnen resulteren in gevaarlijker gedrag (Twisk, 2014). Bovendien laten de ongevalsstatistieken zien dat het risico op ernstig letsel sterk stijgt in de beginjaren van de middelbare school (Twisk et al., 2016). Om deze redenen is in de eerste fase van de ontwikkeling van de WEVER meetlat gekozen voor de doelgroep 'Fietsende leerlingen in groep 8 (FietsmeetlatBO 1.0).

Doel van de FietsmeetlatBO is het zo betrouwbaar mogelijk in kaart brengen van het gewenste eindgedrag door middel van een zo direct mogelijke meting. Dat eindgedrag staat dan voor verkeersgedrag, waarvan we uit onderzoek weten dat het (vermoedelijk) een sterke relatie heeft met verkeersveiligheid. Die gedragingen kunnen worden gezien als indicatoren van risico's die in het verkeerssysteem aanwezig zijn en een belangrijke oorzaak (kunnen) zijn van het ontstaan van verkeersongevallen. Daarin is dus een sterke parallel te trekken met Safety Performance Indicators zoals die in PROMEV worden gebruikt (Aarts et al., 2015). Uitgangspunt bij de ontwikkeling van de FietsmeetlatBO was om opgaves te maken die zo 'goed als mogelijk' lijken op wat leerlingen in het verkeer gaan meemaken als ze van huis naar de middelbare school gaan fietsen, en waarvan het aannemelijk dat er een relatie is met onveiligheid. Kortom, het doel is dat de FietsmeetlatBO inzicht geeft in hoeverre jonge fietsers zijn voorbereid op de verkeerstaken die zij aan het einde van de basisschool bij de overgang naar het voortgezet onderwijs, zelfstandig, veilig en verantwoord moeten kunnen uitvoeren. Vervolgens zal dan met de FietsmeetlatBO – op termijn – verkeerseducatieprogramma's op effectiviteit met elkaar kunnen worden vergeleken.

Naast de eerder genoemde notitie over de haalbaarheid van een FietsmeetlatBO (Vissers et al., 2015), de beschrijving van de fietstoets zelf (Buijs and Witvoet, 2016) en het rapport over de prototype FietsmeetlatBO 1.0 (2016) bevat het nu voorliggende rapport de resultaten uit de proef met de FietsmeetlatBO 3.0 onder 335 leerlingen. Deze proef heeft tot doel om inzicht te krijgen in de kwaliteit van de meetlat, in de verkeerscompetenties van leerlingen aan het begin van groep 8 en in de invloed van externe factoren op het toetsresultaat.

Het eerste deel van dit rapport beschrijft kort de ontwikkeling en de kenmerken van de FietsmeetlatBO 3.0 (h. 2). Voor een meer gedetailleerd beschrijving van de deoltoetsen en de veranderingen ten opzichte van FietsmeetlatBO 1.0 en 2.0 verwijzen we naar het document 'Het ontwerp van de fietstoets' (Cito, 2016). Het volgende deel beschrijft de wijze van afname van de FietsmeetlatBO 3.0 bij basisschoolleerlingen van groep 8 (h.3) en de resultaten daarvan (h.4). Deze resultaten worden in het vervolgonderzoek van WEVER benut om de kwaliteit van de FietsmeetlatBO te vergroten. Om die reden zijn alle resultaten geanalyseerd en inzichtelijk gemaakt zonder dat extreme scores (outliers) en 'slecht scorende' items zijn verwijderd. Alleen scores die onbetrouwbaar waren vanwege technische mankementen zijn niet in de analyses betrokken. Het rapport wordt afgesloten met conclusies, discussie (h.5) en aanbevelingen (h.6). Verder is het doel van het voorliggende rapport om alle procedures, informatiebronnen, formulieren, brieven, kwalitatieve observaties van proefleiders en instructies te borgen voor gebruik in het vervolgonderzoek. Om dit rapport leesbaar te houden zijn deze opgenomen in de bijlagen.

2. Ontwikkeling en kenmerken FietsmeetlatBO

Voor elk van de competenties (Situatiebewustzijn, Beslissen in complexe situaties, Sociaal en moreel handelen en Gevaarherkenning) zijn voor de FietsmeetlatBO 3.0 één of twee deeldoetsen ontwikkeld. Dat is gedaan op basis van de resultaten van de prototype test met de FietsmeetlatBO 1.0. De toetsdomeinen Situatiebewustzijn en Beslissen in complexe situaties zijn gesplitst in elk twee deeldoetsen. Daardoor bestaat de FietsmeetlatBO 3.0 uit in totaal zes deeldoetsen. De ontwikkeling van elke deeldoets verschilt, zowel in proces als in onderbouwing. In de notitie 'Ontwikkeling van de fietstoets' (Buijs & Witvoet, 2016) wordt per deeldoets het ontwikkeltraject, de onderbouwing, de testkenmerken, de verschillen tussen versies en de technische aanbestedingswijze beschreven. Deze paragraaf beperkt zich tot de hoofdlijnen. Tabel 2.1 bevat voor elke deeldoets voorbeeldopgaven. De ontwikkeling van de deeldoetsen is uitgevoerd volgens de eisen die gesteld worden aan testontwikkeling (Vissers et al., 2015). Als eerste stap zijn de deeldoetsen afgenomen bij kleine groepjes leerlingen waarbij getrainde proefleiders systematisch zijn nagegaan of de opgaven gesnapt werden, of de instructies duidelijk, begrijpelijk en leesbaar waren, en of de testen niet te lang waren. De FietsmeetlatBO bevat deeldoetsen gericht op de volgende competenties:

- *Situatiebewustzijn*: het vermogen om veranderingen op te merken in verkeerssituaties.
- *Beslissen in complexe situaties*: het vermogen om complexe situaties te kunnen overzien en daar naar te handelen.
- *Sociaal en moreel handelen*: wel of geen ruimte geven aan andere verkeersdeelnemers, motieven om zich wel of niet aan verkeersregels te houden en de mate waarin men eigen ongewenst gedrag onschuldiger voorstelt dan het is.
- *Gevaarherkenning*: het vermogen om potentieel gevaarlijke situaties tijdig te kunnen identificeren.

Of leerlingen deze competenties ook hebben is mede afhankelijk van een goede kennis van de verkeersregels, de vaardigheid deze toe te kunnen passen, een goede motorische fietsvaardigheid, een goede vaardigheid om te kunnen fietsen in het verkeer, het vermogen zich in te leven in anderen en gemotiveerd zijn hun belangen niet te schaden. Deze voorwaarden worden niet expliciet door de FietsmeetlatBO getest. De aanname is dat deze zich impliciet doorvertalen in de prestatie op de FietsmeetlatBO (Vissers et al., 2015). Tabel 2.1 geeft een overzicht van de opgaven per deeldoets.

Competentie/deeltoets	Items	Voorbeeld item	Vraagstam	Respons
Situatiebewustzijn 12				
Waarnemen	6	Fietspad met zebrapad 	Waar moet je nu goed op letten? Klik op de plaats in de foto.	Aanklikken verkeersdeelnemer in foto
Besluitvaardigheid	6	Auto van rechts 	Wat is de juiste volgorde van voor laten gaan? Klik de volgorde aan.	Klik op verkeersdeelnemers in volgorde van voorrang
Complexe situaties 13				
Dode hoek	6	Vrachtwagen bij verkeerslicht 	Klik met de muis op een plek waar je veilig kunt wachten.	Aanklikken plek
Oversteken	7	Oversteekplaats 	Kun je nu veilig oversteken?	Toetsenbord V: veilig O: onveilig
Sociaal en moreel 10				
		Sluitende bomen overweg 	Van de 10 keer dat dit voorkomt, hoe vaak fiets je nog snel door?	0 t/m 10
Gevaarherkenning 14				
		Potentieel gevaar 	Druk op spatie als je denkt dat de situatie gevaarlijk kan worden.	Aanklikken potentiële gevaar

Tabel 2.1. Overzicht opgaven per deeltoets. Voor een uitgebreide beschrijving en veranderingen in versies 1.0, 2.0 en 3.0 zie het WEVER document: 'Het ontwerp van de deelttests van de FietsmeetlatBO 1.0, 2.0 en 3.0' (Cito en Royal HaskoningDHV, maart 2017)

3. Onderzoeksvragen en methode

Dit hoofdstuk beschrijft welke vragen beantwoord zijn met het afnemen van de FietsmeetlatBO 3.0 en via welke methode dat is onderzocht.

3.1 Onderzoeksvragen

Met de afname zijn de volgende vragen beantwoord over bruikbaarheid, betrouwbaarheid en validiteit van de FietsmeetlatBO 3.0. Voor een deel zijn het dezelfde vragen zoals die aan de orde gekomen zijn in de FietsmeetlatBO 1.0 en 2.0 (zie respectievelijk Witvoet, 2016 en Wesseling, 2016). Daarnaast zijn er voor WEVER 3.0 extra vragen beantwoord over de invloed van externe factoren op de toetsscores.

Bruikbaarheid

- Hoe functioneert WEVER 3.0 in de praktijk?
- Hoeveel tijd hebben de leerlingen nodig om de toets te maken?
- Blijven de leerlingen genoeg gemotiveerd gedurende de toets, of verliezen zij hun aandacht?

Psychometrische eigenschappen, betrouwbaarheid en beïnvloeders

- Wat is de kwaliteit en betrouwbaarheid van de Fietsmeetlat in WEVER 3.0?
- Wat is de moeilijkheidsgraad en het onderscheidend vermogen van de afzonderlijke toets opgaven?
- Worden met de 6 deeltaetsen de verschillende competenties gemeten?
- In welke mate hangen de opgaven binnen de deeltaetsen met elkaar samen? Als er samenhang blijkt te zijn, hoe is dit te verklaren?
- Kunnen met behulp van de toets verschillende niveaus van competenties worden aangetoond?

Aan vullende vragen over de invloed van externe factoren

- Worden de toets scores beïnvloed door andere factoren dan competentie:
 - o Fietservaring van het individuele kind,
 - o Waardering van de FietsmeetlatBO 3.0 en ervaren moeilijkheidsgraad per individueel kind,
 - o eerder ontvangen verkeerseducatie op klasniveau,
 - o Cito-scores op school niveau

3.2 Methode

In deze paragraaf is de onderzoeksmethode van de afname van de FietsmeetlatBO 3.0 beschreven.

3.2.1 Deelnemers

Twaalf basisscholen verspreid over Nederland zijn bereid gevonden om deel te nemen aan het onderzoek. Tabel 3.1 geeft een overzicht van de karakteristieken van de deelnemende scholen, en de verdeling van de deelnemers over de scholen.

School	Locatie	Stedelijkheid	Cito-score	Cito Categorie	Aantal deelnemers
Brede Hei	Amersfoort	Stad	538.1	Hoog	37
Noorderschool	Giethoorn	Platteland	539.4	Hoog	12
Bongerd	Ede	Platteland	520.2	Laag	6
De Hazessprong	Nijmegen	Stad	537.6	Hoog	53
Dorpschool	Bathmen	Platteland	537.1*	Laag	24
St. Jan	Oosteind	Platteland	531.9*	Laag	11
De Meander	Leersum	Platteland	534.9*	Hoog	24
De Schelp	Wevershoof	Platteland	530.0	Laag	28
Europaschool	Amsterdam	Stad	538.6	Hoog	27
Parkschool	Utrecht	Stad	526.5	Laag	20
School op de Berg	Amersfoort	Stad	541.6	Hoog	62
Rehoboth	Ochten	Platteland	538.3*	Hoog	35

*Van deze scholen zijn de cito-scores van eerdere jaren (2013-2014 of 2014-2015) gebruikt, omdat ze voor het jaar 2015-2016 een andere eindtoets (IEP of Route-8) dan Cito hebben gebruikt. Dit betekent ook dat de categoriën bij deze gemiddelden gebaseerd zijn op het landelijk gemiddelde van de eerdere jaren, waardoor een score in categorie 'laag' van de ene school hoger kan zijn dan een score in categorie 'hoog' van een andere school.

Tabel 3.1. Verdeling van leerlingen over scholen naar urbanisatiegraad en gemiddelde Cito-scores per school.

In totaal werd bij 335 leerlingen uit groep 8 de FietmeetlatB0 3.0 afgenomen: 51% jongens en 49% meisjes. De Cito-score per school is vergeleken met het landelijk gemiddelde. Scholen boven het gemiddelde zitten in de Cito-categorie 'hoog' en onder het gemiddelde in de categorie 'laag'. Stedelijkheid is bepaald op basis van aantal inwoners van de plaats waarin de school is gevestigd. Vanaf een inwonersaantal van 150.000 of meer is de plaats als "stad" gecategoriseerd.

Tabel 3.2 geeft een overzicht van de verkeerseducatie-activiteiten op de scholen. Deze informatie ontbreekt van de scholen Brede Hei en Noorderschool. Tevens hebben niet alle docenten de formulieren correct en duidelijk ingevuld. Om de gegevens toch te kunnen gebruiken heeft daarom een interpretatieslag plaatsgevonden.

School	Lang Program- ma	Per week (in min.)	# Korte Pro-gram- ma's	Praktijk Oefening	Extra Aanda cht Verkee rsveilig heid	Verkeers ouder	Veilige Schoolom -geving	Label
Bongerd	VVN	60	1	Incidenteel	N.v.t.	Nee	Ja	Nee
De	Noordhof	40	4	Incidenteel	Ja	Ja	Nee	Nee
Hazessprong								
Dorpschool	VVN	20	5	Wekelijks	Ja	Ja	Nee	Nee
St. Jan	VVN	35	5	Incidenteel	Ja	Ja	Nee	Ja
De Meander	VVN	30	0	Wekelijks	Nee	Nee	Nee	Ja
De Schelp	VVN	30	4	Wekelijks	Ja	Nee	Ja	Nee
Europaschool	Malmberg	30	2	Dagelijks	Ja	Nee	Nee	Nee
Parkschool	VVN	45	2	Wekelijks	Nee	Nee	Nee	Nee
School op de Berg	Noordhof	30	0	Wekelijks	Ja	Nee	Ja/Nee	Nee
Rehoboth	Malmberg	37.5	2	Wekelijks	Ja	Onbeken d	Nee	Nee

Tabel 3.2. Overzicht van verkeersveiligheidsactiviteiten per school

3.2.2 Inhoud en afname FietmeetlatBO 3.0

De 6 deeltaetsen werden geïnstalleerd op 24 HP Probook 4730s laptops, met een i5 processor, 17,3" scherm en een resolutie van 1600X900 pixels. Tabel 3..3. geeft voor elk onderdeel een voorbeeldvraag en daarnaast informatie over het aantal items, de antwoordwijze, de scoringsmethode en het aantal succesvolle afnames. Alle testen waren zelfstandig door de leerlingen uit te voeren, door voorbeeldopgaves en door het combineren van geschreven en gesproken instructie. Om volgorde effecten uit te sluiten, verschilden de volgorde van de deeltaetsen binnen FietmeetlatBO 3.0 tussen de leerlingen. De meeste leerlingen wisten de toets inclusief vragenlijst binnen het uur af te ronden.

3.2.3 Vragenlijsten

Voor inzicht in de relatie van de FietmeetlatBO 3.0 met de achtergrond variabelen, fietservaring, verkeerseducatie, en Cito-score van de school, zijn vragenlijsten ontwikkeld. De vragenlijsten over fietservaring en waardering van de fietsmeetlat werden als onderdeel van de FietsmeetlatBO 3.0 op de laptop afgenomen. De vragenlijst over verkeerseducatie werd door de docent van de klas ingevuld tijdens de afname. De vragenlijsten zijn opgenomen in de bijlagen.

3.2.4 Procedure

Scholen zijn geworven via het informele netwerk zodanig dat er voldoende spreiding was in schoolkenmerken zoals landelijk en stedelijk gevestigde scholen. Zie voor de uitnodigingsbrief aan scholen bijlage 1. Ouders werden geïnformeerd over het onderzoek via de school en konden binnen een week na ontvangst van de brief bezwaar maken bij de school tegen deelname van hun kind. Zie voor de informatiebrief voor ouders bijlage 2. Geen van de ouders heeft bezwaar gemaakt. Omdat de kinderen nog onder de 13 jaar waren, was een door hen getekend informed consent niet nodig. Op de dag van afname werden leerlingen klassikaal geïnformeerd over het doel en de procedure van het onderzoek door twee proefleiders. Leerlingen maakten de FietsmeetlatBO 3.0 in speciaal ingerichte klaslokalen met de hele groep tegelijk. Leerlingen kregen een proefpersoonnummer toegewezen waardoor de gegevens niet te herleiden zijn naar individuele leerlingen, maar de testresultaten uit de verschillende deeltaetsen wel koppelbaar zijn. Om volgorde effecten tussen de toetsen te vermijden werd aan elke leerling, op basis van kans, een eigen volgorde gegeven waarin de toetsen uitgevoerd moesten worden. Er waren 4 verschillende volgordes. Volgorde 1, 2, 3, en 4 werden respectievelijk door 75, 63, 68 en 78 kinderen gemaakt. Proefleiders volgden een protocol om te waarborgen dat afname gestandaardiseerd plaats vond (zie bijlage 4). Gedurende de afname konden leerlingen vragen stellen aan de proefleiders bij onduidelijkheden. Opmerkelijke zaken werden genoteerd door de proefleiders.

3.2.5 Scoring en Data handling

Tabel 3.3 geeft voor elke deeltaets een overzicht van aantal items, antwoordwijze, scoringsmethode, schaal constructie, minimum en maximum waarden en aantal correct afgeronde metingen.

De datafiles werden gekopieerd naar een veilige opslagomgeving, geback-up't, gekoppeld op proefpersoonnummer, ingelezen in een softwareprogramma voor statistische analyses (SPSS) en gecontroleerd op datavervuiling. Ook bij FietsmeetlatBO 3.0 zijn door technische problemen niet alle deeltaetsen van de leerlingen goed opgeslagen en/of afgerond. Deze scores zijn voorafgaande aan de analyses verwijderd. Daardoor verschilt het aantal scores per deeltaets.

Deeltoets	# Items	Voorbeeld item	Instructie	Response	N	Totaal score (Min – Max score)	Coding key
Situatie bewustzijn							
	6	Waarnemen 	Klik op het element in de foto waar je direct aandacht aan zou besteden.	Muisklik in de foto (hotspot)	283	Som van coding key (0-6)	Correct element = 1, incorrect. Ergens anders in de foto = 0
	6	Beslissen 	Wat is de juiste voorrangsvolgorde? . Klik in de juiste volgorde op de verkeersdeelnemers.	Muisklikken in de foto.	283	Som van coding key (0-6)	Correcte volgorde = 1, incorrect volgorde = 0
Complex situaties							
Dode hoek van vrachtwagens	6	Vrachtwagen wacht voor rood licht op kruising. 	Gebruik je muis om aan te geven waar je zou wachten	Plaats van de klik	324	Som van coding key (0-11)	Klik buiten dode hoek gebied is correct (1) anders incorrect (0).
Oversteken	24	Oversteekplaats 	Kan je nu veilig oversteken	Toetsenbord V: veilig O: Onveilig	260	Som van coding key (0-24)	Correct element (1), incorrect (0)
Sociaal en Moreel gedrag							
	10		Uit de 10 keer dat dit gebeurt, hoe vaak doe je dan	Keuze op een schaal van 0 tot 10	287	Som van coding key (0-100)	Score is het totaal over de items. Hoge score is minder sociaal en moreel handelen (in de ruwe data).
Gevaarherkenning							
	14	Mogelijk gevaar 	Druk op de spatiebalk als je het gevoel hebt dat het gevaarlijk kan worden.	Reactietijd (% van de beschikbare tijd om nog tijdig te kunnen reageren.	313	Som van coding key (0-14)	Een snellere reactie geeft een hogere score (max 1)

Tabel 3.3. Overzicht van deeltoetsen

3.2.6 Statistische analyses

Om inzicht te krijgen in de psychometrische eigenschappen van de deoltoetsen zijn eerst psychometrische analyses verricht om de betrouwbaarheid van de deoltoetsen als geheel te bepalen en de psychometrische kwaliteit van de afzonderlijke opgaven, dat wil zeggen de moeilijkheidsgraad en het onderscheidend vermogen.

Alle deoltoetsen zijn onderzocht op psychometrische kwaliteit met behulp van klassieke betrouwbaarheidsanalyses, die zijn gebaseerd op variantie in scores. Gekeken is naar de interne consistentie van de deoltoetsen door Cronbach's α te berekenen, of KR20 in het geval van dichotome schalen. Per opgave is bovendien gekeken naar de moeilijkheidsgraad van de opgaven (het aandeel leerlingen dat de opgave correct gemaakt heeft, de zogenaamde p-waarde) en het onderscheidend vermogen van de opgaven (R_{it} ; het vermogen onderscheid te maken tussen hoogscorders en laagscorders).

Bij de deoltoetsen Oversteektaak en Besluitvaardigheid zijn bovendien betrouwbaarheidsanalyses verricht volgens het Item Response Theorie-model. In bijlage 8 wordt in detail uiteengezet wat de uitgangsunten en werkwijzen hierbij zijn. In het kort komt het erop neer dat een toets (of deoltoets), naast het opleveren van een betrouwbare score over de leerling als geheel, ook:

- informatief is over leerlingen met zowel lage, middelmatige als hoge vaardigheidsniveaus. Dit kan niet bepaald worden met behulp van klassieke betrouwbaarheidsanalyses.
- Informatie levert over het absolute niveau van vaardigheid waarop leerlingen ten tijde van de meting functioneren, zodat beschreven kan worden welke verkeersopgaven leerlingen aankunnen. Ook dit kan niet met behulp van klassieke betrouwbaarheidsanalyses.

Het streven is bij het toetsen van fiets- en in een later ontwikkelingsstadium ook rijvaardigheid, opgavensets te ontwikkelen die passen bij verschillende niveaus van de leerling. Dat wil zeggen dat de opgaven informatief zijn voor een leerling op elke ontwikkelingsniveau. Om aan te tonen of we daarin geslaagd zijn hanteren we schalingstechnieken en betrouwbaarheidsanalyses uit de itemresponsetheorie. Het idee achter de itemresponsetheorie is dat het aantal opgaven dat een leerling goed maakt een juiste schatter oplevert van diens vaardigheid. Verondersteld wordt dat alle opgaven in één dimensie oplopen in moeilijkheid. Verder is het idee dat, bij een bepaald bereikt niveau van vaardigheid, alle opgaven goed worden gemaakt die onder het bereikte niveau van de leerling zitten. Vraagt een opgave meer, dan daalt de kans op goed maken snel tot 0.

Deze aannames kunnen worden geschat door een psychometrisch model aan te nemen en dat te toetsen aan de data. De uitkomsten van deze analyses geven inzicht in de vraag of het aantal opgaven correct (een besluit over een voorrangsvolgorde is juist) of anderszins gescoord (zoals een gap geaccepteerd als 'veilig') een afdoende schatter is van het niveau van een vaardigheid of attitude. Een voorwaarde is meestal dat de serie opgaven consistent oploopt in moeilijkheid (of in belasting van acceptatiegraad). Omdat niet alle toetsen exact met dit oogmerk zijn ontworpen, zijn deze analyses alleen op genoemde deoltoetsen toegepast.

Nadat de betrouwbaarheid van de deoltoetsen is bepaald, is er gekeken naar hoe leerlingen hebben gescoord door gemiddelden en standaardafwijkingen te berekenen. Ook is gekeken of er één of meerdere constructen ten grondslag liggen aan de deoltoets, door een principale componenten analyse uit te voeren met varimax rotatie. Vervolgens is gekeken naar de samenhang tussen deoltoetsen door middel van correlaties. Bij het berekenen van correlaties is gekozen voor de non-parametrische Spearman's correlatie coëfficiënt, omdat de data niet normaal verdeeld waren.

Om na te gaan of de testcores zijn beïnvloed door andere factoren dan competentie, en om na te gaan of de testcores samenhangen met fietservaring en motivatie, zijn er lineaire regressie analyses uitgevoerd. Voorspellende factoren (hierna beïnvloeders genoemd) zijn stapsgewijs opgenomen in het regressiemodel. Voor elk van deze regressies is in de resultaten vermeldt welke afhankelijke variabelen en welke beïnvloeders (voorspellers) in het model zijn opgenomen. Beïnvloeders met een significantie niveau hoger dan 0.05 zijn verwijderd uit het model. Hoewel de data niet normaal verdeeld was liet inspectie van de data zien dat de error termen bij benadering wel normaal verdeeld waren. Hierdoor kunnen regressies dus nog steeds worden gebruikt, omdat regressies robuust genoeg zijn tegen kleine schendingen (Field, 2013). Ook was er geen sprake van onacceptabele autocorrelatie of schendingen van lineariteit. Nominale variabelen werden getransformeerd tot dummy variabelen.

Tot slot is de invloed van eerder genoten verkeerseducatie, van stedelijkheid en van de gemiddelde Cito-scores van de scholen (bepaald op schoolniveau) op de testcores geschat. Vanwege sterke afwijkingen in de normaliteit (ligt ver boven de geaccepteerde +/-2 skewness,) zijn non-parametrische toetsen uitgevoerd. Een aantal beïnvloeders is daarbij niet meegenomen of verwijderd.

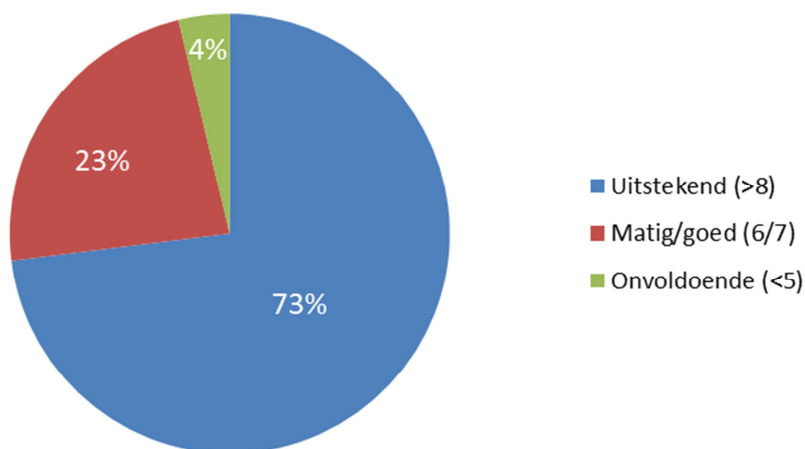
4. Resultaten

Dit hoofdstuk beschrijft de resultaten van de afname van FietsmeetlatBO 3.0. De gebruikerservaringen, psychometrische eigenschappen en invloed van externe factoren komen daarbij aan bod.

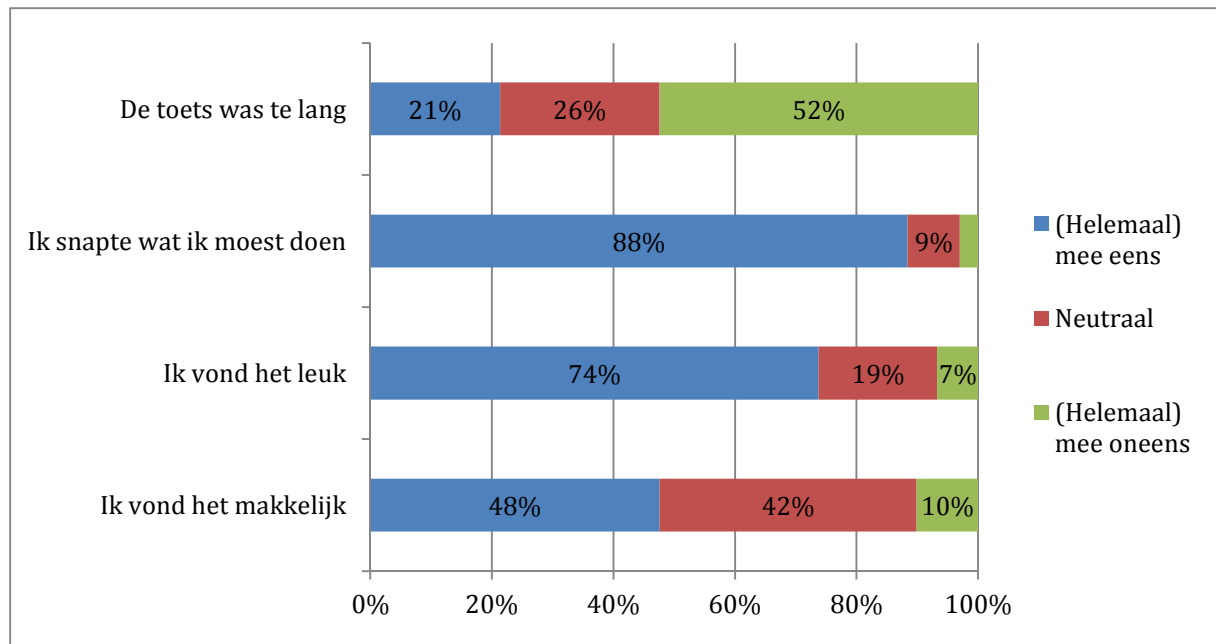
4.1 Gebruikerservaringen

Over het algemeen waren er weinig technische problemen met het laden van de toets. De kwaliteit van de beschikbare lokalen was doorgaans goed. De kwaliteit van de internetverbinding was op sommige scholen minder goed (traag of instabiel). De eerste afname van de FietsmeetlatBO 3.0 ging goed, maar nadat bij de tweede afname internetproblemen ontstonden, is er voor gekozen de onderdelen van de toets volledig offline af te nemen in plaats van gedeeltelijk online. Er is dan alleen een internetverbinding nodig voor het wegschrijven van de data naar een database. Deze overgang loste de problemen die ontstonden door een langzame internetverbinding op en deelttoetsen liepen niet meer vast.

Uit observaties van proefleiders tijdens de afname en uit de beantwoording van de vragen over de toets, blijkt dat de leerlingen over het algemeen positief waren over de FietsmeetlatBO. In Figuur 4.1 is te zien dat 73% van de leerlingen de toets een 8 of hoger geeft. Verder laat Figuur 4.2 zien dat de toets leuk (74%) en duidelijk (88%) gevonden werd. Een klein aantal leerlingen vond de toets moeilijk (10%) en te lang (21%). In de huidige vorm kostte het leerlingen ongeveer een uur om de toets (inclusief instructies en vragenlijst) in zijn geheel te doorlopen.



Figuur 4.1. Rapportcijfer voor FietsmeetlatBO 3.0 in percentages %. De vraag die gesteld werd was: Welk cijfer zou jij de toets geven?



Figuur 4.2. Beoordeling van de FietsmeetlatBO 3.0 in percentages.

4.2 Psychometrische eigenschappen en betrouwbaarheid

Per deelttoets worden in deze paragraaf de psychometrische eigenschappen besproken. Idealiter wordt gestreefd naar een toetsinstrument dat:

1. Deelttoetsen items bevat die verschillen in moeilijkheidsgraad uitdrukt in het aandeel van de leerlingen dat de opgave goed gemaakt heeft: dus in p-waarden.
2. Een hoge samenhang vertoont tussen de items (hoge Cronbach's α , factoranalyse)
3. Leerlingen kan rangordenen naar niveau op verschillende competenties (hoog onderscheidend vermogen). De R_{ir} waarde is het onderscheidend vermogen. Dit is het vermogen van een opgave om onderscheid te maken tussen leerlingen die hoog scoren en leerlingen die laag scoren. Ofwel leerlingen die hoog scoren op een toetsonderdeel, scoren consistent hoog en leerlingen die laag scoren op een toetsonderdeel, scoren consistent laag.
4. Opgebouwd is uit deelttoetsen die elk feitelijk meten dat deze bedoeld zijn te meten, namelijk een enkel onderliggend construct. Daarvoor zijn twee methoden gebruikt: de Cronbach's α en de Principale Componenten Analyse (PCA). De Cronbach's α zegt iets over de interne consistentie van de deelttoetsen (de mate waarin opgaven met elkaar samenhangen). Over het algemeen geldt dat hoe hoger de α is (maximale waarde is 1) hoe beter de samenhang tussen de items. Als grenswaarde wordt doorgaans .70 aangehouden (Field, 2009). De PCA laat zien of er verschillende dimensies zijn binnen een deelttoets. Dus of het gaat over een enkele competentie of een aantal (Field, 2009).

4.2.1 Sitatiebewustzijn

Het onderdeel Sitatiebewustzijn bestaat uit de deoltoetsen waarnemen en besluitvaardigheid. De psychometrische eigenschappen van de twee deoltoetsen worden hieronder besproken.

Waarnemen

Dit onderdeel bestaat uit 8 opgaven, afgenomen in twee boekjes. Bij elk boekje werden er 6 van de 8 opgaven afgenomen. Opgaven W2 en W7 werden alleen bij boekje 1 afgenomen, en opgaven W4 en W6 alleen bij boekje 2.

In Tabel 4.1 zijn de psychometrische eigenschappen van de deoltoets waarnemen weergegeven, op basis van een klassieke betrouwbaarheidsanalyse. Wat opvalt is het grote verschil in p-waarden tussen boekje 1 en 2 voor overlappende opgaven. Het lijkt erop dat de twee groepen die de verschillende versies hebben gemaakt (sterk) verschillen van elkaar op hun prestatie.

Opgave ¹	Boekje 1 (n=147)			Boekje 2 (n=134)		
	p-waarde	R _{ir}	Cronbach's α	p-waarde	R _{ir}	Cronbach's α (bij weglating opgave)
W2	.96	.35	.26			
W3	.80	.19	.31	.53	.29	.49
W4				.95	.23	.52
W6				.51	.34	.46
W7	.86	.16	.32			
W8	.88	.31	.23	.81	.44	.41
W9	.90	.20	.30	.96	.34	.50
W11	.26	-.03	.49	.32	.19	.54

N.B. Alpha boekje 1 = .36; Alpha boekje 2 = .53

Tabel 4.1. Psychometrische eigenschappen van de deoltoets Waarnemen.

Tevens is het opvallend dat de moeilijkheidsgraad voor het merendeel van de opgaven bij beide versies erg laag is, met p-waarden boven de .80. Dit betekent dat deze opgaven erg gemakkelijk zijn voor de leerlingen. Daarnaast kan er worden gesteld dat in het algemeen, het onderscheidend vermogen (onderscheid maken tussen hoog- en laagscorders) laag is. In boekje 1 kan, gebaseerd op de R_{ir}-waarde, alleen bij opgaven W2 en W8 onderscheid worden gemaakt tussen potentiële hoog- en laagscorders. Echter, doordat de opgaven redelijk makkelijk zijn, geldt dit onderscheidend vermogen vooral voor zwakke en nog zwakker presterende leerlingen. In boekje 2 zijn de opgaven, mede door de veel lager presterende groep leerlingen, veel meer onderscheidend voor leerlingen in die groep. Met name de kenmerken van opgave W8 zijn gunstig.

Verder is de interne consistentie van de deoltoets Waarnemen als geheel laag, hoewel dit bij boekje 2 hoger ligt dan bij boekje 1 (respectievelijk $\alpha = .36$ en $\alpha = .53$). Mede gezien deze resultaten (gemakkelijke opgaven, verschillend presterende steekproeven per boekje, lage alfa) kan er worden gesteld dat de toets Waarnemen, in de huidige vorm, voornamelijk geen betrouwbare resultaten oplevert. Onduidelijk is nog wat mogelijke oorzaken zijn hiervoor.

¹ De opgaven hier zijn een selectie van de opgaven die zijn afgenomen bij FietsmeetlatBO 2.0.

Besluitvaardigheid

In Tabel 4.2 zijn de psychometrische eigenschappen van de deelttoets besluitvaardigheid te vinden op basis van een klassieke betrouwbaarheidsanalyse. De p-waarden voor de overlappende opgaven van de twee boekjes zijn redelijk gelijk. Dit wijst erop dat de twee boekjes niet sterk verschillen op hun prestaties met betrekking tot de overlappende opgaven.

Opgave	Boekje 1 (n=127)			Boekje 2 (n=133)		
	p-waarde	R_{ir}	Cronbach's α (bij weglating van opgave)	p-waarde	R_{ir}	Cronbach's α (bij weglating van opgave)
B3	.41	.28	.67	.40	.25	.56
B4				.33	.17	.59
B5				.78	.41	.49
B6	.48	.51	.59	.58	.38	.50
B8	.91	.33	.65	.89	.49	.48
B9	.68	.56	.58			
B12	.43	.33	.67			
B13	.64	.43	.62	.70	.28	.55

Alpha boekje 1 = .67; Alpha boekje 2=.57

Tabel 4.2. Psychometrische eigenschappen van de deelttest besluitvaardigheid na een klassieke betrouwbaarheidsanalyse.

Daarnaast is in Tabel 4.2 te zien dat de p-waarden voor boekje 1 tussen .41 en .91 liggen en voor boekje 2 tussen .33 en .89 liggen, wat een goede spreiding van moeilijke en makkelijke vragen aangeeft. Ook is te zien dat over het algemeen het onderscheidend vermogen van de opgaven relatief goed is, met uitzondering van opgave B4 ($R_{ir} = .17$). Het onderscheidend vermogen van opgaven B6 en B9 van boekje 1 is zelfs hoog (beide groter dan .50). Dit geeft aan dat er over het algemeen goed onderscheid kan worden gemaakt tussen hoog-en laagscorders.

Opgave	M	SD
B8	7.24	6.55
B5	8.81	6.72
B13	10.82	8.24
B3	11.99	9.59
B9	15.48	10.30
B12	17.77	12.79
B6	18.22	11.85
B4	20.72	13.96
Gemiddeld	13.88	10.00

Tabel 4.3. Antwoorttijden op de opgaven van de deelttest Besluitvaardigheid in seconden

In Tabel 4.3 zijn de antwoorttijden op de afzonderlijke opgaven van de deelttest besluitvaardigheid vermeld. Duidelijk is dat de antwoorttijden omgekeerd samenhangen met de moeilijkheidsgraad: hoe moeilijker de opgave, des te langer de antwoorttijd.

Verder is de interne consistentie (op basis van correcte respons) van boekje 1 bijna acceptabel ($\alpha = .67$, minimum = .70), terwijl die van boekje 2 relatief laag is ($\alpha = .57$). Om tot een acceptabele betrouwbaarheid te komen, zou de test verlengd moeten worden tot 12 à 13 vergelijkbare opgaven.

Uit Tabel 4.4 kan worden afgelezen dat het itemresponsemodel overall goed past op de data, nadat voor alle opgaven verschillende discriminatieparameters zijn geschat, die rekening houden met het onderscheidend vermogen van een opgave.

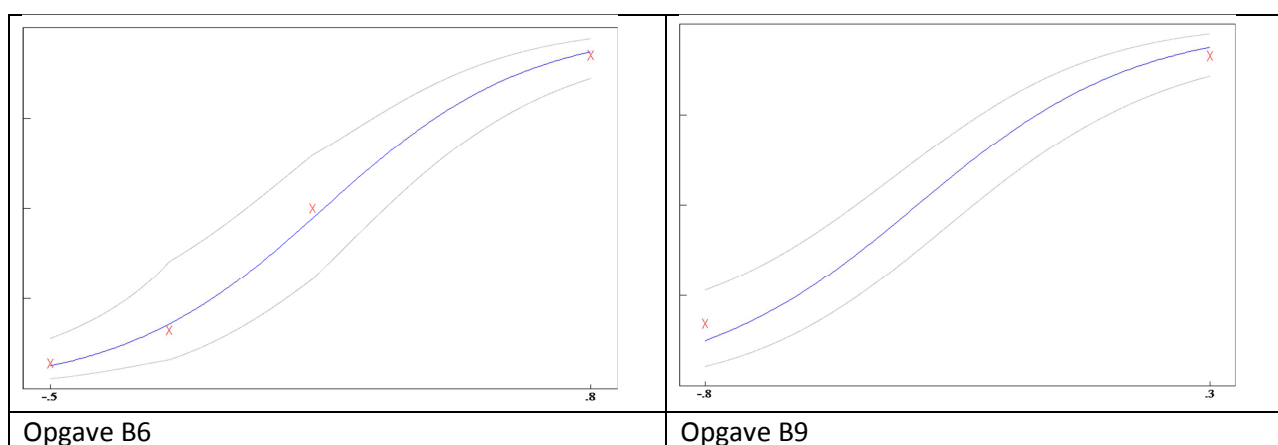
In zijn algemeenheid betekent dit dat de verschillende besluitsituaties goed schalen op de mate van besluitvaardigheid. Het aantal correct beantwoorde besluitopgaven kan daarmee worden gebruikt als een afdoende schatter van de besluitvaardigheid van de leerling.

Item	Label	A-parameter	Beta	SE(B)	S-statistiek	df	p	M-statistiek
5	B8	3	-0.926	0.102	1.515	1	0.218	
1	B9	4	-0.333	0.085	3.642	1	0.056	1.733
8	B5	4	-0.211	0.083	0.010	1	0.920	-0.946
6	B13	3	-0.037	0.102	0.000	0	NB	
4	B6	4	0.168	0.053	0.552	3	0.907	-0.224
2	B12	2	0.320	0.099	3.885	1	0.049	1.971
3	B3	2	0.402	0.073	16.565	5	0.005	0.052
7	B4	2	0.616	0.104	1.671	1	0.196	

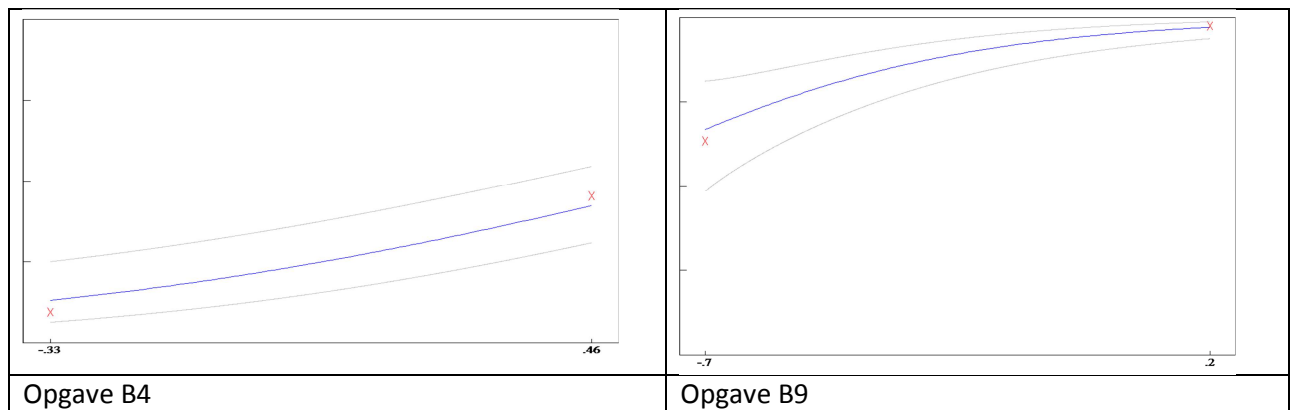
Tabel 4.4. Calibratie van een één-parameter itemresponsemodel op itemscores van de deeltests Besluitvaardigheid (opgaven geordend naar moeilijkheidsgraad, bèta).

N.B. Overall Fit IRT-model: $R^2=0.17$; $df=21$; $p=0.6949$. P-waarde: kans dat voorspelde data IRT-model afwijken van oorspronkelijke data. Measurement of accuracy (verwachte alpha-betrouwbaarheidscoëfficiënt) Boekje 1: .74; Boekje 2: .62.

Figuur 4.3 en Figuur 4.4 geven een beeld van de best en minst onderscheidende opgaven in de deelttest en zijn een illustratie van wat uit Tabel 4.4 kan worden opgemaakt. De twee best onderscheidende opgaven (hoge a-waarden) laten een steile itemsresponsecurve zien, wat inhoudt dat bij deze opgaven een duidelijk omslagpunt is te zien bij leerlingen met een bepaald niveau van besluitvaardigheid. Opgaven B6 en B9 zijn hierbij voorbeeldige items.



Figuur 4.3. Geobserveerde proporties correct besluit en de verwachte kansproportie (Y-as) op grond van verschillende niveaus van besluitvaardigheid (geschatte θ op de X-as), voor de twee best onderscheidende items van de deelttest Besluitvaardigheid.



Figuur 4.4. Geobserveerde proporties correct besluit en de verwachte kansproportie (Y-as) op grond van verschillende niveaus van besluitvaardigheid (geschatte θ op de X-as), voor de twee minst onderscheidende items van de deelttest Besluitvaardigheid.

De twee minst onderscheidende opgave laten bij verschillende vaardigheidsniveaus slechts weinig verschil zien in de proportie correcte antwoorden. Opgave B4 en B9 dragen relatief weinig bij aan het onderscheid in vaardigheid tussen leerlingen, ofschoon er nog wel sprake is van eng onderscheidend vermogen bij deze opgaven.

4.2.2 Beslissen in complexe situaties

Dode hoek taak

In Tabel 4.5 staan de psychometrische eigenschappen van de dode hoek taak. Opgave 6 is bedoeld als basis opgave en gaat om inzicht in dode hoeken rondom een vrachtwagen.

Zoals verwacht scoren leerlingen beter op simpele opgaven dan op complexe opgaven. Alleen opgave 4 maakt voldoende onderscheid tussen hoogscoorders en laagscoorders. De interne consistentie van de onderdelen is laag met een α van 0.23. Een principale componentenanalyse (Tabel 4.6) laat zien dat de opgaven geen eenduidige schaal vormen. De opgaven laden op 3 verschillende componenten. Dit is te verklaren doordat de dode hoek taak uit drie verschillende types opgaven bestaat, namelijk: vrachtwagens en verkeerslichten (opgave 1 en opgave 2), fietsen rondom een vrachtwagen (opgave 4 en opgave 5), en het herkennen van dode hoeken (opgave 6).

Opgave	Scoring	Gem. (SD)	R_{ir}	α
Basis Opgave 6	0 t/m 6	3.76 (1.26)		.08
Simpele Opgave 1	0 t/m 3	1.94(.87)		
Opgave 3	goed/fout	.56(.50)		
Opgave 4	goed/fout	.86(.35)		
Complex Opgave 2	goed/fout	.52(.50)		.26
Opgave 5	0 t/m 2	.65(.87)		
Opgave 6	goed/fout	.24(.43)		
Totaal	0 t/m 11	.41(.71)		.19
		6.14 (1.91)		
				.23

Tabel 4.5. Psychometrie van de dode hoek taak

Factorloadingen	1(25.23%)	2(18.22%)	3(17.10%)
Opgave 1	-.01	.84	.21
Opgave 2	.07	.55	-.47
Opgave 3	.47	.28	-.10
Opgave 4	.79	-.01	.18
Opgave 5	.78	-.08	-.03
Opgave 6	.07	.11	.87

Tabel 4.6. Factorloadingen van de dode hoek taak

Oversteektaak

Op de data van de oversteektaak zijn twee analyses uitgevoerd. Naast de reguliere betrouwbaarheidsanalyse zoals ook voor de andere onderdelen is gedaan, is een OPLM analyse verricht. Voor de reguliere analyse is er een vast interval gekozen waarbinnen de leerlingen aangeven of de gap (het 'tijdgat' waarin kan worden overgestoken) veilig of onveilig is en wordt er gescoord het antwoord goed of fout is. In de OPLM analyse is er gekeken naar wat leerlingen een acceptabele gap vinden. Beide analyses worden hier beschreven.

In Tabel 4.7 staan de psychometrische eigenschappen van de oversteektaak. Er valt op dat de interne consistentie van de onderdelen laag is met een alpha van 0.37. Ook een principale componentenanalyse (Tabel 4.8) laat zien dat de opgaven geen eenduidige schaal vormen. Er zijn groepen met veilige/onveilige gaps die heel subtiel verschillen. Bijvoorbeeld onveilige gaps van 3 seconden laden op één factor, en onveilige gaps van 4 seconden laden op een andere factor. Doordat deze verschillen echter subtiel zijn is er een aanvullende OPLM analyse uitgevoerd.

Opgave	Scoring	Gem. (SD)	p	R _{ir}	α
Situatie 1	0 t/m 4	2.34(.88)			
Gap 1	goed/fout		.31	.31	
Gap 2	goed/fout		.82	-.20	
Gap 3	goed/fout		.90	.34	
Gap 4	goed/fout		.31	.15	
Situatie 2	0 t/m 4	2.12(.64)			
Gap 1	goed/fout		.91	-.21	
Gap 2	goed/fout		.98	.17	
Gap 3	goed/fout		.14	.27	
Gap 4	goed/fout		.23	.28	
Situatie 3	0 t/m 3	2.30(.79)			
Gap 1	goed/fout		.85	-.15	
Gap 2	goed/fout		.80	.24	
Gap 3	goed/fout		.65	-.07	
Situatie 4	0 t/m 4	2.32(.90)			
Gap 1	goed/fout		.13	.14	
Gap 2	goed/fout		.90	.21	
Gap 3	goed/fout		.56	.35	
Gap 4	goed/fout		.90	-.23	
Situatie 5	0 t/m 3	2.38(.74)			
Gap 1	goed/fout		.82	.36	
Gap 2	goed/fout		.79	-.43	
Gap 3	goed/fout		.89	-.05	
Situatie 6	0 t/m 3	2.02(.56)			
Gap 1	goed/fout		.27	.18	
Gap 2	goed/fout		.91	.07	
Gap 3	goed/fout		.89	-.32	
Situatie 7	0 t/m 3	1.91(.84)			
Gap 1	goed/fout		.77	.37	
Gap 2	goed/fout		.90	.15	
Gap 3	goed/fout		.47	.27	
Totaal	0 t/m 24	15.30(2.89)			

.37

Tabel 4.7. Psychometrie Beslissen in complexe situaties onderdeel B 'oversteken'. Per gap wordt de scoring aangegeven, de p-waarde en het onderscheidend vermogen. Per situatie, die bestaat uit verschillende gaps, wordt de scoring, het gemiddelde met standaard deviatie tussen haakjes weergegeven. Cronbach's alpha over de gehele toets wordt weergegeven.

Factorloadingen	1 (23,19%)	2 (9,60%)	3 (7,15%)	4 (6,38%)	5 (5,43%)	6 (5,16%)	7 (4,52%)
opgave1_g1	.33	.61	-.03	-.11	.03	.44	.07
opgave1_g2	-.08	-.72	-.11	.09	.11	-.02	.13
opgave1_g3	.04	-.03	.08	.06	.72	.08	-.04
opgave1_g4	.68	.30	-.12	.02	-.01	.05	.10
opgave2_g1	-.09	.05	-.04	-.02	.01	-.05	-.84
opgave2_g2	-.05	-.22	.78	.17	.19	-.22	.10
opgave2_g3	.86	-.07	.17	-.07	.01	.15	-.07
opgave2_g4	.45	.37	-.14	-.15	.26	-.16	.41
opgave3_g1	-.05	-.24	.22	.62	.08	-.13	-.27
opgave3_g2	-.02	.16	.04	.19	.11	.69	.06
opgave3_g3	-.08	.02	-.02	.64	-.07	.23	.07
opgave4_g1	.61	.14	-.08	-.39	.10	-.23	.25
opgave4_g2	-.08	-.07	.09	.02	.65	-.06	.22
opgave4_g3	.20	.45	-.04	-.30	.29	.24	.35
opgave4_g4	-.03	-.65	.38	-.05	-.07	-.06	-.09
opgave5_g1	.20	.25	-.07	-.09	.57	.18	-.16
opgave5_g2	-.66	-.27	.13	.07	-.18	.14	-.24
opgave5_g3	-.28	-.06	.30	.54	.19	-.32	.09
opgave6_g1	.40	.05	-.24	-.10	.07	.41	.42
opgave6_g2	-.06	.08	.65	.37	.08	-.09	-.30
opgave6_g3	-.64	.05	.24	.20	.14	-.40	.11
opgave7_g1	.04	.08	.06	-.36	.44	.54	.00
opgave7_g2	-.09	-.004	.78	-.08	-.04	.24	.04
opgave7_g3	.24	.404	-.01	-.40	.27	.07	.01

Tabel 4.8. Factorloadingen oversteektaak

In

Tabel 4.9 staan de resultaten van de OPLM-analyses op de intervalveiligheidsinschattingen door leerlingen bij de verschillende aangeboden 'gaps' in de oversteeksituaties. Hierbij is een inschatting veilig als 1 gescoord en 0 als onveilig. De opgaven zijn hierbij geordend naar de Bèta-parameter, die de moeilijkheidsgraad aanduidt, wat in dit geval de graad belasting van de gap-acceptatie betekent.

Het itemresponsemodel past overall perfect op de data, nadat voor alle opgaven discriminatieparameters zijn geschat, die rekening houden met het onderscheidend vermogen van een opgave.

In zijn algemeenheid betekent dit dat de verschillende oversteekintervallen bijna perfect schalen op de mate van ingeschatte veiligheid naar afnemende intervaltijd. Het aantal als veilig geaccepteerde gaps (tijdsintervallen) kan daarmee worden gebruikt als een afdoende schatter van de gapacceptatie door de leerling. Tevens kan de score worden gerelateerd aan de situatie waarin een leerling een interval nog net veilig acht om over te steken. Voor een belangrijk deel zal deze inschatting te maken hebben met de duur van het interval, maar mogelijk ook met de aard van het volgende voertuig. Dit laatste is nog niet onderzocht, maar kan via een combinatie van situatiekenmerken en itemparameters worden achterhaald.

Verder kan naar aanleiding van

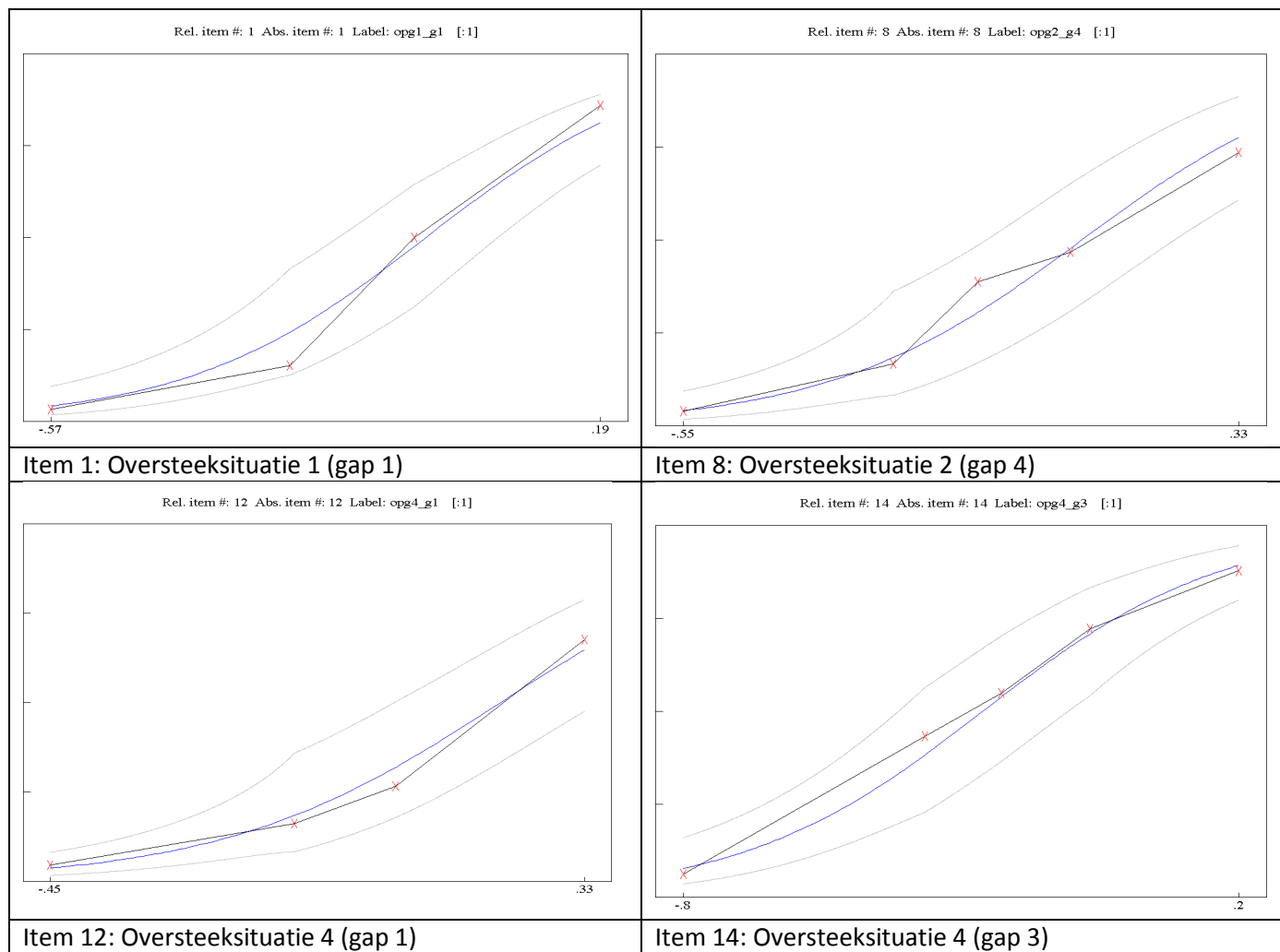
Tabel 4.9 opgemerkt worden dat de betrouwbaarheid van de IRT-geschaalde items hoog is (.88). Deze waarde is te beschouwen als de verwachte alpha-coëfficiënt, die gebruikt wordt bij betrouwbaarheidsanalyses volgens de klassieke testtheorie.

Betalingsaanpak analyses volgens de klassieke testtheorie								
Item	Label	A- parameter	Bèta	SE(B)	S- statistiek	df	p	M- statistiek
13	opg4_g2	1	-2.41	0.198	2.409	3	0.492	0.94
3	opg1_g3	1	-2.22	0.185	6.501	4	0.165	-0.03
10	opg3_g2	2	-0.95	0.082	6.309	5	0.277	-0.24
16	opg5_g1	4	-0.77	0.052	1.303	3	0.728	-0.20
22	opg7_g1	3	-0.57	0.053	4.964	5	0.42	-1.79
14	opg4_g3	5	-0.27	0.037	0.894	4	0.925	0.04
24	opg7_g3	4	-0.14	0.041	2.569	5	0.766	-0.52
1	opg1_g1	6	-0.05	0.035	2.930	3	0.403	-1.11
4	opg1_g4	4	-0.04	0.042	4.716	5	0.452	-0.12
8	opg2_g4	5	0.08	0.039	1.415	4	0.842	0.27
19	opg6_g1	4	0.08	0.044	3.538	4	0.472	0.11
17	opg5_g2	4	0.17	0.046	2.891	4	0.576	-1.15
12	opg4_g1	5	0.21	0.043	0.811	3	0.847	-0.68
7	opg2_g3	4	0.33	0.052	6.312	3	0.097	0.26
21	opg6_g3	4	0.40	0.055	0.395	2	0.821	-1.11
9	opg3_g1	4	0.41	0.056	0.409	2	0.815	0.95
15	opg4_g4	3	0.46	0.065	1.435	3	0.697	0.75
18	opg5_g3	3	0.46	0.065	0.387	3	0.943	-0.68
23	opg7_g2	3	0.47	0.066	2.679	3	0.444	1.15
20	opg6_g2	4	0.49	0.061	0.468	2	0.792	-0.61
11	opg3_g3	1	0.55	0.134	3.511	6	0.742	-0.62
2	opg1_g2	2	0.68	0.090	2.956	4	0.565	0.00
5	opg2_g1	2	1.02	0.113	1.558	2	0.459	0.85
6	opg2_g2	2	1.62	0.186	0.000	0		
Overall Fit IRT-model								
Aantal items	Aantal groepen	Aantal deviaten getest		Kwadratenso m		R1c	Df	P
24	4	92		43,17		41,5	69	0,997

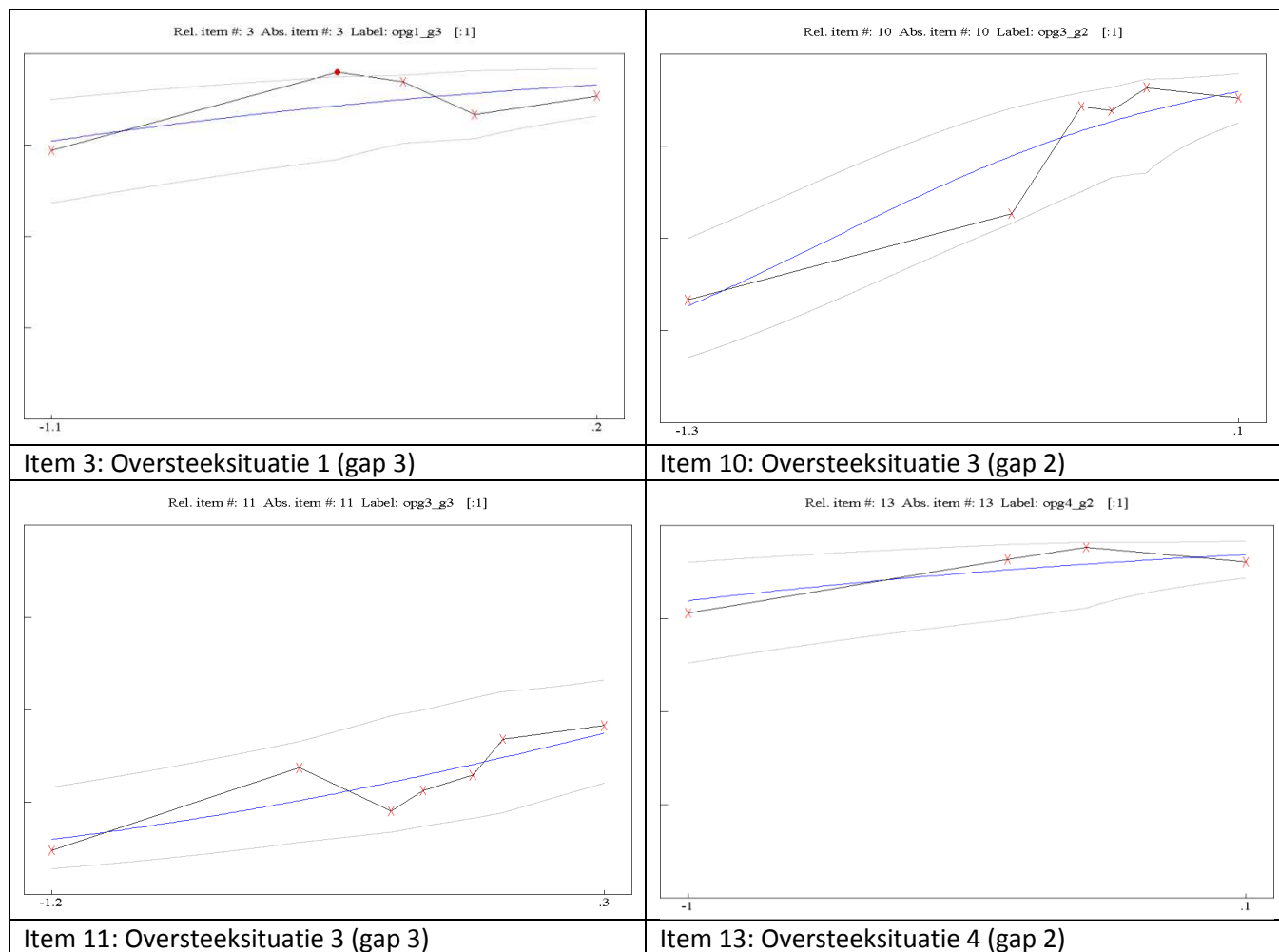
Tabel 4.9. Calibratie van een één-parameter itemresponsemodel op de intervalveiligheidsinschattingen van de oversteektaak (opgaven geordend naar toenemende belasting van gapacceptatie). *Noot.* Measurement of accuracy (verwachte alpha) = .88; P-waarde: kans dat voorspelde data IRT-model afwijken van oorspronkelijke data.

Figuur 4.5 en Figuur 4.6 geven een beeld van de best en minst onderscheidende opgaven in de deeltoets en zijn een illustratie van wat uit

Tabel 4.9 kan worden opgemaakt. De vier best onderscheidende opgaven (hoge a-waarden) laten een steile itemsresponsecurve zien, wat inhoudt dat bij deze opgaven een duidelijk omslagpunt is te zien bij leerlingen met een bepaald niveau van gap-acceptatie. Item 1 en 14 zijn hierbij voorbeeldige items (Gap 1 van oversteeksituatie 1). De vier minst onderscheidende opgave laten bij verschillende gap-acceptatieniveaus slechts weinig verschil zien in de proportie acceptatie als veilig. Item 3 (Gap 1 van oversteeksituatie 3) en 13 (Gap 2 van oversteeksituatie 4) zijn hiervan duidelijke voorbeelden en dragen dan ook weinig bij aan het onderscheid tussen leerlingen.



Figuur 4.5. Geobserveerde proporties 'oordeel veilig' en de verwachte kansproportie (Y-as) op grond van verschillende niveaus van gapacceptatie (geschatte θ op de X-as), voor de vier best onderscheidende items van de deelttest Oversteken.



Figuur 4.6. Geobserveerde proporties 'oordeel veilig' en de verwachte kansproportie (Y-as) op grond van verschillende niveaus van gapacceptatie (geschatte θ op de X-as), voor de vier minst onderscheidende items van de deelttest Oversteken

In Tabel 4.10 is de omrekening van ruwe score naar latente score en getransformeerde latente score weergegeven. Deze laatste score is gekozen, maar zou ook een andere willekeurige lineaire transformatie van de geschatte latente score kunnen zijn, afhankelijk van de wensen van de testgebruiker. Gekozen is voor een score met een gemiddelde rondom de 100 en een standaard deviatie van 22. Hoge scores vertegenwoordigen acceptatie van zeer kleine intervallen bij oversteken, lage scores daarentegen uitsluitend grote intervallen.

Ruwe gapscore	Latente score	Getransformeerde latente score (intervalacceptatie)
0	-3.89	33
1	-2.51	57
2	-1.43	76
3	-0.98	84
4	-0.72	88
5	-0.54	92
6	-0.40	94
7	-0.29	96
8	-0.19	98
9	-0.11	99
10	-0.03	100
11	0.04	102
12	0.11	103
13	0.18	104
14	0.25	105
15	0.33	107
16	0.41	108
17	0.49	110
18	0.59	111
19	0.70	113
20	0.83	116
21	1.02	119
22	1.30	124
23	1.81	133
24	2.35	142
Gemiddelde	-0.03	100.51
SD	1	22

Tabel 4.10. Omzettingstabel van de gapscore naar schaal, van 33 tot 142

	Pearson Correlatie	N
Totaalscore deelttest gevaarherkenning	-.12	238
Totaalscore deelttest Waarnemen	-.13*	244
Totaalscore deelttest Besluitvaardigheid	-.07	244
Totaalscore deelttest Sociaal moreel handelen (mate waarin leerling niet-sociaal fietsgedrag vertoont)	.23**	244
Totaalscore deelttest game	-.04	245
Eigen overtrekend fietsgedrag		
Hoe vaak in de afgelopen maand de hand niet uitgestoken wanneer je afsloeg.	.13*	239
Groepsdruk		
Groepsdruk: Als anderen door rood fietsen, fiets ik meestal ook door rood	.11	239
Groepsdruk: Als anderen met z'n drieën naast elkaar willen fietsen, vind ik het flauw om erachter te blijven fietsen	.13*	239
Voorbeeldgedrag belangrijke anderen		
Rijdt een van je ouders wel eens zonder gordel om in de auto?	.19**	239
Rijdt een van je ouders wel eens met de fiets of met de auto na het drinken van teveel alcohol?	-.14*	239
Rijdt een van je ouders wel eens met de fiets of met de auto door rood?	.12	239
Rijden je vrienden wel eens met de fiets of met de auto door rood?	.12	239
Rijdt je broer of /zus wel eens met de fiets of met de auto door rood?	.13*	239

Tabel 4.11. Correlaties tussen gaptaakscore en overige variabelen (alleen andere deelttests en enkele variabelen met bijna significante verbanden)

In Tabel 4.11 zijn correlaties weergegeven tussen de latente gapacceptatiescore en de scores op de overige deelttoetsen. De negatieve correlatie met Gevaarherkenning, -.12, is in de verwachte richting maar net niet significant. De correlatie met de score op waarnemen, -.13 (aandacht richten op een gevaar in een verkeerssituatie) is net wel significant. Er is een positief verband ($r = .23$) met de mate waarin de leerling niet-sociaal fietsgedrag vertoont. Verder is er ook een positief verband ($r = .19$) met gordelovertradingen van de ouders, als indicator voor negatief voorbeeldgedrag. Ook door rood rijden door broer of zus hangt positief samen met gapacceptatie door de leerling ($r = .13$). Met alcohol op rijden (fiets of auto) door de ouders hangt juist negatief samen met gap-acceptatie door de leerling. De correlaties tussen de andere indicatoren van voorbeeldgedrag en gapacceptatie hebben een positief teken, maar zijn net niet significant.

4.2.3 Sociaal en moreel handelen

Het onderdeel sociaal en moreel handelen bestaat uit 10 opgaven. Voor dit onderdeel zijn er geen goede of foute antwoorden. De score geeft een indicatie van het niveau van sociaal en moreel handelen. De scores van sociaal en moreel handelen zijn gespiegeld voor de psychometrische analyses, zodat de resultaten beter te interpreteren zijn. Dit betekent dat een leerling met een hoge score (10) socialer en moreel handelt dan een leerling met een lage score (0).

In tabel Tabel 4.12 zijn de psychometrische eigenschappen te vinden voor het onderdeel Sociaal en Moreel Handelen. Hieruit blijkt dat alle vragen goed bijdragen aan de totaalscore voor dit onderdeel. Met een gemiddelde score van 7.55 kan gezegd worden dat de leerlingen redelijk sociaal en moreel handelen in het verkeer. Opgave 2 en opgave 10 hebben hoge scores: het gaat hier over het oversteken van een spoorweg terwijl de rode lampen branden en het beantwoorden van berichten op de telefoon tijdens het fietsen.

	Itemscore		R_{ir}	α	Antwoordtijd	
	Gem	SD			Gem	SD
				Als item wordt weggelaten		
Opgave 1: Met z'n drieën naast elkaar fietsen	6.88	2.66	.53	.76	15.98	12.23
Opgave 2: Spoorweg overfietsen als de rode lampen al knipperen	9.72	1.16	.29	.78	7.95	4.65
Opgave 3: Over de stoep fietsen om af te snijden	6.21	2.79	.54	.75	8.94	7.8
Opgave 4: Doorfietsen als een oudere voetganger oversteekt op een andere locatie dan bij het zebrapad	6.40	2.99	.40	.77	10.87	7.09
Opgave 5: Voor een voetganger langsfietsen die oversteekt bij een zebrapad	8.51	2.31	.46	.77	12.64	7.38
Opgave 6: Fietsen zonder licht	6.65	3.17	.43	.77	14.76	10.58
Opgave 7: Afslaan zonder richting aangeven	6.64	3.19	.44	.77	12.40	7.15
Opgave 8: Tegen de richting in rijden op een rotonde	8.76	2.33	.43	.77	17.49	9.94
Opgave 9: Fietsen op een voetpad	6.59	3.07	.59	.75	13.15	10.40
Opgave 10: Telefoon gebruiken tijdens het fietsen	9.18	1.80	.45	.77	9.61	10.23

Cronbach's alpha deelttest: .78

Tabel 4.12. Psychometrische eigenschappen deelttest sociaal en moreel handelen (n=283)

In de laatste twee kolommen van Tabel 4.12 zijn de reactietijden opgenomen. Gemiddeld hadden de leerlingen 12.37 seconden nodig om de opgave te beantwoorden. De leerlingen hebben de meeste tijd nodig om opgave 8 te maken. Het gaat hierbij om de vraag hoe vaak ze tegen de rijrichting in rijden op een rotonde. Het juist interpreteren van de afbeelding kan een oorzaak zijn dat leerlingen hier meer tijd voor nodig hebben.

Verder is de interne consistentie van het onderdeel Sociaal en Moreel Handelen acceptabel ($\alpha = .78$). Dit betekent dat de scores van dit onderdeel betrouwbare resultaten opleveren.

Aanvullend is er nog een principale componentenanalyse uitgevoerd, om te kijken of de verschillende items op gemeenschappelijke constructen laden (zie Tabel 4.13). Na varimax-rotatie resteren 3 factoren. Dit lijkt erop te wijzen dat het huidige Sociaal en Moreel Handelen geen éédimensionaal construct is. Factor één lijkt te wijzen op het eigen gedrag. Factor twee lijkt te wijzen op directe interacties met andere weggebruikers. Opgave twee staat meer op zichzelf en lijkt te wijzen op een derde factor die iets zegt over het (directe) ongevalsrisico dat het voorgestelde gedrag met zich meebrengt.

	1 (34,90%)	2 (10,35%)	3 (10,07%)
Opgave 3: Over de stoep fietsen om af te snijden	.787		
Opgave 9: Fietsen op een voetpad	.688	.378	
Opgave 1: Met z'n drieën naast elkaar fietsen	.663		
Opgave 10: Telefoon gebruiken tijdens het fietsen	.608		.406
Opgave 7: Afslaan zonder richting aangeven	.570		
Opgave 8: Tegen de richting in rijden op een rotonde	.501		
Opgave 5: Voor een voetganger langsfietsen die oversteekt bij een zebrapad		.456	.498
Opgave 6: Fietsen zonder licht		.712	
Opgave 4: Doorfietsen als een oudere voetganger oversteekt op een andere locatie dan bij het zebrapad		.808	
Opgave 2: Spoorweg overfietsen als de rode lampen al knipperen			.837

Tabel 4.13. Factorladingen deelttest sociaal en moreel handelen na een principale componentenanalyse (tussen haakjes de percentages verklaarde variantie, gerelateerd aan de factor). Criterium weglaten factorlading is .35.

4.2.4 Gevaarherkenning

Het onderdeel gevaarherkenning bestaat uit 18 opgaven. Het onderdeel werd in twee versies afgenomen bestaande uit 14 opgaves in elke versie. Opgaves 3, 5 en 8 werden alleen bij versie 1 afgenomen, en opgaves 12, 14 en 18 alleen bij versie 2. In Tabel 4.14 wordt een overzicht gegeven van de psychometrische resultaten van de afname van de gevaarherkenningstoets.

Wanneer we kijken naar de gemiddelde scores van de twee versies valt direct op dat het onderdeel gevaarherkenning moeilijk was. Al met al is de gemiddelde score laag, namelijk 3.34 voor versie 1 en 3.42 voor versie 2, van de 14 punten die te behalen waren. De gemiddelde scores op opgave 8 van versie 1 en opgave 1 van versie 2 zijn opvallend laag. Dit betekent dat weinig leerlingen deze twee opgaven goed hadden. Het onderscheidend vermogen is echter voor alle opgaves goed. De interne consistentie van de deelttoets gevaarherkenning is goed met een alpha van 0.83 voor beide versies.

	V1 (n=164)	V2 (n=149)	V1	V2	V1	V2	V1	V2
Opgave	Gem. (SD)	Gem. (SD)	R ^{ir}	R ^{ir}	α	α	Oordeel gevaar	Oordeel gevaar
Opgave 1	.16(.17)	.02(.06)	.40	.26			1.01	.89
Opgave 3	.15(.25)	-	.43	-			.81	-
Opgave 4	.16(.22)	.16(.25)	.21	.43			.70	.66
Opgave 5	.42(.31)	-	.27	-			1.24	-
Opgave 6	.38(.29)	.26(.25)	.47	.48			.96	.72
Opgave 7	.32(.23)	.30(.23)	.48	.51			1.31	1.16
Opgave 8	.05(.14)	-	.38	-			.33	-
Opgave 9	.22(.28)	.23(.27)	.52	.59			.63	.50
Opgave 10	.31(.30)	.29(.32)	.46	.49			1.12	.87
Opgave 11	.30(.21)	.31(.21)	.69	.69			1.34	1.24
Opgave 12	-	.20(.27)	-	.34			-	1.23
Opgave 13	.42(.22)	.46(.27)	.63	.49			1.41	1.39
Opgave 14	-	.19(.26)	-	.31			-	1.17
Opgave 15	.21(.23)	.32(.28)	.59	.70			.70	.68
Opgave 16	.28(.24)	.47(.27)	.61	.50			1.20	1.15
Opgave 17	.17(.22)	.30(.26)	.42	.34			.66	.64
Opgave 18	-	.47(.20)	-	.43			-	1.67
Totaal gem. (SD)	3.34(1.79)	3.42(1.83)					.96	1.00
					.83	.83		

Tabel 4.14. Psychometrie gevaarherkenningstoets. Gemiddelde scores (0 t/m 1) per opgave op de toets met standaard deviatie tussen haakjes, het onderscheidend vermogen per opgave, en beoordeling van gevaar (0 t/m 2) en Cronbach's alpha over de gehele toets

4.2.5 Samenhang deoltoetsen

Om te onderzoeken in hoeverre prestaties van leerlingen op de verschillende deoltoetsen onderling samenhangen zijn correlaties berekend. Wanneer een correlatie de waarde 1 nadert betekent dat, dat voor leerlingen geldt dat als ze hoog scoren op de ene deoltoets zij ook hoog scoren op de andere, en andersom. Een negatieve waarde betekent juist dat leerlingen over het algemeen laag op de ene toets maar hoog op de andere scoren. Een hoge samenhang kan betekenen dat er een gezamenlijke onderliggende factor is, bijvoorbeeld fietservaring. Maar ook kan het betekenen dat de deoltoetsen dezelfde competenties aanspreken. Op grond van correlaties alleen is het niet mogelijk de oorzaak van de gevonden samenhang te bepalen. In Tabel 4.15 is te zien dat de oversteektaak positief samenhangt met de deoltoets sociaal en moreel handelen: kinderen die hoog scoren op de oversteektaak, scoren ook hoog op sociaal en moreel handelen. Deze samenhang is echter niet sterk. Tussen de andere deoltoetsen is er geen verband.

	1. GH	2a. CG	2a. LG	3. DHk	4. SOC	5. WN
1. Gevaarherkenning (GH)						
2a. Oversteektaak (correcte gap; CG)	-.07					
2a. Oversteektaak (geaccepteerde gaplengte (LG))	-.10	.46**				
3. Dode hoek taak (DH)	.04	-.00	-.03			
4. Sociaal en moreel handelen (SOC)	-.15*	.18**	.21**	-.01		
5. Waarnemen (WN)	-.01	-.02	-.08	.09	.03	
6. Besluitvaardigheid (BS)	.10	-.02	-.07	.11	-.01	.01

Tabel 4.15. Correlaties tussen de deoltoetsen. Noot. *p<0.05 **p<0.01

4.3 Beïnvloeding testcores

Om na te gaan welke externe factoren de testcores beïnvloeden zijn er correlaties en lineaire regressie analyses uitgevoerd. De achtergrondvariabelen waar naar gekeken wordt zijn motivatie voor de test (zie bijlage 5), fietservaring (zie bijlage 6), Cito-score, verkeerseducatie (zie bijlage 4) en stedelijkheid. Bevindingen zijn in onderstaande tabel weergegeven, en in het vervolg van deze paragraaf per aspect verder uitgewerkt.

		Deoltoetsen					
		GH	OT	DHT	SOCM	WN	BV
Factoren	Motivatie	Niet sig.	Sig.	Niet sig.	Sig.	Sig.	Niet sig.
	Fietservaring	Sig.	Sig.	Sig.	Sig.	Sig.	Sig.
	Cito	Niet sig.	Niet sig.	Niet sig.	Niet sig.	Niet sig.	Niet sig.
	Verkeerseducatie	Sig.	Niet sig.	Niet sig.	Sig.	Niet sig.	Sig.
	Stedelijkheid	Niet sig.	Sig.	Niet sig.	Niet sig.	Niet sig.	Sig.

Tabel 4.16. Overzicht van de gevonden resultaten per deoltoets. N.B. GH = Gevaarherkenning, OT = Oversteektaak, DHT = Dode Hoek Taak, SOCM = Sociaal en Moreel Handelen, WN = Waarnemen, BV = Besluitvaardigheid.

4.3.1 Motivatie

In Tabel 4.17 wordt een overzicht gegeven van de correlaties tussen de motivatie van leerlingen en de scores op de verschillende deoltoetsen. Voor de deoltoetsen gevaarherkenning, dode hoek taak en besluitvaardigheid zijn er geen verbanden tussen motivatie en prestatie op de taak. De significante verbanden worden verder beschreven.

	1. GH	2a. CG	2a. LG	3. DHk	4. SOC	5. WN	6. BS
Ik vond het makkelijk	.09	-.01	.00	-.02	.04	.14*	-.04
Ik vond het leuk	.09	.12	-.06	.04	-.23**	.10	.01
Ik snapte wat ik moest doen	.12	.06	-.07	-.06	-.10	.07	.09
De toets was te lang	-.07	-.14*	.13*	-.10	.09	-.02	-.09
Welk cijfer zou jij de toets geven?	.08	.04	-.10	-.01	-.14*	.06	.04

Tabel 4.17. Correlaties tussen motivatie en scores op de deoltoetsen. N.B. *p<0.05 **p<0.01; GH= Gevaarherkenning; CG= correcte gaps Oversteektaak; LG= Lengte gacepeerde gaps Oversteektaak; DHk= Dode hoek taak; WN= waarnemen; BS= Besluitvaardigheid

Waarnemen

De toets makkelijk vinden heeft een significant verband met de deelttoets waarnemen. Naarmate leerlingen hoger scoren op waarnemen vinden ze de toets als geheel gemakkelijker.

Oversteektaak

De mate waarin leerlingen de toets leuk of te lang vonden hangt samen met de behaalde score 'correcte gapkeuzes' op de oversteektaak. Leerlingen die de toets leuk vonden scoren hoger op de oversteektaak en leerlingen die de toets te lang vonden scoren lager op de oversteektaak. Daarentegen, naarmate leerlingen kortere oversteekintervallen (gaps) accepteren bij de oversteektaak, vinden ze de toets minder leuk.

Sociaal en moreel handelen

Naarmate leerlingen minder sociaal handelen vinden ze de toets minder leuk ($r = -.23$) en geven ze de toets een wat lager cijfer.

4.3.2 Fietservaring

In deze paragraaf wordt de invloed van fietservaring en fietsomstandigheden op de testcores beschreven. De vragenlijst omvat vragen waarmee leerlingen hun fietservaring omschrijven en in welke omstandigheden zij fietsen. Lineaire regressieanalyses zijn uitgevoerd om te bepalen wat significante voorspellers zijn. Significante verbanden² worden beschreven per deelttoets.

Waarnemen

De resultaten van de regressie analyse tonen aan dat de leeftijd waarop leerlingen leren fietsen een significante voorspeller is voor de score op waarnemen, 1% van de variantie wordt verklaard door deze voorspeller ($R^2 = .01$, $\beta = .133$, $p = .049$). Leerlingen die later hebben leren fietsen scoren beter op waarnemen. Dit komt overeen met de bevindingen voor de oversteektaak, maar staat haaks op de bevindingen voor gevaarherkenning.

Besluitvaardigheid

Uit de resultaten van de regressie analyse blijkt dat hoe vaak leerlingen een bekende route fietsen een significante voorspeller is voor de score op besluitvaardigheid, 2% van de variantie wordt verklaard door deze voorspeller ($R^2 = .02$, $\beta = -.425$, $p = .025$). Leerlingen die vaker een bekende route fietsen, scoren lager op besluitvaardigheid. Dit komt overeen met de bevindingen voor de dode hoek taak.

Dode hoek taak

De resultaten van de regressie analyse tonen aan dat hoe vaak leerlingen een bekende route fietsen een significante voorspeller is voor de score op de dode hoek taak, 1% van de variantie wordt verklaard door deze voorspeller ($R^2 = .01$, $\beta = -.407$, $p = .048$). Leerlingen die vaker een bekende route fietsen, scoren slechter op de dode hoek taak.

² De betekenis van de p-waarden die hier (en in het vervolg) wordt besproken verschillen van de p-waarden bij de eerdere psychometrie. Hier is een p-waarde de kans dat een gevonden effect ook echt bestaat, op basis van de data die is verzameld. Als een p-waarde dus gelijk of kleiner is dan .05, dan betekent dit dat er gronden zijn om de nulhypothese (bijvoorbeeld: 2 groepen presteren even goed) te verwerpen en dat er steun is voor de alternatieve hypothese (bijvoorbeeld: groep 1 presteert beter dan groep 2). Dit betekent echter niet dat de alternatieve hypothese is bewezen (dat is een misinterpretatie). Er kunnen namelijk nog andere verklaringen zijn.

Oversteektaak

Uit de resultaten van de regressie analyse blijkt dat twee variabelen de score op de oversteektaak (scoringsvorm: keuze veilige gaps) significant voorspellen. 2% van de variantie wordt verklaard door de voorspeller leeftijd ($R^2 = .02$, $\beta = -.369$, $p = .025$) en 2% van de variantie wordt verklaard door de voorspeller hoe vaak leerlingen in drukke verkeerssituaties fietsen ($R^2 = .02$, $\beta = -.507$, $p = .013$). Hoe ouder de leerling was toen hij/zij leerde fietsen des te beter de score op de oversteektaak (in contrast met gevaarherkenning). Leerlingen die vaker in drukke verkeerssituaties fietsen scoren beter op de oversteektaak.

Sociaal en moreel handelen

Significante voorspellers voor de score op sociaal en moreel handelen zijn hoeveel dagen leerlingen gemiddeld per week fietsen ($\beta = 2.251$, $p = .048$), hoe vaak leerlingen in het donker fietsen ($\beta = 4.256$, $p < .001$) en dat leerlingen meestal alleen fietsen ($\beta = -3.974$, $p = .029$). Deze voorspellers te samen verklaren 2% van de variantie ($R^2 = .02$). Leerlingen die meer fietsen gedragen zich minder sociaal in het verkeer, leerlingen die vaker in het donker fietsen gedragen zich minder sociaal in het verkeer, leerlingen die meestal alleen fietsen gedragen zich socialer in het verkeer.

Gevaarherkenning

De resultaten van de regressie analyse tonen aan dat de leeftijd waarop leerlingen leren fietsen een significante voorspeller is voor de score op gevaarherkenning, 1% van de variantie wordt verklaard door de voorspeller leeftijd ($R^2 = .01$, $\beta = -.204$, $p = .041$). Leerlingen die op latere leeftijd hebben leren fietsen scoren lager op gevaarherkenning.

4.3.3 Cito-scores

Deze paragraaf beschrijft het verband tussen de Cito-score van de school en het toetsresultaat. De Cito-score is bekend voor de school, maar niet per klas of per leerling. Het gaat dus om een vergelijking op basis van een algemeen presetatieniveau van de school. Om een vergelijking te kunnen maken zijn scholen verdeeld in laag of hoog op basis van Cito-score. Bij de deelttoets gevaarherkenning is een t-toets uitgevoerd voor onafhankelijke steekproeven (Tabel 4.19 **Error! Reference source not found.**) en voor de overige deelttoetsen de Mann-Whitney U toets (Tabel 4.18). Uit deze analyses bleken geen verschillen in Cito-score op de toetsresultaten voor de deelttoetsen.

Onderdeel	Aantal deelnemers	U	Z	P	r
Waarnemen	280	7968.00	-.64	.52	-.04
Besluitvaardigheid	280	7958.00	-.63	.53	-.04
Dode hoek	324	9855.00	-.62	.54	-.03
Oversteektaak	259	6589.00	-.86	.39	-.05
Sociaal en moreel handelen	282	7575.00	-1.27	.20	-.08

Tabel 4.18. Resultaten van de Mann-Whitney U analyses, met Cito als factor van invloed op de verschillende onderdelen. *Noot.* r = effectgrootte, waarbij $<.3$ een klein effect is, tussen $.3$ en $.5$ een medium effect is, en $>.5$ een groot effect is. r is berekend met de formule $Z/\sqrt{N}$. De formule hiervan komt uit Field (2013).

Afhangelijke Variabele	t	df	p	Cohen's d	95% Betrouwbaarheidsinterval	
					Lager	Hoger
Gevaarherkenning	-1.63	310	.10	-.205	-.866	.080

Tabel 4.19. T-toets van Cito (hoog/laag) op Gevaarherkenning.

4.3.4 Verkeerseducatie

In deze paragraaf wordt de invloed van verkeerseducatie op de toetsresultaten beschreven. Er is informatie ingewonnen over verschillende aspecten van verkeerseducatie (tijdsbesteding per week, verkeerslabel, gevolgde programma's etc.), maar dit bleek niet altijd bruikbaar. Daarom is nu alleen gekeken naar verbanden met het door de klas gevolgde doorlopende verkeersprogramma. De gevonden resultaten zijn slechts indicatief. Ook wat betreft dit aspect was de verzamelde data niet altijd betrouwbaar en compleet, en zijn er meerdere factoren die het gevonden resultaat kunnen beïnvloeden.

Om toch een eerste analyse te kunnen doen zijn de one-way ANOVA en de Kruskal-Wallis test gebruikt. De Kruskal-Wallis toets is een non-parametrische variant van de one-way ANOVA. Deze toets staat dus toe dat de niveau's van een factor onderling worden vergeleken, wanneer de data niet normaal is verdeeld. Tevens zijn non-parametrische toetsen minder gevoelig voor verschillen in het aantal deelnemers (Field, 2013). Echter, mogelijk kan het verschil in deelnemers hier nog steeds een rol spelen doordat er relatief grote verschillen zijn tussen de programma's (één programma heeft 57 deelnemers, terwijl de andere twee programma's respectievelijk 109 en 104 deelnemers hebben).

Voor de deoltoetsen Oversteektaak, Dode Hoek Taak en Waarnemen zijn er geen verbanden gevonden. Significante effecten worden verder beschreven.

Sociaal en moreel handelen

Het verband tussen verkeersprogramma en sociaal en moreel handelen is geanalyseerd met behulp van de Kruskal-Wallis toets. De resultaten van de algemene test (zie Tabel 4.20) geven aan dat de programma's onderling van elkaar verschillen. Dit geeft echter niet aan of alle programma's van elkaar verschillen, of dat het gaat om een verschil tussen slechts twee van de programma's.

Factor	χ^2	DF	P	η^2
Verkeersprogramma's	6.19	2	.05	.02

Tabel 4.20. Kruskal-Wallis toets met Verkeersprogramma op Sociaal en Moreel Handelen. *Noot.* $N = 270$. η^2 is berekend met de formule $\chi^2 / (N - 1)$. De formule om η^2 voor de Kruskal-Wallis te berekenen komt van Green en Salkind (2005).

Gevaarherkenning

Het effect van verkeersprogramma op Gevaarherkenning is geanalyseerd met behulp van een one-way ANOVA. Met deze toets kan worden vergeleken hoe de niveau's van een factor onderling verschillen, mits de data normaal verdeeld is (wat hier het geval is). De resultaten van de algemene test (zie **Error! Reference source not found.**) geven aan dat de programma's onderling van elkaar verschillen. Voor verdere interpretatie van de gevonden resultaten zijn er post-hoc analyses uitgevoerd. Echter, doordat er ook nog andere factoren zijn die kunnen meespelen, naast het gebruikte programma, waren de uitkomsten niet verhelderend. De post-hoc toetsen zijn daarom niet

gerapporteerd (hetzelfde geldt voor de onderdelen *sociaal en moreel handelen* en *besluitvaardigheid*).

Factor	Df	F	P	η^2
Verkeersprogramma	2	4.40	.01	.03

Tabel 4.21. Resultaten van de one-way ANOVA met Verkeerseducatie als onafhankelijke variabele en Gevaarherkenning als afhankelijke variabele. *Noot.* $N = 266$, type III Sum of Squares.

4.3.5 Stedelijkheid

Deze paragraaf beschrijft de invloed van urbanisatiegraad op de toetsresultaten. Daarvoor zijn de plaatsen waar de scholen zijn gevestigd gecategoriseerd als stad of platteland. Voor de deelttoetsen gevaarherkenning, dode hoek taak, sociaal en moreel handelen en waarnemen zijn er geen verbanden gevonden. Significante effecten worden verder beschreven.

Besluitvaardigheid

Voor het testen van een mogelijk effect van stedelijkheid op de scores van Besluitvaardigheid is er een Mann-Whitney U toets gebruikt (zie **Error! Reference source not found.**). Hieruit blijkt dat er een marginaal significant effect is van stedelijkheid op Besluitvaardigheid. Dit wijst op de mogelijkheid dat leerlingen uit steden en van het platteland van elkaar kunnen verschillen op hun scores van Besluitvaardigheid. Tabel 4.24 **Error! Reference source not found.** toont de gemiddelde ranken, wat aangeeft hoe de mogelijke verschillen eruit kunnen zien. Hierin is te zien dat leerlingen van het platteland lager scoren dan leerlingen uit de stad.

Factor	U	Z	P	r
Stedelijkheid	8618.50	-1.76	.08	-.11

Noot. $N = 280$.

Tabel 4.22. Mann-Whitney U toets van Stedelijkheid (Platteland/Stad) op Besluitvaardigheid.

Variabele	Aantal deelnemers	Mean Rank
Besluitvaardigheid	Platteland 135	131.84
	Stad 145	148.56

Tabel 4.23. Gemiddelde ranken van platteland/stad op Besluitvaardigheid.

Oversteektaak

Voor het testen van een mogelijk effect van stedelijkheid op de scores van de Oversteektaak is er een Mann-Whitney U toets gebruikt. De resultaten van deze analyse staan in Tabel 4.24. Hieruit blijkt dat er een marginaal significant effect is van stedelijkheid op de Oversteektaak. Dit wijst op de mogelijkheid dat leerlingen uit steden en van het platteland van elkaar kunnen verschillen. In Tabel 4.25 staan de gemiddelde ranken, die aangeven dat leerlingen uit de stad mogelijk beter scoren dan leerlingen van het platteland.

Factor	U	Z	P	r
Stedelijkheid	7157.00	-1.89	.06	-.12

Noot. $N = 259$.

Tabel 4.24. Mann-Whitney U toets van Stedelijkheid (Platteland/Stad) op de Oversteektaak.

Variabele	Aantal deelnemers	Mean Rank
-----------	-------------------	-----------

Oversteektaak	Platteland	115	120.23
	Stad	144	137.80

Tabel 4.25. Gemiddelde ranken van platteland/stad op de oversteektaak.

5. Discussie en Conclusie

De FietsmeetlatBO is driemaal getoetst (versie 1.0, 2.0 en 3.0) bij basisschoolleerlingen en verbeterd na elke afname. In dit rapport is de afname van de FietsmeetlatBO 3.0 besproken. Het doel van deze afname was om inzicht te krijgen in de bruikbaarheid, de psychometrische eigenschappen, betrouwbaarheid en beïnvloeders van de FietsmeetlatBO. In paragraaf 5.1 en 5.2 beschrijven we de belangrijkste bevindingen. De conclusies volgen in paragraaf 5.3.

5.1 De bruikbaarheid van de FietsmeetlatBO

De onderzoeksvragen die gesteld zijn over de bruikbaarheid van de FietsmeetlatBO zijn: hoe de toets functioneert in de praktijk; hoeveel tijd leerlingen nodig hebben om de toets te maken; en of leerlingen genoeg gemotiveerd blijven gedurende de toets. Deze vragen worden in de volgende paragrafen beantwoord.

5.1.1 Hoe functioneert FietsmeetlatBO 3.0 in de praktijk?

De technische verbeteringen en aangescherpte instructies in FietsmeetlatBO 3.0 zorgden ervoor dat de leerlingen minder assistentie nodig hadden vergeleken met de eerste afname. Doordat uit de eerdere afnames bleek dat er een goede stabiele internetverbinding nodig was, zijn de taken die een internetverbinding nodig hadden omgezet naar lokale taken. Hierdoor werden er met betrekking tot het internet geen problemen meer ondervonden. Er is overigens wel een internetverbinding nodig voor het wegschrijven van data, maar dit is minder belastend en zorgt niet voor problemen.

5.1.2 Hoeveel tijd hebben de leerlingen nodig om de toets te maken?

Leerlingen doen gemiddeld een uur over het doorlopen van de FietsmeetlatBO 3.0. Dit vindt 21% van de leerlingen te lang en 26% van de leerlingen staat daar neutraal tegenover. De rest van de leerlingen is positief over de duur van de toets.

5.1.3 Blijven de leerlingen voldoende gemotiveerd gedurende de toets?

De FietsmeetlatBO wordt positief gewaardeerd door leerlingen. Evenals bij eerdere afnames krijgt de toets een hoog cijfer, vinden kinderen het leuk en snappen ze wat ze moeten doen. De ervaringen met de FietsmeetlatBO 3.0 zijn nog een stuk positiever ten opzichte van FietsmeetlatBO 1.0 en 2.0. In totaal 73% van de leerlingen geeft de FietsmeetlatBO 3.0 een score van 8 of hoger ten opzichte van 63% bij FietsmeetlatBO 1.0.

Wanneer de toets (te) lang is, kan de volgorde van waarin de toetsonderdelen worden aangeboden invloed heeft op het resultaat. Om te meten of dit zo is, zijn vier verschillende volgordes gemaakt. Uit de analyses blijkt een effect van de verschillende volgordes voor de toetsonderdelen gevaarherkenning en besluitvaardigheid. Voor beide onderdelen verschillen twee volgordes significant van elkaar. De richting van het effect varieert, waardoor we op basis van deze bevindingen geen conclusies kunnen trekken over de oorzaak van het verschil. Belangrijk is wel de bevestiging dat de volgorde inderdaad een invloed heeft, en dat er dus goed aan wordt gedaan ook in bij een volgende versie van de FietsmeetlatBO de toetsonderdelen in volgorde te variëren.

5.1.4 Investerings in FietsmeetlatBO als basis voor toekomstige meetlatten

De ervaringen die nu zijn opgedaan met het ontwikkelen van de Fietsmeetlat zijn bruikbaar voor het ontwikkelen van meetlatten van andere doelgroepen. Twee andere doelgroepen die binnen de werkgroep en IPO zijn genoemd en in aanmerking komen voor een meetlat zijn middelbare scholieren en ouders. Voor beide is nu een brede basis gelegd. De overlap met ouders is groot (schatting 2/3), voor middelbare scholieren spelen andere competenties, waarvoor meer nieuwe ontwikkeling binnen WEVER nodig is. Voordat we andere meetlatten gaan ontwikkelen willen we eerst meer duidelijkheid en inzicht krijgen in de implementatie van WEVER (FietsmeetlatBO) (financiering, haalbaarheid, kosten voor producenten etc.).

5.2 Psychometrische eigenschappen, betrouwbaarheid en beïnvloeders

De onderzoeksvragen die gesteld zijn over de psychometrische eigenschappen, betrouwbaarheid en beïnvloeders van de FietsmeetlatBO hebben betrekking op de kwaliteit en betrouwbaarheid, de moeilijk en het onderscheidend vermogen van de opgaven, de (verschillende niveaus van) competenties die worden gemeten en de invloed van achtergrondvariabelen. Deze vragen zijn hierna per onderdeel van de FietsmeetlatBO beantwoord.

5.2.1 Sitatiebewustzijn

Voor sitatiebewustzijn is getoetst met de twee deeltolsten 'waarnemen' en besluitvaardigheid', beide bestaande uit filmpjes die stil zijn gezet met een vraag. Bij 'waarnemen' klikken kinderen op verkeersdeelnemers gerelateerd aan gevaar, bij 'besluitvaardigheid' op verkeersdeelnemers in volgorde van voorrang.

Waarnemen

Psychometrische eigenschappen: Bij het onderdeel waarnemen is de moeilijkheidsgraad van de meeste opgaven erg laag, verschillen de prestaties per boekje (de twee versies met grotendeels overlappende opgaven) en is de betrouwbaarheid laag. Dit toetsonderdeel levert vooralsnog dus geen betrouwbare resultaten op.

Beïnvloedende factoren: Resultaten tonen aan dat leerlingen die later hebben leren fietsen, hoger scoren op het onderdeel waarnemen. Ook leerlingen die de toets makkelijker vonden scoren hoger. Voor de andere onderzochte beïnvloeders zijn op dit onderdeel geen significante verbanden gevonden.

Besluitvaardigheid

Psychometrische eigenschappen: De opgaven van dit toetsonderdeel onderscheiden goed tussen leerlingen met meer en minder besluitvaardigheid, geven een redelijk spreiding qua moeilijkheidsgraad en lijken één construct te meten. Wel kan de betrouwbaarheid van de totaalscore nog beter.

Beïnvloedende factoren: De resultaten op dit toetsonderdeel verschillen afhankelijk van het doorlopende programma verkeerseducatie dat wordt gevolgd in de klas. Dit is een eerste indicatie van een verband, maar moet met grote voorzichtigheid worden geïnterpreteerd: de ingewonnen informatie over verkeerseducatie is beperkt (10 scholen) en niet altijd betrouwbaar. Er kan namelijk niet met zekerheid worden vastgesteld dat de docent die de vragenlijst heeft ingevuld volledig op de hoogte is van welke verkeerseducatie de kinderen gedurende hun hele basisschoolloopbaan hebben gehad.

5.2.2 Sociaal en moreel handelen

Voor sociaal moreel handelen is een deeltoets ingezet waarbij kinderen aangeven hoe vaak ze van de 10 keer een bepaald gedrag vertonen.

Psychometrische eigenschappen

De interne consistentie van het onderdeel Sociaal en Moreel Handelen is acceptabel. Dit betekent dat deze deeltoets betrouwbare resultaten oplevert. Verder lijkt het huidige Sociaal en Moreel Handelen geen éédimensionaal construct te zijn, maar te bestaan uit meerdere factoren. Eén factor lijkt te wijzen op het eigen gedrag en de andere factor op directe interacties met andere weggebruikers. Mogelijk is er nog een derde factor die nu slechts door één opgave wordt belicht.

Beïnvloedende factoren

Uit de resultaten blijkt dat er een positief significant verband is tussen verschillende factoren van fietservaring en sociaal en moreel handelen, zoals: hoe mee fietservaring, hoe minder sociaal en moreel het gerapporteerde gedrag. Dit is een bekend effect dat ook bij andere competenties wordt gevonden, en geeft maar weer aan dat kunnen en willen twee verschillende zaken zijn. Verder gedragen leerlingen die de toets leuk vonden of een hoger cijfer gaven zich vaker socialer in het verkeer. Als laatste verschillen de resultaten op dit toetsonderdeel afhankelijk van het doorlopende programma dat wordt gevolgd. Net als bij andere toetsonderdelen moeten deze bevindingen eerst nader worden onderzocht en voorzichtig worden geïnterpreteerd.

Al met al kan er worden gesteld dat het onderdeel Sociaal en Moreel handelen betrouwbare scores oplevert.

5.2.3 Gevaarherkenning

Voor gevaarherkenning is een deeltoets ingezet, waarin kinderen op de spatiebalk moeten klikken als ze denken dat een situatie gevaarlijk kan worden.

Psychometrische eigenschappen: Uit de resultaten blijkt dat het onderdeel gevaarherkenning een hoge mate van kwaliteit en betrouwbaarheid heeft. De onderdelen hangen met elkaar samen en het onderscheidend vermogen van de toetsopgaven is hoog. Dit betekent dat verschillende niveaus van de competentie 'gevaarherkenning' kunnen worden aangetoond. Deze resultaten tonen aan dat het onderdeel zoals deze in FietsmeetlatBO 1.0 is afgenomen, sterk is verbeterd.

Beïnvloedende factoren: Uit de resultaten blijkt dat leerlingen die op latere leeftijd hebben leren fietsen lager scoren op gevaarherkenning en dat leerlingen die vinden dat ze ervaren fietsers zijn hoger scoren op gevaarherkenning. Wat betreft verkeerseducatie lijkt het toetsresultaat verband te houden met het gevolgde doorlopende programma. Net als bij andere toetsonderdelen moeten deze bevindingen eerst nader worden onderzocht en voorzichtig worden geïnterpreteerd. Leerlingen die aangeven een gevaar gevaarlijker te vinden, scoren hoger op gevaarherkenning. Er zijn geen verbanden gevonden tussen scores op dit toetsonderdeel en motivatie, Cito-scores of stedelijkheid.

5.2.4 Beslissen in complexe situaties

Voor beslissen in complexe situaties zijn twee deeltoetsen ingezet. Voor de eerste is game ontwikkeld, waarmee ook is ingegaan op het verzoek van IPO om te werken met innovatieve toetsmethoden die het werkelijke gedrag zo dicht mogelijk benaderen. De andere deeltoets zijn filmpjes die worden stilgezet, waarna de leerling moet aangeven of het nog veilig is om over te steken of niet.

Dode hoek

Psychometrische eigenschappen: Het dode hoek onderdeel scoort laag op betrouwbaarheid en blijkt drie deelconstructen te meten. Dit zou een lage score op de betrouwbaarheid kunnen verklaren. Hierdoor is ook het onderscheidend vermogen op de verschillende opgaven veelal laag. De toets meet drie verschillende competenties, namelijk rekening houden met vrachtwagens, waarbij de fietser moet stoppen voor een verkeerslicht, het fietsen rondom een vrachtwagen waarbij de fietser niet hoeft te stoppen en het herkennen van dode hoeken.

Beïnvloedende factoren: Leerlingen die vaker een bekende route fietsen, scoren lager op het dode hoek onderdeel. De andere gemeten achtergrondvariabelen hebben geen invloed op de scores.

Oversteektaak

Psychometrische eigenschappen: Uit de analyse van de oversteektaak blijkt dat er een subtiel verschil zit in de beleving van een gap (de ruimte tussen twee auto's bij oversteken) en dat de manier van scoren wellicht op een betere manier kan. Wanneer de items geanalyseerd worden door middel van een betrouwbaarheidsanalyse blijkt dat de betrouwbaarheid laag is. Uit een factoranalyse blijken de gaps te laden op verschillende factoren. Hoe deze gaps laden op de verschillende factoren (of ze samen 1 construct meten) blijkt daarnaast ook afhankelijk te zijn van de lengte van de gap en of deze veilig of onveilig is. Verder hebben sommige items een hoge onderscheidend vermogen en andere een lage onderscheidend vermogen (hoe goed er onderscheid gemaakt kan worden tussen hoog/laag scorende leerlingen). Ook dit lijkt te maken te hebben met de lengte van de gap.

Uit een aanvullende OPLM analyse blijkt dat de betrouwbaarheid van de test zeer hoog is wanneer er gekeken wordt naar wat leerlingen een acceptabele gap vinden in plaats van dit voorafgaand vast te hebben gesteld.

Beïnvloedende factoren: De scores op de oversteektaak worden beïnvloed door hoe leuk leerlingen de toets vinden, de leeftijd waarop de leerling leerde fietsen, en of leerlingen vaak in drukke verkeerssituaties fietsen. De relatie tussen deze factoren en de score op de oversteektaak is echter zeer zwak. Tevens een zwak verband is er met stedelijkheid: leerlingen uit de stad lijken hoger te scoren dan leerlingen van het platteland. Voor de overige gemeten achtergrondkenmerken is er geen verband gevonden met de score op dit toetsonderdeel.

5.2.5 Worden met de toetsonderdelen verschillende competenties gemeten?

De resultaten tonen aan dat er met de zes toetsonderdelen verschillende competenties worden gemeten binnen de vier domeinen. De oversteektaak vertoont samenhang met de toetsonderdelen sociaal en moreel handelen en waarnemen, de andere toetsonderdelen vertonen geen samenhang. De toetsonderdelen die samenhang tonen correleren echter zeer zwak met elkaar, wat betekent dat de samenhang zeer gering is.

5.3 Conclusie

De bruikbaarheid van de toets is sterk verbeterd ten opzichte van de eerste afname door technische verbeteringen en het verbeteren van de opgaven. Leerlingen blijken de toets leuk te vinden en 80% van de leerlingen vindt de toets niet te lang. Leerlingen die de toets leuk vinden scoren ook significant hoger op veel taken. Het verband dat hier gevonden wordt is echter zeer zwak, wat betekent dat motivatie uiteindelijk weinig invloed heeft op de scores van leerlingen. Daarnaast lijkt de volgorde waarop de toetsonderdelen aangeboden worden een rol te spelen bij de onderdelen Gevaarherkenning en Besluitvaardigheid. Wat deze invloed precies is, is echter nog onduidelijk. Voor nu is het vooral belangrijk om te onderkennen dat dit verband er is, en ook in het vervolg verschillende volgordes te gebruiken om dit effect zo klein mogelijk te maken. Al met al kunnen we concluderen dat – hoewel er nog verbeteringen mogelijk zijn – de toets nu naar behoren functioneert.

Uit de resultaten blijkt dat de deoltoetsen verbeterd zijn ten opzichte van de eerste afname. Iedere deoltoets meet iets unieks en er kunnen verschillende niveaus van competentie worden aangetoond. We kunnen op basis van de psychometrische eigenschappen concluderen dat drie deoltoetsen al goed functioneren. De betrouwbaarheid van de deoltoets Gevaarherkenning, de Oversteektaak (in de variant acceptatie van gaplengte) en Sociaal en moreel handelen zijn hoog. De overige drie deoltoetsen zijn in de huidige vorm nog niet goed genoeg. Van de deoltoets Besluitvaardigheid moet de betrouwbaarheid nog iets hoger, en van de deoltoetsen Dode hoek en Waarnemen is de betrouwbaarheid nog veel te laag.

Voor een aantal beïnvloedende factoren is een verband gevonden met de toetsresultaten. Met name het type verkeerssituaties waarin leerlingen fietsen (zoals drukke wegen, onbekende routes) lijken een kleine, maar significante invloed te hebben op de testcores van verschillende onderdelen.

Of de leerlingen in de stad of op het platteland wonen lijkt geen grote rol te spelen. Alleen bij de deoltoetsen Besluitvaardigheid en Oversteektaak zijn er marginaal significant effecten gevonden, waarbij leerlingen uit de stad iets hoger scoren. Dit sluit aan bij de redenering dat leerlingen uit steden (misschien) vaker in aanraking komen met druk of veel verkeer, waardoor ze al meer ervaring hebben en de toets mogelijk beter maken. Dat er weinig effect is gevonden kan erop duiden dat WEVER voldoende gevarieerd is qua opgaven.

Wat betreft verkeerseducatie werd een verband gevonden met toetsscore op de deoltoetsen Gevaarherkenning, Sociaal en Moreel Handelen, en Besluitvaardigheid. De toetsresultaten verschillen op deze deoltoetsen afhankelijk van het gevolgde doorlopende programma in de klas. Echter, de data waarop deze bevindingen zijn gebaseerd was niet betrouwbaar genoeg om conclusies te trekken en er zijn veel andere factoren die het verband kunnen beïnvloeden.

De Cito-score van de school lijkt geen invloed te hebben de toetsscore op de verschillende deoltoetsen. Dit geeft een indicatie dat de toets bij leerlingen van alle niveau's kan worden afgenomen, zonder dat leerlingen de toets beter of slechter maakt op basis van het niveau van de school. Echter, dit is momenteel alleen getest in groepen (hoog vs laag) en met de Cito-score per school. Dit zegt dus nog niets over het algehele prestatieniveau van de klas of de individuele leerlingen.

6. Aanbevelingen voor vervolg

Op basis van de resultaten van de afname van FietsmeetlatBO 3.0, hebben we aanbevelingen opgesteld over het verbeteren van de bruikbaarheid van de toets, de psychometrische eigenschappen en het onderzoeken van de beïnvloedende factoren. Daarnaast geven we ook aanbevelingen voor de verdere ontwikkeling en implementatie van WEVER.

Bruikbaarheid toets

De bruikbaarheid van de toets is sterk verbeterd ten opzichte van eerdere versies, maar er is zeker nog winst te behalen. Het doel is uiteindelijk om de toetsafname geheel geautomatiseerd te laten verlopen. Er zijn verschillende stappen mogelijk om dat te bereiken:

- Het automatische randomiseren van toetsonderdelen. Leerlingen krijgen de deeltolsten in verschillende volgorde aangeboden en worden na afronden van een onderdeel automatisch doorverwezen naar het volgende onderdeel
- De toets kunnen installeren op computers van de school. Het is dan niet langer nodig om laptops te huren voor de afname.
- Instructies bij afname laten geven door de docent. Hiervoor moet een duidelijk protocol en uitleg voor de de docent worden geschreven.
- Het verder automatiseren van data-analyse.

Deze aanpassingen maken de afname van de toets gemakkelijker en minder arbeidsintensief. Voor deze fase van het onderzoek functioneert de toets echter goed genoeg, en treden bij de afnames geen problemen op. Voor vervolgafnames kan daarom voorlopig dezelfde afnameprocedure gebruikt worden. Het verbeteren van de functionaliteit van de toets wordt vooral relevant wanneer de plannen om WEVER groot uit te rollen concreter zijn.

Psychometrische eigenschappen

In deze test van de FietsmeetlatBO 3.0 zijn alle deeltolsten uitgebreid toegepast. Belangrijk voor de volgende afname is dat er voor elke competentie (Situatiebewustzijn, Gevaarherkenning, Sociaal en moreel handelen, Beslissen in complexe situaties) een goed functionerende deeltolst is. We beschrijven daarom de aanbevelingen per competentie.

- Situatiebewustzijn: Voor deze competentie zijn er nu twee deeltolsten, Waarnemen en Besluitvaardigheid. Waarnemen functioneert niet goed en er moeten waarschijnlijk ingrijpende aanpassingen worden gedaan aan de vraagvorm om dit te verbeteren. Alleen als dit onderdeel blijkt te onderscheiden tussen onervaren en ervaren fietsers, zouden we dit onderdeel moeten handhaven. In de opdracht WEVER2017 wordt de WEVER toets daarom ook bij volwassen/ervaren fietsers afgenomen. Mocht uit de vergelijking van de scores van de volwassenen met die van de kinderen blijken dat volwassenen dit WEVER onderdeel niet beter maken dan de kinderen, dan betekent dit dat dit onderdeel niet iets meet dat 'belangrijk is voor de veiligheid. In dat geval bevelen we aan om deze deeltolst in het geheel weg te laten, en zo meer ruimte te maken voor de deeltolst Besluitvaardigheid. Deze deeltolst functioneert al wel goed, maar om een betrouwbare totaalscore te verkrijgen zijn er meer opgaven nodig (12 à 13). Bij het maken van nieuwe opgaven voor Besluitvaardigheid kan als inspiratie uit worden gegaan van de bestaande opgaven die nu al een hoog onderscheidend vermogen bezitten.
- Gevaarherkenning: Deze deeltolst presteert goed en wijzigingen in de vraagvorm of aan de opgaven zijn niet nodig.

- Sociaal en moreel handelen: Deze deoltoets presteert goed en wijzigingen in de vraagvorm of aan de opgaven zijn niet nodig. Aangezien uit de analyse blijkt dat de deoltoets bestaat uit meerdere factoren, is het wel interessant om te onderzoeken wat die factoren zijn en daarmee deze deoltoets nog verder aan te scherpen.
- Beslissen in complexe situaties: Voor deze competentie zijn er nu twee deoltoetsen, Dode hoek en Oversteektaak. Dode hoek functioneert niet goed, mogelijk doordat het drie verschillende factoren meet. Om deze deoltoets te verbeteren moeten voor elke factor meer opgaven gemaakt worden, om op die manier de betrouwbaarheid vast te stellen. Alleen als de dodehoe taak blijkt te onderscheiden tussen onervaren en ervaren fietsers, en vorderingen van leerlingen zichtbaar maakt nadat zij een dode hoek programma hebben gevolgd zouden we dit onderdeel moeten handhaven. Mocht dat niet het geval zijn, bevelen we aan om deze deoltoets in het geheel weg te laten. Daarom wordt in WEVER2017 dit toetsonderdeel afgenomen bij volwassen ervaren fietsers, en bovendien wordt een voor en nastudie met controle groep uitgevoerd onder leerlingen die een hoek programma hebben gevolgd.
De competentie Beslissen in complexe situaties wordt goed gemeten door de deoltoets Oversteektaak, die naar behoren functioneert. Voor het vervolg houden we voor de Oversteektaak bij voorkeur de scoringsmethode aan die de meest betrouwbare resultaten geeft, namelijk de gap die leerlingen acceptabel vinden, zonder dat de onderzoekers dat vooraf zelf vaststellen. Met die aanpak is het niet nodig om de vraagvorm aan te passen of nieuwe opgaven te maken.

Invloed externe factoren

Bij de analyse van FietsmeetlatBO 3.0 zijn verbanden tussen verschillende externe factoren en het toetsresultaat onderzocht. Er is een aantal relaties gevonden. In het algemeen geldt dat om hierover harde uitspraken te kunnen doen, eerst alle deoltoetsen goed moeten functioneren. We beschrijven voor elk onderzocht aspect de aanbevelingen:

- Motivatie: Uit de aangetoonde verbanden concluderen we dat deze vragenlijst ook voor het vervolg onderdeel moet zijn van de afname. De huidige vraagstellingen leveren de benodigde informatie op en de vragenlijst hoeft daarom niet aangepast te worden. Verder is het belangrijk om verschillende volgordes van de deoltoetsen aan te blijven houden, omdat dit verband houdt met de toetsscore.
- Fietservering: Er zijn met verschillende deoltoetsen verbanden gevonden, en het is dus belangrijk om achtergrondinformatie over de fietservering van leerlingen in te blijven winnen. De gevonden verbanden zijn in enkele gevallen verwarrend, tegenstrijdig en niet zoals men zou verwachten. Meer onderzoek met verbeterende deoltoetsen is nodig om alle bevindingen goed te duiden.
- Cito-score school: Nu uit de analyse gebleken is dat de Cito-score van de school geen verband houdt met het toetsresultaat, hoeft dit niet verder onderzocht te worden. Wel is er nu alleen gekeken naar de score van de school. Om specifiekere uitspraken te doen over het effect van het algemene prestatieniveau op het toetsresultaat, is het aan te bevelen om voor zover dat mogelijk is naar verbanden met de klassikale of individuele Cito-scores.

- Verkeerseducatie: Het in kaart brengen van de verkeerseducatie die de leerlingen tijdens hun basisschoolperiode hebben gehad bleek erg lastig. De huidige vragenlijst die wordt ingevuld door de docent levert onbetrouwbare en incomplete informatie op. Daarnaast spelen er naast de bevraagde onderwerpen nog andere factoren een rol, zoals de invloed van de ouders en docent op vaardigheden van de leerling. Om dit alles goed in beeld te krijgen bevelen wij aan om alternatieve onderzoeksmethoden te testen. Er kan bijvoorbeeld gedacht worden aan een interview met de docent. Voor de uiteindelijke effectenevaluaties is het van groot belang dat er een methode wordt gevonden die alle relevante aspecten rondom verkeerseducatie op betrouwbare wijze in beeld brengt.
- Stedelijkheid: De invloed van de urbanisatiegraad op het toetsresultaat was gering, en het is voor de nieuwe versie van de toets daarom niet nodig om de opgaven meer toe te spitsen op situaties in het platteland dan wel de stad. Wel bevelen we aan om de invloed van de urbanisatiegraad te blijven monitoren.

Bovenstaande aanbevelingen geven aanknopingspunten om de FietsmeetlatBO te verbeteren. Daarnaast spelen er in het ontwikkeltraject van WEVER nog andere zaken. We bevelen dan ook aan om naast het aanpassen van de toets, de energie en focus voor het huidige jaar vooral te leggen bij:

- Het verhelderen van de toepasbaarheid van WEVER in de praktijk. Het eindresultaat is voor verschillende betrokkenen nog onduidelijk. Ook speelt de vraag wat het verschil is tussen mogelijkheden van effectmeting tussen kortdurende en langlopende interventies. Om doelgericht vervolgstappen te kunnen zetten is het belangrijk de ontwikkelrichting van WEVER helder te krijgen.
- Een evaluatiekader opstellen voor de verschillende onderzoeksvragen die WEVER gaat beantwoorden. Een model waarin alle beïnvloedende factoren een plek krijgen helpt daarbij.
- Evalueren van educatieprogramma's. De FietsmeetlatBO is grotendeels klaar voor een volgende fase waarin het als instrument voor een effectmeting wordt ingezet. Dit is een belangrijke doel van het WEVER project.
- Het ontwikkelen van een businessplan. Uiteindelijk is het de bedoeling dat de FietsmeetlatBO in de praktijk gebruikt kan worden ook na afloop van het WEVER-project. Daarvoor moet worden gekeken naar mogelijkheden voor financiering en doorontwikkeling in de toekomst.
- Voor verder toekomst, het ontwikkelen van WEVER voor andere doelgroepen, zoals middelbare scholieren en ouderen.

Literatuur

- Aarts, L.T., Wijnhuizen, G.J., & Dijkstra, A. (2015). Monitoring veiligheidskwaliteit weg- en fietsinfrastructuur. *Nationaal verkeerskundecongres 2015*.
- Bianchi, A., & Summala, H. (2004). The "genetics" of driving behavior: parents' driving style predicts their children's driving style. *Accident Analysis & Prevention*, 36, 655-659.
- Buijs, E.C., & Witvoet, J.S. (2016). Ontwerp van een fietstoets.
- Divera Twisk, J.W., SWOV, Jan Vissers, N.H., & Geertje Hegeman, Royal HaskoningDHV, Erik Roelofs, E.B., Cito 2016. WEVER: 'op Weg naar Effectieve VERkeerseducatie' (voortgangsverslag 2015-2016).
- Field, A. (2009). *Discovering statistics using SPSS (3 ed.)*. London, UK: Sage publications Ltd.
- Field, A. (2013). *Discovering statistics using SPSS (4th ed.)*. California, CA: Sage Publications Ltd.
- Green, S. B. & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (4thth edition). New Jersey: Pearson
- Hegeman, G., Vissers, J.R., & Slinger, W. (2016). Road Safety education checklist. Implementation of a checklist to measure the quality of educational interventions. *International conference of Traffic and Transport Psychology*, Accepted for presentation, Brisbane.
- Scott-Parker, B., Watson, B., & King, M. J. (2009). Understanding the psychosocial factors influencing the risky behaviour of young drivers. *Transportation research part F: traffic psychology and behaviour*, 12, 470-482.
- Slinger, W., Koen, F., Vissers, J., & Twisk, D., (2016). *Op weg naar effectieve verkeerseducatie: het WEVER-project*. Paper presented at the Nationaal Verkeersveiligheidscongres NVVC.
- Twisk, D., Bos, N., & Weijermars, W.A.M. (2016). Road injuries, health burden, but not fatalities make 12 to 17 year olds a high risk group in the Netherlands. Submitted manuscript (9 June, 2016) to the European Journal of Public Health.
- Twisk, D.A.M. (2014). Protecting pre-license teens from road risk: Identifying risk- contributing factors and quantifying effects of intervention strategies *Dissertation Thesis*. University of Maastricht [www.swov.nl/rapport/Proefschriften/Divera Twisk.pdf](http://www.swov.nl/rapport/Proefschriften/Divera_Twisk.pdf), Maastricht.
- Vissers, J.R., de Groot-Mesken, J.S., Twisk, D.S., & Roelofs, E.C. (2015). Naar structurele effectmeting van verkeerseducatie. Ontwikkeling van een toets voor veilig fietsen in het basisonderwijs. RHDHV.
- Vissers, J.R., & Hegeman, G. (2015). Het meten van effecten van verkeerseducatie - Tien gouden regels voor effectmeting. .

Bijlagen

Bijlage 1 Informatiebrief scholen

Uitnodiging deelname onderzoek

WEVER: op weg naar effectieve verkeerseducatie

Voor een onderzoek over het verbeteren van de verkeersveiligheid van jonge fietsers, zijn CITO, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid (SWOV) en RoyalHaskoningDHV op zoek naar scholen waar wij een toets kunnen afnemen om de verkeersvaardigheden van jonge fietsers in kaart te brengen. **Aan u de vraag of wij ons onderzoek mogen uitvoeren op uw school bij de leerlingen van groep 8.** Door mee te helpen aan dit onderzoek draagt u bij aan het ontwikkelen van een goede toets waarmee de kwaliteit van verkeerseducatie verbeterd kan worden, en waardoor kinderen uiteindelijk veiliger kunnen deelnemen aan het verkeer. In deze brief en de bijlage leest u wat het onderzoek inhoudt.

Wat wordt er van u verwacht?

Het onderzoek bestaat uit een computertaak en vragenlijst waarmee verkeersvaardigheden van groep 8 leerlingen worden gemeten. Aangezien wij beseffen dat tijd een kostbaar goed is, streven wij er naar dat u als school zo min mogelijk hoeft te doen om aan het onderzoek deel te nemen. Deelnemen aan het onderzoek betekent het volgende:

- We nemen de fietstoets en vragenlijst af in één (of meerdere) van uw groep 8 klassen. Hiervoor gebruiken we graag een ruimte op uw school. Laptops worden georganiseerd door ons.
- De toets duurt 1 uur en wordt deels online afgenomen. Een voorwaarde voor deelname is dan ook toegang tot internet.
- Om ouders in te lichten over het onderzoek willen wij u vragen leerlingen een informatiebrief mee naar huis te laten nemen. Graag informeren we ouders daarnaast via een eventuele nieuwsbrief of uw website. Hier leveren wij een tekst voor aan die u voor dit doeleinde kunt gebruiken.
- Op de dag van de afname zijn twee proefleiders aanwezig. Zij zetten alles klaar voor het onderzoek, en kunnen vragen van kinderen beantwoorden.

- Proefleiders die aanwezig zijn bij het onderzoek zijn in het bezit van een Verklaring Omtrent Gedrag (VOG).

Wanneer vindt het onderzoek plaats?

We voeren ons onderzoek uit tussen 9 januari 2017 en 17 februari 2017 . Mocht u mee willen werken aan het onderzoek dan prikken wij in overleg met u een datum die schikt.

Voordelen

- Deelname aan het onderzoek helpt om de verkeersveiligheid van kinderen te verbeteren, waardoor kinderen veiliger kunnen deelnemen aan het verkeer.
- U kunt aan ouders en andere betrokkenen aangeven dat uw school meehelpt om de kwaliteit van verkeerseducatie te verbeteren.
- Uit de vorige afname is gebleken dat kinderen het maken van de toets leuk vinden.
- Indien u het op prijs stelt dan houden wij u graag op de hoogte van de resultaten van het onderzoek.

Vergoeding

Als blijk van waardering ontvangt u bij deelname aan het onderzoek een VVV-bon ter waarde van 50 euro, vrij te besteden aan iets leuks voor de klas of de leerlingen.

Contact

Wij stellen het op prijs stellen als u wilt bijdragen aan ons onderzoek. U kunt aan ons laten weten dat u mee wilt doen door contact op te nemen met een van onderstaande contactpersonen. Wij nemen dan binnen 1 week telefonisch contact met u op om na te gaan of deelname voor uw school mogelijk is. Mocht u nog verdere vragen hebben of meer informatie willen over dit onderzoek, neemt u dan gerust contact op met Niki Hukker van Royal HaskoningDHV (niki.hukker@rhdhv.com, 0612508956)

Wat is het doel van het onderzoek?

Dit onderzoek maakt deel uit van een groter onderzoeksproject: het WEVER-project. Momenteel wordt door Cito, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid (SWOV) en Royal HaskoningDHV gewerkt aan het maken van een fietstoets voor leerlingen in groep 8 van de basisschool. Deze toets is bedoeld om de verkeersvaardigheden van jonge fietsers in kaart te brengen. Uiteindelijk dient de toets als meetlat om de effectiviteit van verschillende verkeerseducatie programma's te evalueren. Effectieve verkeerseducatie is immers een belangrijke stap in het vergroten van de verkeersveiligheid van jonge fietsers. In deze fase van het onderzoeksproject willen we graag de tweede versie van de fietstoets afnemen om deze te kunnen evalueren.

Wat houdt het onderzoek in?

Leerlingen die mee doen aan het onderzoek maken eerst een computertoets met filmpjes en animaties. De toets bestaat uit verschillende onderdelen, namelijk: 'gevaarherkenning', 'situatiebewustzijn', 'sociaal en moreel handelen' en 'beslissingen in complexe verkeerssituaties'. De toets bevat 40 opgaven en duurt ongeveer een uur. De verschillende onderdelen zijn hieronder kort beschreven:

- Bij het onderdeel 'gevaarherkenning' gaat het erom hoe goed jonge fietsers zijn in het herkennen van (potentieel) gevaarlijke situaties in het verkeer. Leerlingen krijgen filmpjes te zien vanuit het perspectief van een fietser en moeten de spatiebalk indrukken wanneer zij een verkeerssituatie als gevaarlijk beoordelen.
- Voor het onderdeel 'situatiebewustzijn' zijn we benieuwd naar waar jonge fietsers op letten in het verkeer. Zij krijgen filmpjes te zien vanuit het perspectief van een fietser en moeten vragen beantwoorden over waar zij op hebben gelet (bijv. verkeersborden, andere weggebruikers etc.).
- 'Sociaal en moreel handelen' richt zich op wat leerlingen denken dat goed en slecht gedrag is in het verkeer. Zo zullen zij bijvoorbeeld een filmpje zien waarin iemand tegen het verkeer in fietst op een rotonde. Zij moeten dan aangeven of zij vinden dat dit mag of niet.
- Het onderdeel 'beslissingen in complexe situaties' meet in hoeverre jonge fietsers in staat zijn de meest veilige beslissing te maken in complexe verkeerssituaties. Leerlingen krijgen onder andere animaties te zien van verkeerssituaties met vrachtwagens. Gekeken wordt of zij veilig kunnen omgaan met de dode hoek van de vrachtwagen.

Na ieder onderdeel wordt aan leerlingen gevraagd aan te geven hoe goed zij denken het onderdeel gemaakt te hebben. Ook vullen leerlingen een korte vragenlijst in (maximaal 5 minuten) met vragen over hun ervaring met fietsen, bijvoorbeeld hoe vaak zij fietsen.

Nog een fietstoets?

Wellicht bent u als school al bekend met het fietsexamen van Veilig Verkeer Nederland (VVN). Wij kunnen ons dan ook voorstellen dat u misschien denkt, waarom weer een fietstoets? De WEVER-toets verschilt echter zowel inhoudelijk als qua doel van het fietsexamen.

Het doel van het fietsexamen van VVN is om scholen te stimuleren praktisch en theoretisch verkeersonderwijs te geven. Het schriftelijke deel van het examen toetst de verkeerskennis en het verkeersinzicht, en het praktijkdeel toetst of kinderen deze kennis ook kunnen toepassen in het verkeer.

Het doel van de fietstoets in het WEVER-project is om verkeerseducatieprogramma's (zoals de programma's van VVN, maar ook producten van ANWB en andere aanbieders) te kunnen evalueren. Verkeerseducatieprogramma's worden nu vaak niet of gebrekkig geëvalueerd. De bedoeling is dat met de fietstoets vergelijkbare educatieprogramma's voor jonge fietsers op effectiviteit met elkaar kunnen worden vergeleken. Daarbij richt de WEVER fietstoets zich meer op hogere orde verkeersvaardigheden, zoals gevaarherkenning. Van deze vaardigheden is aangetoond dat zij cruciaal zijn voor veilige deelname aan het verkeer.

Bijlage 2 Informatiebrief ouders

Geachte ouder(s)/verzorger(s),

U ontvangt deze brief omdat wij binnenkort een onderzoek uitvoeren op de school van uw kind. Met het onderzoek willen wij meer inzicht krijgen in de verkeersvaardigheden van kinderen in groep 8. Deze inzichten dragen bij aan het verbeteren van verkeerseducatie aan kinderen om ze uiteindelijk veiliger aan het verkeer te laten deelnemen.

De school van uw kind verleent medewerking aan dit onderzoek en wij zouden het heel fijn vinden als uw kind mee mag doen aan dit onderzoek. Het onderzoek wordt uitgevoerd door onderzoekers van CITO, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid (SWOV) en Royal HaskoningDHV.

Wat houdt het onderzoek in?

Het onderzoek bestaat uit een computertoets met filmpjes en animaties en een vragenlijst. De toets richt zich op verkeerssituaties die lastig zijn voor jonge fietsers (bijvoorbeeld het oversteken van een weg). De korte vragenlijst gaat over hoeveel ervaring kinderen hebben met fietsen en over hoe kinderen en hun sociale omgeving (ouders en vrienden) zich gedragen in het verkeer. Het onderzoek duurt ongeveer 1 uur.

Bezwaar deelname onderzoek

Indien u bezwaar heeft tegen deelname van uw kind aan dit onderzoek, vragen wij u dit **na ontvangst van deze brief** te laten weten aan de school van uw kind. Dit kunt u doen door per e-mail aan te geven dat u bezwaar maakt voor deelname van uw kind aan het onderzoek en hierbij de naam en geboortedatum van uw kind te vermelden. Als u niet reageert, gaat u automatisch akkoord met de deelname van uw kind aan het onderzoek.

Vrijwilligheid

Deelname aan dit onderzoek is volledig vrijwillig. Als uw kind niet aan het onderzoek wil meedoen, of als u niet wilt dat uw kind aan het onderzoek deelneemt, of als uw kind gaandeweg besluit dat zij/hij wil stoppen, dan kan dat op elk moment. Mocht uw kind willen stoppen met het onderzoek, of mocht u binnen 24 uur uw toestemming intrekken, dan verwijderen wij de gegevens van uw kind uit onze bestanden.

Vertrouwelijkheid van onderzoeksgegevens

De gegevens van uw kind worden alleen gebruikt tijdens het onderzoek en worden volledig anoniem verwerkt. Het is daardoor niet mogelijk om uit rapportages op te maken wie aan het onderzoek hebben deelgenomen.

Contact

Mocht u nog verdere vragen hebben of meer informatie willen over dit onderzoek, dan kunt u altijd contact opnemen met de één van de onderzoekers van het onderzoek. Als u niet wilt dat uw kind deelneemt aan het onderzoek dan kunt u dit per e-mail doorgeven aan de school van uw kind.

Hopende u hiermee voldoende te hebben geïnformeerd.

Met vriendelijke groet,

Namens het onderzoeksteam,



Geertje Hegeman

Projectleider Royal HaskoningDHV

Divera Twisk (projectleider SWOV)

divera.twisk@swov.nl

Jessica Witvoet (proefleider SWOV)

jessica.witvoet@swov.nl

Geertje Hegeman (projectleider Royal HaskoningDHV) geertje.hegeman@rhdhv.com

Niki Hukker (proefleider Royal HaskoningDHV)

niki.hukker@rhdhv.com

Bijlage 3 Onderzoeksprotocol afname WEVER 3.0

1. Procedure

Volgorde	Handelingen
Stap 1	Kennis maken met docent en afspraken maken
Stap 2	Opzetten testmateriaal in testruimte
Stap 3	Klassikaal proefleiders voorstellen aan leerlingen en korte introductie over het onderzoek
Stap 4	Afname toets+vragenlijst bij groepjes leerlingen
Stap 5	Afsluiting: school en kinderen bedanken voor deelname

2. Handelingen

2.1 Kennis maken docent

Naast onszelf voorstellen aan de leerkracht(en) moeten een aantal zaken besproken worden voor de afname:

- testruimte (waar kunnen de leerlingen rustig werken?)
- internetverbinding
- indeling van de dag
- vormen kinderen zelf groepjes (indien van toepassing) of stelt de leerkracht deze groepjes samen?
- bij wie kunnen we terecht als er tijdens de afname iets voorvalt?

2.2 Opzetten testmateriaal

Op de testdag moet verbinding worden gemaakt met internet. Er moet getest worden of alle onderdelen goed werken via deze internetverbinding. Alle laptops moeten klaar gezet worden met externe muizen, koptelefoons, en moeten aan de oplader zitten. Het inlogportal moet open staan op iedere laptop. Ook moet er een volgordebriefje + pen klaarliggen voor iedere leerling.

2.3 Voorstellen aan leerlingen

Voor afname van het prototype zullen de proefleiders zich voorstellen aan de kinderen. Het is aan proefleiders om te bedenken wat zij kinderen willen vertellen over zichzelf.

We moeten kinderen in ieder geval vertellen dat zij bij ons terecht kunnen voor vragen of andere zaken. Gedurende de testdag zijn de proefleiders een aanspreekpunt. Daarnaast zullen we kort wat vertellen over het onderzoek en de procedure. Het is van belang dat we vertellen dat de toets vrijwillig is en anoniem. Hieronder staat een voorbeeld:

2.4 Afsluiting

Als de toets is afgenomen zullen we de kinderen en de school nog eens bedanken voor het meedoen. Als bedankje hebben we een vvv-bon voor de klas. Afhankelijk van de dag en de tijd kunnen we nog even napraten met kinderen over de toets.

BELANGRIJK:

- VERZAMEL NA DE AFNAME DE VOLGORDE BRIEFJES! Zo kunnen we zien welke leerling welke volgorde heeft gemaakt.
- SCHRIJF DE VERSIE DIE LEERLINGEN HEBBEN GEMAAKT OP HET VOLGORDE BRIEFJE. Op iedere laptop zit een sticker met een 1 of een 2 (de 2 versies dus).

Bijlage 4 Vragenlijst Verkeerseducatie voor docent

Met deze vragenlijst willen wij graag een indicatie krijgen wat de kinderen uit uw klas allemaal voor verkeerseducatie hebben gekregen in hun gehele schooltijd. Het gaat er hierbij dus specifiek om de **leerlingen die meedoen aan deze verkeerstoets** en niet om de verkeerseducatie van de school op dit moment. Bij meerkeuze vragen kunt u het juiste antwoord omcirkelen. Alvast hartelijk dank voor het invullen!

1. Welke verkeerseducatie hebben de leerlingen die vandaag meedoen aan de verkeerstoets gehad? Kunt u ook aangeven in welke groep dit heeft plaatsgevonden?
 - a. Een doorlopend verkeerseducatie programma, namelijk: (groep bij vraag 2 invullen)
 - i. de verkeerseducatielijnen (verkeerskrantjes) van VVN
 - ii. Let's go (of Wegwijs) van Malmberg
 - iii. Wijzer door het verkeer of Klaar...over van Noordhof
 - iv. Afgesproken van De Boer
 - v. Andere namelijk:.....
 - b. Een dode hoekles, in groep.....
 - i. Weet u de producent/uitvoerder?
 - c. Streetwise van ANWB in groep.....
 - d. Verkeersoneelstuk, namelijk:..... in groep.....
 - e. Fietsencheck in groep
 - f. Aandacht (in de klas) aan verkeerscampagnes, namelijk:.....
 - g. in groep....
 - h. Gordelcontrole in groep.....
 - i. Ik stop voor Zebra's in groep....
 - j. Andere, namelijk:..... in groep....
2. Indien u bij vraag 1 a heeft omcirkeld, hoeveel uur werd er gemiddeld **per week** ongeveer besteed aan het programma in de verschillende groepen?

Groep 1	
Groep 2	
Groep 3	
Groep 4	
Groep 5	
Groep 6	
Groep 7	
Groep 8	

3. Hoeveel heeft de school met deze leerlingen in de praktijk geoefend met veilig lopen en fietsen? Geef per groep aan welke activiteiten lopend of fietsend zijn gedaan. Bovenaan de tabel is een voorbeeld gegeven.

<i>Voorbeeld Groep 6</i>	<i>Lopend naar de bibliotheek, op de fiets naar de moestuinen (1 x per 4 weken), op de fiets naar een museum</i>
Groep 1	
Groep 2	
Groep 3	
Groep 4	
Groep 5	
Groep 6	
Groep 7	
Groep 8	

4. Werd er vooraf aan deze activiteiten in de klas nog extra aandacht besteed aan verkeersveilig gedrag onderweg van en naar deze activiteiten?
 - a. Ja,
 - b. Nee

5. Heeft de school een verkeersouder?
 - a. Ja
 - b. Nee

6. Is de schoolomgeving veilig ingericht? zo ja hoe/waarmee? Denk bijvoorbeeld aan gekleurde paaltjes, een (gekleurd) hek voor de ingang, een oversteek met zebra of verkeerslichten, een waarschuwingsbord voor spelende kinderen, klaarovers, enz.
 - a. Ja, namelijk.....
.....
 - b. Nee

7. Heeft u verder nog opmerkingen over verkeerseducatie op uw school?
.....
.....

Dit is het einde van de vragenlijst, hartelijk dank voor uw medewerking.

Bijlage 5 Vragenlijst Evaluatie FietsmeetlatBO

Je hebt nu alle onderdelen van de fietstoets gemaakt. We willen graag weten wat jij van de fietstoets vond.

Wat vond jij van de toets? Klik je antwoorden aan.

	Helemaal mee oneens	Mee oneens	Niet mee eens/ niet mee oneens	Mee eens	Helemaal mee eens
Ik vond de toets makkelijk	☐	☐	☐	☐	☐
Ik vond de toets leuk	☐	☐	☐	☐	☐
Ik snapte wat ik moest doen	☐	☐	☐	☐	☐
De toets was te lang	☐	☐	☐	☐	☐

Welk cijfer zou jij de toets geven?

- ☐ 0
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7
- ☐ 8
- ☐ 9
- ☐ 10

Wil je nog iets kwijt over de toets?

Bijlage 6 Vragenlijst Fietservaring

We willen graag wat weten over jouw ervaring met fietsen.

Heb je een fiets?

- ☐ Ja
☐ Nee

Hoe oud was je (ongeveer) toen je leerde fietsen?

Hoeveel dagen fiets je gemiddeld per week?

- ☐ 5, 6 of 7 dagen per week
☐ 3 of 4 dagen per week
☐ 1 of 2 dagen per week
☐ Minder dan 1 dag per week

Hoe vaak fiets je in één van deze situaties:

	Nooit	Weinig	Soms	Vaak	Altijd
In het donker	☐	☐	☐	☐	☐
Druk verkeer	☐	☐	☐	☐	☐
Rustige wegen	☐	☐	☐	☐	☐
Bekende route	☐	☐	☐	☐	☐
Onbekende, nieuwe route	☐	☐	☐	☐	☐

Fiets je meestal.....?

- ☐ Alleen
☐ Met iemand anders
☐ In een groepje

Als je met iemand anders fietst, met wie is dit dan meestal?

- ☐ Met vrienden
☐ Met je ouders
☐ Met je broer(tje)/zus(je)

Geef aan of je het eens bent met de stellingen hieronder.

	Helemaal oneens	Oneens	Neutraal	Mee eens	Helemaal eens
Ik ben een goede fietser	☐	☐	☐	☐	☐
Ik vind het leuk om te fietsen	☐	☐	☐	☐	☐
Ik voel me veilig op de fiets	☐	☐	☐	☐	☐

Ik heb veel ervaring met fietsen	☐	☐	☐	☐	☐
----------------------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Ben je in het afgelopen jaar:

	Nee	Bijna	Ja, één keer	Ja, meerdere keren
Met de fiets ergens tegenaan gebotst	☐	☐	☐	☐
Met de fiets gebotst tegen iemand anders	☐	☐	☐	☐

Moest je naar de dokter nadat je ergens tegenaan/ tegen iemand aan was gebotst?

- ☐ Ja
☐ Nee

Bijlage 7 Vragenlijst Ervaringen in het verkeer

Geef aan hoe vaak je dit de afgelopen maand hebt gedaan.

	Nooit	Weinig	Soms	Vaak	Heel vaak
Door rood fietsen	☐	☐	☐	☐	☐
Fietsen in het donker zonder licht	☐	☐	☐	☐	☐
Je telefoon gebruiken terwijl je fietst	☐	☐	☐	☐	☐
Zonder gordel in de auto zitten	☐	☐	☐	☐	☐
Ergens fietsen waar dat niet mag	☐	☐	☐	☐	☐
Met opzet je hand niet uitsteken wanneer je afslaat	☐	☐	☐	☐	☐

Geef aan of je het hiermee eens bent:

	Helemaal oneens	Niet mee een/ niet mee oneens	Helemaal eens
Als anderen door rood fietsen, fiets ik meestal ook door rood	☐	☐	☐
Als anderen met z'n drieën naast elkaar willen fietsen, vind ik het flauw om erachter te blijven fietsen	☐	☐	☐

Als passagier maak je van alles mee. Wat heb jij meegemaakt?
Denk aan de afgelopen maand.

1. Heb je wel eens bij iemand in de auto gezeten die te hard reed?

- ☐ Ja
☐ Nee

Bij wie was dat?

- Één van je ouders
Één van je vrienden
Broer/zus
Bij iemand anders

2. Heb je bij iemand in de auto gezeten die teveel alcohol had gedronken?

- ☐ Ja
☐ Nee

Bij wie was dat?

Één van je ouders

Één van je vrienden

Broer/zus

Bij iemand anders

3. Heb je weleens bij iemand in de auto gezeten die zo gevaarlijk reed, dat je je niet meer veilig voelde?

☐ Ja

☐ Nee

Bij wie was dat?

Één van je ouders

Één van je vrienden

Broer/zus

Bij iemand anders

Wie rijdt er wel eens zonder gordel om in de auto?

	Ja	Nee
Je ouders?	☐	☐
Je vrienden?	☐	☐
Je broers of zussen?	☐	☐

Wie rijdt er wel eens met de fiets of met de auto na het drinken van teveel alcohol?

	Ja	Nee
Je ouders?	☐	☐
Je vrienden?	☐	☐
Je broers of zussen?	☐	☐

Wie rijdt er wel eens met de fiets of met de auto door rood?

	Ja	Nee
Je ouders?	☐	☐
Je vrienden?	☐	☐
Je broers of zussen?	☐	☐