

Het toetsen van gevaarherkenning met behulp van bewegende beelden

Dr. W.P. Vlakveld

R-2014-7

Het toetsen van gevaarherkenning met behulp van bewegende beelden

Onderzoek naar een responsmethode die geschikt is voor het theorie-
examen

Documentbeschrijving

Rapportnummer:	R-2014-7
Titel:	Het toetsen van gevaarherkenning met behulp van bewegende beelden
Ondertitel:	Onderzoek naar een responsmethode die geschikt is voor het theorie-examen
Auteur(s):	Dr. W.P. Vlakveld
Projectleider:	Dr. W.P. Vlakveld
Projectnummer SWOV:	C07.03
Trefwoord(en):	Risk, perception, driving (veh), test, driver training, education, driving test, image processing, skill (road user), reaction (human), international, Netherlands, SWOV.
Projectinhoud:	Dit rapport doet verslag van een onderzoek naar het gebruik van bewegende beelden om gevaarherkenning in het verkeer te toetsen. Voor dit onderzoek zijn twee tests ontwikkeld op basis van korte animatiefilms waarin gevaren verborgen zijn. De twee responsmethoden verschillen. De resultaten van dit onderzoek kunnen door het CBR worden gebruikt bij de ontwikkeling van een nieuwe gevaarherkenningstoets voor het theorie-examen.
Aantal pagina's:	56 + 8
Prijs:	€ 12,50
Uitgave:	SWOV, Den Haag, 2014

De informatie in deze publicatie is openbaar.
Overname is echter alleen toegestaan met bronvermelding.

Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV
Postbus 93113
2509 AC Den Haag
Telefoon 070 317 33 33
Telefax 070 320 12 61
E-mail info@swov.nl
Internet www.swov.nl

Samenvatting

Dit rapport doet verslag van een onderzoek naar het gebruik van bewegende beelden om gevaarherkenning in het verkeer te toetsen. Het onderzoek is uitgevoerd vanuit een samenwerkingsverband tussen de SWOV en het CBR. De resultaten van dit onderzoek kunnen door het CBR worden gebruikt bij de ontwikkeling van een nieuwe gevaarherkenningstoets voor het theorie-examen.

Gevaarherkenning is het vermogen om weg- en verkeerssituaties te detecteren en herkennen die zouden kunnen uitgroeien tot acuut gevaarlijke situaties waarin de kans op een ongeval groot is. Beginnende bestuurders zijn daar minder goed in dan ervaren bestuurders. Door training in de rijopleiding kan het leerproces om gevaren in het verkeer te herkennen, versneld worden. Als er in het rijexamen op een adequate wijze op gevaarherkenning wordt getest, zal hier ook tijdens de rijopleiding meer aandacht aan worden besteed.

Gevaarherkenning is beter te testen aan de hand van bewegende beelden dan met foto's. In Nederland wordt gevaarherkenning nog getest met foto's. In het Verenigd Koninkrijk dienen kandidaten een gevaarherkenningstest te maken waarin gebruik wordt gemaakt van films vanuit het perspectief van een bestuurder. Het is de vraag of die methode veel heeft bijgedragen aan de verkeersveiligheid. De validiteit van deze test is nooit overtuigend aangetoond en de methode is fraudegevoelig. Eerder Nederlands onderzoek naar een toets met bewegende beelden die niet de bezwaren had van de Engelse methode, leverde helaas niet de gewenste resultaten op.

De onderzoeksvraag die in dit rapport wordt beantwoord, luidt: "Is er een valide, betrouwbare en fraudebestendige methode te vinden om aan de hand van bewegende beelden gevaarherkenning te toetsen, die geschikt is om onderdeel te worden van het toekomstige theorie-examen?"

In een test waarin vaardigheden getoetst worden – en in dit geval is die vaardigheid 'gevaarherkenning' – gaat het erom kandidaten in zulke omstandigheden te brengen dat uit hun reacties afgeleid kan worden dat ze een bepaalde vaardigheid beheersen. Die omstandigheden zijn de *stimuli* en de reacties zijn de *responsen*. Op basis van de literatuur en de ervaringen uit het eerdere onderzoek in Nederland zijn twee nieuwe testvormen bedacht. De stimuli zijn in beide testvormen gelijk en bestaan uit korte animatiefilms waarin gevaren verborgen zijn. De twee responsmethoden verschillen.

Bij responsmethode 1 krijgen kandidaten telkens korte animatiefilms te zien met in elke film een dominant potentieel gevaar en enkele minder belangrijke nevengevaren. Nadat zij een film gezien hebben, moeten ze bedenken wat volgens hen het grootste potentiële gevaar in die film was. Vervolgens starten ze de film voor de tweede keer en stoppen de film door op de spatiebalk te drukken op het moment dat ze in gedachten hebben. Daarna wijzen ze op het stilstaande beeld het potentiële gevaar aan door er met de muis op te klikken.

Bij responsmethode 2 krijgen kandidaten dezelfde animatiefilms te zien. Zij zien elk filmpje nu echter maar één keer. Wanneer ze tijdens het zien van een film voelen dat het gevaarlijk zou kunnen worden, drukken ze op de spatiebalk. Onder in beeld loopt een tijdbalk. Als de kandidaten op de spatiebalk hebben gedrukt, horen ze een toon en komt er op de tijdbalk een witte markering te staan op het tijdstip waarop is gedrukt. De film blijft echter gewoon doorlopen. Men kan maximaal vier keer per film op de spatiebalk drukken. Direct na afloop van elke film krijgen de kandidaten de stilstaande beelden te zien van de momenten waarop ze hebben gedrukt. Van die momenten dienen ze het beeld uit te kiezen waarop het grootste potentiële gevaar te zien is. Vervolgens moeten ze op dat geselecteerde beeld, dat groot op het scherm komt te staan, met de muis dat potentiële gevaar aanklikken.

De tests zijn uitgevoerd door twee groepen proefpersonen: rijsschoolleerlingen en professionals (rijexaminatoren en rijinstructeurs). Bij beide responsmethoden scoorden professionals significant beter dan rijsschoolleerlingen. Het verschil was echter aanmerkelijk groter en robuuster bij responsmethode 1.

Er zijn twee verschillende scoringsmethoden gebruikt. Bij de ene methode kon men een film goed of fout beoordelen (scoringsmethode 1). Bij de andere methode werd de hoogte van de score mede bepaald door het moment waarop de kandidaat, binnen de tijd dat het potentiële gevaar er was, het potentiële gevaar aanklikte (scoringsmethode 2). Het verschil tussen de rijsschoolleerlingen en professionals is bij responsmethode 1 groter als scoringsmethode 2 gebruikt wordt dan als scoringsmethode 1 gebruikt wordt. De verschillen tussen de scoringsmethoden zijn echter klein.

Als naar de scores op elke film afzonderlijk gekeken wordt, zijn er veel meer films bij de test met responsmethode 1 die een significant verschil tussen beide groepen laten zien dan bij de test met responsmethode 2. Opvallend is wel dat de moeilijkheidsgraad per film sterk fluctueert en dat op een aantal films zeer slecht wordt gescoord. Dit is er waarschijnlijk de oorzaak van dat de interne consistentie (de Cronbach's α) aan de lage kant is, wat erop duidt dat de films niet allemaal hetzelfde kenmerk (gevaarherkenning) meten.

Rijsschoolleerlingen die veelvuldig computergames spelen, maakten de test met responsmethode 2 significant beter dan rijsschoolleerlingen die weinig tot nooit computergames spelen. De mate waarin rijsschoolleerlingen computergames spelen, had geen invloed op de hoogte van de eindscores bij de test met responsmethode 1. Dit is nog een reden om de test met responsmethode 1 te verkiezen boven de test met responsmethode 2.

De lage eindscores bij de test met responsmethode 2 kunnen deels worden verklaard uit het feit dat rijsschoolleerlingen in 11,5% van de gevallen bij het zien van een film in het geheel niet op de spatiebalk drukten om een potentieel gevaar te markeren. Bij de professionals was dit 12,4%. Het zou kunnen dat bij deze test niet alleen het vermogen gevaren te herkennen, maar ook spelstrategie die niets met gevaarherkenning te maken heeft, een rol heeft gespeeld. Wellicht zijn deze mensen bijvoorbeeld minder geneigd om aan het begin van een film al op de spatiebalk te drukken, omdat ze

bang zijn de vier mogelijkheden te snel te verspelen. Aan het einde van de film komen ze er dan achter dat ze nog mogelijkheden over hebben.

De conclusie luidt dat de test met responsmethode 1 duidelijk geschikter is om verder door te ontwikkelen tot een test in het theorie-examen dan de test met responsmethode 2.

Het CBR wordt aanbevolen om onderzoek te verrichten naar de psychometrische kwaliteiten van de test met responsmethode 1. Ook wordt het CBR aanbevolen onderzoek te verrichten naar de trainbaarheid, waarbij het niet louter gaat om het trainen in het doen van de test. Ten slotte wordt het CBR aangeraden meerdere tests van gelijke zwaarte te ontwikkelen voordat de nieuwe test geïmplementeerd wordt in het theorie-examen.

Summary

Testing hazard perception with moving images; Research into a response method that can be used in the knowledge test

This report presents a study into the use of moving images to test hazard perception in traffic. The study was carried out in cooperation between SWOV and The Dutch Driving Test Organisation (CBR). CBR can use the results of the study to develop a new hazard perception test for the knowledge test.

Hazard perception is the ability to detect and recognize road and traffic situations that could develop into situations of acute danger with a large risk of a crash. Novice drivers are not as skilled as experienced drivers in doing so. Training during the driver education can speed up the learning process of hazard perception in traffic. If the driving exam adequately tests hazard perception, this will also receive extra attention during the driver education.

It is easier to test hazard perception with moving images than with photographs. The Netherlands still uses photographs to test hazard perception. In the United Kingdom, candidates are required to take a hazard perception test in which films are used from the driver perspective. It remains to be seen whether this method has made a considerable contribution to road safety. The validity of this test has never been convincingly demonstrated and the method is susceptible to fraud. An earlier Dutch study into a test with moving images that did not have similar drawbacks as the UK method, unfortunately did not deliver the desired results.

The research question that will be answered in this report is: “Can a valid, reliable, and tamper-proof method be found which uses moving images to test hazard perception and is suitable for use in the future knowledge test?”

A test of someone’s skills – in this case the skill ‘hazard perception’ – requires exposing the candidates to conditions in which their reactions can be used to determine whether they control a certain skill. These conditions are the stimuli and the reactions are the responses. Based on literature and experiences from the earlier Dutch study, two new types of tests were developed. Both test types use the same stimuli which consist of brief animated films in which hazards are hidden. The two response methods, however, differ.

In response method 1, the candidates are presented with several brief animated films, each film containing a dominant potential hazard and some less important additional hazards. After having seen a film, they have to decide which was the most important potential hazard in that film. They then start the film a second time and by pressing the spacebar stop the film at the moment they have in mind. Next, they indicate the potential hazard on the screen capture by clicking the mouse.

In response method 2, the candidates are presented with the same animated films. However, they are shown each film only once. When, while watching the film, they feel that the situation may become dangerous, they press the spacebar. At the bottom of the screen a time bar is visible. When the candidates press the spacebar, they hear a sound and the time bar is marked at the time the spacebar was pressed. The film, however, continues. The spacebar can be pressed a maximum of four times per film. Immediately at the end of each film the candidates are presented with the screen captures of the moments at which they pressed the spacebar. They are then asked to select the moment with the most important potential hazard. Next, they have to indicate that potential hazard with a mouse click in the chosen image which is now shown full-screen.

The tests were performed by two groups of subjects: driving school pupils and professionals (driving examiners and driving instructors). The professionals had a significantly higher score than the driving school pupils on both response methods. However, the difference was considerably greater and more robust on response method 1.

Two different scoring methods were used. In one method, a film could be passed or failed (scoring method 1). In the other method, the height of the score was partly determined by the time that had elapsed since the potential hazard could be detected and the candidate clicked the potential hazard (scoring method 2). The difference between driving school pupils and professionals is higher on response method 1 when scoring method 2 is used rather than scoring method 1. However, the difference between the scoring methods is small.

When the scores are considered for each individual film, many more films show a significant difference between the two groups on the test with response method 1 than on the test with response method 2. It is remarkable, however, that the difficulty level fluctuates considerably per film and that the score is very low on some of the films. This is probably the reason that the internal consistency (the Cronbach's α) is somewhat weak, indicating that the films do not all measure the same characteristic (hazard perception).

Driving school pupils who often play computer games, do significantly better on the test with response method 2 than driving school pupils who rarely or never play computer games. The extent to which driving school pupils play computer games had no effect on the height of the final scores on the test with response method 1. This is one more reason to choose the test with response method 1 rather than the test with response method 2.

The low final scores on the test with response method 2 can partly be explained by the fact that in 11.5% of the cases driving school pupils did not press the spacebar at all to mark a potential hazard while seeing the film. The percentage of the professionals was 12.4%. This could be explained by the fact that in this test it was not only the hazard perception skill playing a role, but also game strategy which has nothing to do hazard perception. These people may, for example, be less inclined to press the spacebar early in the film, because they are afraid to use the four possibilities too quickly. At the end of the film they then find themselves having possibilities left.

The conclusion is that the test with response method 1 is clearly more suitable to continue developing into a test in the knowledge test than the test with response method 2.

CBR is advised to carry out research into the psychometric qualities of the test with response method 1. CBR is also advised to perform research into the trainability, this not only involving training to take the test. Finally, CBR is advised to develop several tests of equal difficulty before implementing the test in the knowledge test.

Inhoud

1. Inleiding	11
1.1. Ontstaansgeschiedenis van het rapport	11
1.2. Achtergrond	11
1.3. Onderzoeksvraag	12
1.4. Aanpak en taakverdeling	12
1.5. Leeswijzer	13
2. Gevaarherkenningstests voor het rijexamen	14
2.1. Beginnende automobilisten, de rijtaak en gevaarherkenning	14
2.2. Definitie van gevaarherkenning	15
2.3. Het meten van gevaarherkenning	17
2.4. Gevaarherkenningstests in het rijexamen	19
2.5. Het testen van gevaarherkenning in Nederland	23
2.6. Eerder Nederlands onderzoek met bewegende beelden	25
3. Methode van onderzoek	28
3.1. De animatiefilms voor de twee responsmethoden	28
3.2. De twee geëvalueerde responsmethoden	29
3.2.1. Responsmethode 1	29
3.2.2. Responsmethode 2	30
3.3. Deelnemers aan het onderzoek	32
3.4. Materiaal en apparatuur	33
3.5. Procedure	33
3.6. Van respons naar score	34
3.6.1. Scoringsmethode 1	35
3.6.2. Scoringsmethode 2	35
3.6.3. Het automatisch bepalen van de score	36
3.7. Design en analyse van de data	36
4. Resultaten	38
4.1. Eindscores responsmethoden 1 en 2 bij scoringsmethode 1	38
4.2. Eindscores responsmethoden 1 en 2 bij scoringsmethode 2	39
4.3. Conclusie eindscores en responsmethoden	40
4.4. De scores per film bij gebruik van scoringsmethode 1	41
4.5. De scores per film bij gebruik van scoringsmethode 2	43
4.6. Conclusies van de scores per film	45
4.7. De vragenlijsten	45
4.8. Het markeren van gevaren bij responsmethode 2	47
4.9. De interne consistentie van de tests	48
5. Conclusies en aanbevelingen	49
5.1. Conclusies	49
5.2. Aanbevelingen	51
Literatuur	52
Bijlage Dominante potentiële gevaren per film	57

1. Inleiding

1.1. Ontstaansgeschiedenis van het rapport

Dit rapport doet verslag van een onderzoek naar het gebruik van bewegende beelden om gevaarherkenning in het verkeer te toetsen. Het onderzoek is uitgevoerd in het kader van een intentieverklaring tot samenwerking tussen de SWOV en het CBR. Het CBR heeft onder andere tot taak toekomstige gemotoriseerde weggebruikers te toetsen op hun rijvaardigheid en de kennis die voor veilige verkeersdeelname noodzakelijk is. De SWOV heeft tot taak om aan de hand van resultaten uit wetenschappelijk onderzoek een bijdrage te leveren aan de verkeersveiligheid. Dankzij wetenschappelijk onderzoek is het mogelijk het rijexamen te verbeteren, zodat het een grotere bijdrage aan de verkeersveiligheid levert dan momenteel het geval is.

In 2005 heeft het CBR aan wat toen nog het Ministerie van Verkeer en Waterstaat heette, voorgesteld het rijexamen drastisch te moderniseren. Deze plannen staan verwoord in het rapport *Van goed naar beter* (CBR, 2005). In de kern pleit dat rapport voor een getrapte vorm van examinering met deeltolsten die aansluiten op de Rijopleiding in Stappen (RIS). In *Van goed naar beter* wordt ook voorgesteld het theorie-examen uit te breiden met een gevaarherkenningstoets. Toenmalig minister Peijs wees het idee van meerdere toetsmomenten af, maar wilde wel dat een gevaarherkenningstoets in het theorie-examen zou worden opgenomen.

Het verzoek van de toenmalige minister heeft ertoe geleid dat de SWOV en het CBR een gevaarherkenningstoets hebben ontwikkeld op basis van foto's (Vlakveld, 2008). Deze toets is in 2009 in het theorie-examen voor het rijbewijs B opgenomen. Door de software die in gebruik was bij het CBR en de klassikale manier waarop het theorie-examen wordt afgenomen, was het niet mogelijk om op dat moment een gevaarherkenningstoets met bewegende beelden in te voeren. Een gevaarherkenningstoets met bewegende beelden heeft echter veel voordelen boven een fototoets (zie de *Paragrafen 2.4, 2.5 en 2.6*). Nu het de bedoeling is om in de nabije toekomst theoriekandidaten individueel te toetsen op pc's die op de examencentra zijn opgesteld, wordt het wel mogelijk om kandidaten te toetsen met bewegende beelden. Het onderzoek waar dit rapport verslag van doet, is de eerste stap op weg naar een dergelijke toets.

1.2. Achtergrond

Gevaarherkenning is het vermogen om weg- en verkeerssituaties te detecteren en herkennen die zouden kunnen uitgroeien tot acuut gevaarlijke situaties waarin de kans op een ongeval groot is. Jonge beginnende bestuurders zijn daar minder goed in dan ervaren bestuurders. Uit een aantal onderzoeken is gebleken dat het vermogen om potentiële gevaren te herkennen, door training is aan te leren (zie voor een overzicht van deze onderzoeken Vlakveld, 2011). Ook is gebleken dat gevaarherkenning accuraat kan worden gemeten in een rijnsimulator als daarbij een apparaat gebruikt wordt dat oogbewegingen registreert (een zogenoemde eyetracker; zie onder meer Crundall et al., 2012; Vlakveld et al., 2011).

Methoden waarbij een rij simulator en een eyetracker noodzakelijk zijn, zijn echter (nog) te duur en te omslachtig in gebruik voor toepassing op het rijexamen. Methoden met foto's of films kunnen gevaarherkenning op een eenvoudiger manier testen. Aan elk van die methoden schort echter wel het een en ander. Het is vooral niet duidelijk of deze eenvoudiger vormen wel echt iemands gevaarherkenningscapaciteiten meten. Dit geldt in het bijzonder voor methoden waarbij foto's gebruikt worden. Voor de methoden met films is gebleken dat de gebruikte responsmethode niet geschikt is voor toepassing op het theorie-examen (Vlakveld, 2011). Daarom wordt in Nederland vooralsnog gebruikgemaakt van de toets op basis van foto's.

Als er in het rijexamen op een adequate wijze op gevaarherkenning wordt getest, zal hier ook tijdens de rijopleiding meer aandacht aan worden besteed. Dit kan het leerproces om gevaren in het verkeer te herkennen versnellen. Het is dus belangrijk dat er een goede methode wordt ontwikkeld om gevaarherkenning te toetsen. Omdat het in de nabije toekomst mogelijk wordt om hierbij bewegende beelden te gebruiken, ligt het voor de hand om het onderzoek naar zo'n methode daarop toe te spitsen.

1.3. **Onderzoeksvraag**

De onderzoeksvraag die in dit rapport wordt beantwoord, luidt: "Is er een valide, betrouwbare en fraudebestendige methode te vinden om aan de hand van bewegende beelden gevaarherkenning te toetsen, die geschikt is om onderdeel te worden van het toekomstige theorie-examen?"

Als eerste indicatie voor de validiteit dient uit onderzoek te blijken dat professionals (rijinstructeurs en rijexaminatoren) de nieuw ontwikkelde toetsvormen in ieder geval significant beter maken dan rij schoolleerlingen, met een grote effectgrootte. De nieuw ontwikkelde toets hoeft in het kader van dit eerste onderzoek nog niet aan alle psychometrische criteria te voldoen, maar moet wel uitzicht bieden op een hoge interne consistentie, waardoor het verantwoord wordt om individuele leerlingen met een lage score voor gevaarherkenning voor dit onderdeel te laten zakken.

1.4. **Aanpak en taakverdeling**

Voor dit onderzoek heeft de SWOV twee toetsmethoden uitgewerkt. De animatiefilms die hierbij zijn gemaakt, zijn gebruikt om te testen of proefpersonen een potentieel gevaar herkennen. Deze animatiefilms zijn voor het grootste deel ontwikkeld op basis van animatiefilms die al voor eerder onderzoek waren vervaardigd (Vlakveld, 2011). Het CBR en de SWOV hebben samen ideeën voor verbetering van het bestaand materiaal ontwikkeld en een scenario opgesteld voor een geheel nieuwe animatiefilm. Het CBR heeft het werk voor de verbetering van het bestaande materiaal en de ontwikkeling van de nieuwe animatiefilm bekostigd. Ook de speciaal ontwikkelde software waarmee het mogelijk wordt om de twee testmethoden op een pc af te nemen, heeft het CBR bekostigd. Daarnaast heeft het CBR gezorgd voor de proefpersonen en hun vervolgens de tests afgenomen. De SWOV heeft de twee manieren van toetsen met twee verschillende responsmethodes ontwikkeld. Daarnaast heeft de SWOV de software ontwikkeld waarmee de scores op de tests berekend kunnen worden, de data geanalyseerd en dit rapport opgesteld.

1.5. Leeswijzer

Hoofdstuk 2 gaat in op gevaarherkenningstests voor het rijexamen. Het hoofdstuk behandelt definities van gevaarherkenning, het belang van gevaarherkenning voor de verkeersveiligheid en eerder onderzoek naar het onderwerp. In *Hoofdstuk 3* wordt besproken welke twee testmethoden voor dit onderzoek ontwikkeld zijn, hoe het onderzoek is opgezet en wie de proefpersonen waren. Ook wordt in dat hoofdstuk uiteengezet hoe de tests zijn afgenomen en hoe de gegevens zijn geanalyseerd. De resultaten van het onderzoek staan in *Hoofdstuk 4*. *Hoofdstuk 5* beschrijft wat de resultaten betekenen. In dit laatste hoofdstuk met conclusies staan ook aanbevelingen voor verder onderzoek. Verder onderzoek is noodzakelijk om tot een nieuwe, wetenschappelijk verantwoorde gevaarherkenningstoets te komen die in het theorie-examen kan worden opgenomen.

2. Gevaarherkenningstests voor het rijexamen

Dit hoofdstuk gaat verder in op de noodzaak van een betere gevaarherkenningstoets voor het rijexamen. Jonge beginnende automobilisten hebben een hoog ongevalsrisico. Voor een deel komt dit doordat beginnende automobilisten niet goed in staat zijn om potentieel gevaarlijke situaties in het verkeer te detecteren en herkennen. *Paragraaf 2.1* behandelt de rol die gevaarherkenning speelt bij de het hoge ongevalsrisico van jongeren. *Paragraaf 2.2* beschrijft de definitie van gevaarherkenning. Welke methoden er zijn ontwikkeld om gevaarherkenning te meten en wat de voor- en nadelen van die methoden zijn, staat beschreven in *Paragraaf 2.3*. Niet elke methode waarmee op wetenschappelijke wijze gevaarherkenning gemeten kan worden, is geschikt voor gebruik tijdens het rijexamen. Voor tests op het rijexamen geldt onder meer dat ze praktisch in gebruik moeten zijn, dat de uitkomsten direct beschikbaar moeten zijn en dat de tests niet fraudegevoelig mogen zijn. Welke ontwikkelingen er wereldwijd zijn met betrekking tot het testen van gevaarherkenning tijdens het rijexamen, staat beschreven in *Paragraaf 2.4*. In de *Paragrafen 2.5 en 2.6* worden de ontwikkelingen in Nederland besproken. *Paragraaf 2.5* gaat over de fototest die in 2009 onderdeel is geworden van het theorie-examen en in *Paragraaf 2.6* wordt eerder Nederlands onderzoek besproken naar een gevaarherkenningstoets met bewegende beelden.

2.1. Beginnende automobilisten, de rijtaak en gevaarherkenning

Jonge beginnende automobilisten (18-24 jaar) hebben in Nederland een relatief grote kans om betrokken te raken bij een ernstig verkeersongeval (met ten minste één dode of één ernstig verkeersgewonde – dat wil zeggen, een gewonde met een letselnst volgens MAIS van 2 of meer¹). Per gereden kilometer is die kans ruim vijf keer zo groot als voor automobilisten van 30 tot 59 jaar. Voor jonge mannelijke automobilisten is die kans zelfs ruim zeven keer zo groot (SWOV, 2012). Dit relatief hoge ongevalsrisico is overigens niet uniek voor Nederland, maar komt overal ter wereld voor waar jongeren na het behalen van hun rijbewijs als zelfstandig autobestuurder aan het verkeer mogen deelnemen (OECD, 2006).

Dat jonge beginnende automobilisten zo'n hoog ongevalsrisico hebben, heeft veel verschillende oorzaken (zie voor een recent overzicht van die oorzaken Vlakveld, 2011). De twee hoofdoorzaken zijn leeftijd en gebrek aan rijervaring. Met leeftijd worden de 'wilde haren' van de adolescentiefase bedoeld. Jongeren zijn, mede doordat hun hersenen nog niet volgroeid zijn, vaak wat meer dan gemiddeld tegendraads en reageren wat impulsiever. Ze nemen meer dan gemiddeld risico om voor hen belangrijke doelen te bereiken en ze laten zich gemakkelijk beïnvloeden door vrienden (o.a. Crone & Dahl, 2012). Door de andere hoofdoorzaak, onvoldoende rijervaring, hebben beginnende automobilisten vooral een gebrek aan zogenoemde hogereordevaardigheden. Hogereordevaardigheden zijn onder meer het herkennen van potentieel gevaarlijke verkeerssituaties, het inschatten van risico's, je eigen grenzen kennen en anticiperend rijden. Uit buitenlands

¹ De MAIS is een internationaal gebruikte maat om de ernst van letsel aan te duiden en is afgeleid uit de letsels die bij de patiënt gecodeerd zijn.

diepteonderzoek naar de achterliggende oorzaken van ongevallen waar jonge automobilisten bij betrokken zijn, is gebleken dat gebrek aan gevaarherkenning de belangrijkste achterliggende oorzaak is (Curry et al., 2011; McKnight & McKnight, 2003).

Automobilisten kunnen voor een groot deel zelf bepalen hoe zwaar en/of gevaarlijk de rijtaak is. Wanneer je bijvoorbeeld harder gaat rijden, wordt de rijtaak zwaarder en neemt het risico op een ongeval toe. Om controle te houden moeten automobilisten voortdurend inschatten wat de risico's zijn en niet meer risico nemen dan ze aankunnen. Het in balans brengen van de risico's die je bereid bent te nemen met je eigen vermogens, wordt kalibratie genoemd (De Craen, 2010). Bij ervaren bestuurders verloopt die kalibratie meestal automatisch. Door de mentale representaties (schemata) die zij op basis van rijervaring hebben ontwikkeld, herkennen ze de situatie en richten ze hun aandacht, vaak zonder dat ze zich daarvan bewust zijn, op objecten en situaties in de verkeersomgeving die potentieel gevaarlijk zijn. Om nog tijdig een ongeval te kunnen vermijden – indien die mogelijke gevaren werkelijkheid worden – houden ervaren bestuurders niet alleen bepaalde elementen in de verkeersomgeving in het oog, maar passen ze meestal ook hun snelheid aan door bijvoorbeeld het gas los te laten. Dit wordt anticiperend rijden genoemd. Een bestuurder rijdt anticiperend en is goed gekalibreerd als hij de gevaren weet te herkennen, de risico's van die gevaren juist weet in te schatten, zichzelf niet overschat en alleen maar bereid is weinig risico te nemen.

Nog weer anders gesteld: op basis van hoe bestuurders verwachten dat de verkeerssituatie zich zou kunnen ontwikkelen, verrichten ze meestal automatisch handelingen die van invloed zijn op hoe de weg- en verkeerssituatie zich daadwerkelijk ontwikkelt. Die handelingen kunnen bijvoorbeeld zijn: in richtingen kijken van waaruit een gevaar zou kunnen opdoemen en de snelheid matigen om daarmee de veiligheidsmarge groot genoeg te houden om een ongeval te vermijden, mocht het verwachte gevaar werkelijkheid worden. Jonge beginnende automobilisten zijn niet goed in het proactief controleren van de weg- en verkeerssituatie. Dat is een van de oorzaken waardoor zij oververtegenwoordigd zijn bij ongevallen in het verkeer.

2.2. Definitie van gevaarherkenning

In dit hoofdstuk zijn al impliciet definities van gevaarherkenning gegeven. In de inleiding van dit hoofdstuk gaat het over het vermogen om 'potentieel gevaarlijke situaties in het verkeer te detecteren en herkennen'. In *Paragraaf 2.1* wordt gesproken over 'het proactief controleren van de weg- en verkeerssituatie'. Een veelgebruikte definitie van gevaarherkenning is die van Horswill & McKenna (2004): 'het "lezen" van de weg, inclusief de zich ontvouwende verkeerssituatie, en het anticiperen op aanstaande gebeurtenissen'. Deze definitie gaat verder dan het herkennen van gevaren. Ze omvat ook het nemen van acties om botsingen tijdig te kunnen afwenden, mochten de herkende potentiële gevaren werkelijkheid worden. Voorbeelden hiervan zijn bepaalde plekken extra in het oog houden en het gas loslaten. Je kunt potentiële gevaren immers wel herkennen, maar dat heeft pas nut als je er ook op reageert. Het gaat bij Horswill & McKenna in feite om een combinatie van gevaarherkenning en anticiperend rijden. Beter zou dan ook zijn om in plaats van gevaarherkenning over gevaaranticipatie

te spreken. Vlakveld (2011) stelt dat gevaaranticipatie de volgende processen omvat:

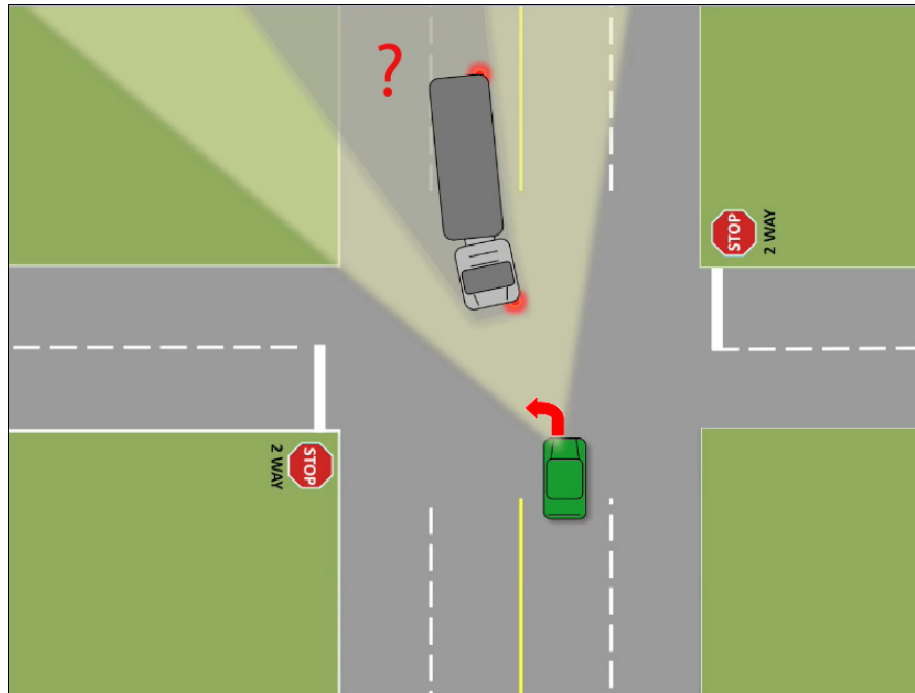
1. detecteren en herkennen van potentieel gevaarlijke weg- en verkeerssituaties;
2. voorspellen hoe deze potentiële gevaren zich zouden kunnen ontwikkelen tot directe bedreigingen;
3. voelen van de risico's die worden opgeroepen door die voorspellingen;
4. selecteren en uitvoeren van handelingen om de gevoelens van gevaar te verminderen en ervoor zorgen dat de veiligheidsmarge groot genoeg wordt om een botsing te voorkomen, mocht het potentiële gevaar zich tot een acuut gevaar ontwikkelen.

Gevaaranticipatie heeft een cognitief aspect. Dit cognitieve aspect bestaat uit de twee eerstgenoemde processen. Gevaaranticipatie heeft ook een emotioneel en motivationeel aspect. Het emotionele aspect zit in het derde proces en voor het vierde proces geldt dat bestuurders ook de wil (motivatie) moeten hebben om acties te nemen. Die bereidheid hebben bestuurders bijvoorbeeld niet wanneer ze een potentieel gevaar herkennen en toch geen anticiperende maatregelen nemen, omdat ze bijvoorbeeld vinden dat andere weggebruikers zelf maar moeten opletten. De vier processen zijn niet sequentieel, maar lopen door elkaar heen.

Het is belangrijk om nu al te vermelden dat het vierde proces, waarbij motivatie en de hoeveelheid risico die iemand bereid is te accepteren een rol spelen, niet goed gemeten kan worden op het rijexamen. Kandidaten kunnen omwille van het behalen van hun rijbewijs tijdens het rijexamen voorzichtig rijden en er daarmee blij van geven weinig risico's te accepteren. Na het rijexamen kunnen deze kandidaten er echter een 'sportieve rijstijl' op nahouden, waarbij ze veel risico's nemen. Omdat motivatie niet goed te meten is op het rijexamen, is het beter om geen tests te ontwikkelen die pretenderen ook het vierde proces objectief te meten. Dit wil overigens niet zeggen dat het vierde proces onbelangrijk is. Het is alleen niet goed te meten op een examen. Daarom is het in de context van het rijexamen toch beter om te blijven spreken over een gevaarherkenningstest (die de eerste drie processen meet) dan over een gevaaranticipatietest (die alle vier de processen meet).

Er bestaan verschillende soorten potentiële gevaren. Twee belangrijke typen zijn *zichtbare potentiële gevaren* en *verborgen potentiële gevaren*. Zichtbare potentiële gevaren zijn zichtbare andere verkeersdeelnemers die zich (nog) niet onveilig gedragen, maar die, gelet op de omstandigheden, zich wel eens zo zouden kunnen gaan gedragen dat ze op botskoers komen. Een automobilist ziet bijvoorbeeld een voetganger op het trottoir en aan de overkant van de straat stopt een bus bij een bushalte. De automobilist moet dan bedenken dat die voetganger wel eens plotseling de straat kan oversteken om nog net op tijd de bus te halen. De voetganger, die de automobilist dus kan zien, is op het moment dat hij nog niet rent (maar zou kunnen gaan rennen) een zichtbaar potentieel gevaar. Verborgen potentiële gevaren zijn mogelijke andere verkeersdeelnemers op botskoers die niet te zien zijn, omdat het zicht erop ontnomen wordt. Wanneer een bestuurder in een kinderrijke omgeving door een straat rijdt met aan weerszijde geparkeerde auto's, moet deze bestuurder zich realiseren dat een kind dat niet te zien is, wel eens tussen de geparkeerde auto's door de straat zou

kunnen oversteken. Dit is een voorbeeld van een verborgen potentieel gevaar. Afdeksituaties zijn ook verborgen potentiële gevaren. *Afbeelding 2.1* geeft een afdeksituatie weer.



Afbeelding 2.1. Voorbeeld van een afdeksituatie.

De automobilist wil in *Afbeelding 2.1* linksaf slaan. De vrachtauto, die ook linksaf wil slaan, staat stil. De automobilist kan vanuit zijn positie niet zien of er op de rijstrook rechts van de vrachtauto, verkeer nadert (de plek met het vraagteken), omdat de vrachtauto hem daarop het zicht ontnemt.

2.3. Het meten van gevaarherkenning

Er wordt al bijna veertig jaar wetenschappelijk onderzoek gedaan naar gevaarherkenning. In de begintijd werd gevaarherkenning hoofdzakelijk 'gemeten' door bestuurders te observeren (met name door te registreren waar bestuurders naar keken) en door bestuurders hardop te laten denken tijdens ritten in het verkeer (Mourant & Rockwell, 1972; Soliday, 1974). Deze methoden waren complex en hadden een lage betrouwbaarheid. Doordat ze in het echte verkeer reden, werden niet alle proefpersonen met dezelfde potentiële gevaren geconfronteerd, waardoor het moeilijk was om prestaties met elkaar te vergelijken.

De eerste onderzoekers die gevaarherkenning niet tijdens het rijden, maar in een laboratorium maten, waren Pelz & Krupat (1974). Zij vroegen proefpersonen om een hendel te bewegen terwijl ze naar een film keken die was opgenomen vanuit het perspectief van de bestuurder. Met de hendel konden de kandidaten de mate van gevaar aangeven. Proefpersonen zonder ongevalsverleden en proefpersonen die nog nooit bekeurd waren, gaven gemiddeld 500 milliseconden eerder aan dat ze een situatie gevaarlijk begonnen te vinden dan proefpersonen die wel een ongevals- en overtredingsverleden hadden. Dezelfde methode is gebruikt door Quimby &

Watts (1981). McKenna & Crick (1997) hebben de methode vereenvoudigd door de hendel te vervangen door een drukknop.

Bij de methode van McKenna & Crick kijken proefpersonen naar korte films die zijn opgenomen vanuit het perspectief van een bestuurder. Elk filmpje bevat minstens één gevaar (bijvoorbeeld een voetganger die tussen geparkeerde auto's door de weg oversteekt). De gevaren ontwikkelen zich in de filmpjes en de dreiging wordt steeds groter. Zodra de proefpersoon een gevaar meent te herkennen, moet hij op een knop drukken. Gemeten wordt hoelang het duurt voordat hij drukt nadat de eerste tekenen van het gevaar zichtbaar zijn geworden. Bij de methode van McKenna & Crick gaat het er dus om hoe vroeg iemand op een knop drukt bij een gevaar dat steeds dreigender wordt. Wat gemeten wordt zijn de reactietijden.

Uit een aantal onderzoeken is gebleken dat personen met een ongevalsverleden de test van McKenna & Crick slechter maken dan personen die nooit een ongeval hebben gehad (Darby, Murray & Raeside, 2009; McKenna & Horswill, 1999; Wells et al., 2008). Er bestaan echter ook onderzoeken waarin geen verband is gevonden tussen ongevalsverleden en scores op een gevaarherkenningstest volgens de methode van McKenna & Crick (Grayson et al., 2003; Grayson & Sexton, 2002; Sagberg & Bjørnskau, 2006). Dat de meeste keren wel een verband is gevonden en enkele keren niet, zou te maken kunnen hebben met de in de films aanwezige gevaren. Zo vonden Sagberg & Bjørnskau (2006) geen overall verschil, maar bleek dat de twee meest complexe gevaren in hun test wel eerder herkend werden door automobilisten zonder ongevalsverleden.

Uit vele onderzoeken met de methode van McKenna & Crick is ook gebleken dat beginners de test significant slechter maken dan ervaren automobilisten (bijvoorbeeld McKenna, Horswill & Alexander, 2006; Sexton, 2000; Wallis & Horswill, 2007). Ook hier geldt echter weer dat er enkele onderzoeken zijn uitgevoerd waarin dit verband niet is gebleken (Chapman & Underwood, 1998; Crundall, Underwood & Chapman, 2002; Sagberg & Bjørnskau, 2006). Dit zou eveneens aan de complexiteit van de in de films aanwezige gevaren kunnen liggen.

De methode van McKenna & Crick is veruit de meest gebruikte methode bij onderzoek naar gevaarherkenning, ondanks het feit dat niet eenduidig is vastgesteld dat deze methode accuraat is. Een van de bezwaren tegen de methode is dat je niet weet waarom proefpersonen op de knop drukken. Het kan het zich ontwikkelend gevaar zijn, maar het kan ook iets anders zijn. Om dit probleem te ondervangen hebben recent onderzoekers in Australië en in de Verenigde Staten apart van elkaar een methode ontwikkeld waarbij proefpersonen niet op een knop drukken, maar met een vinger een touchscreen aanraken op de plek van het gevaar (Scialfa et al., 2011; Wetton et al., 2010). Uit beide onderzoeken bleek dat beginners bij deze responsmethode significant slechter scoorden dan ervaren automobilisten.

Er zijn ook nog heel andere methoden ontwikkeld. Jackson, Chapman & Crundall (2009) lieten proefpersonen ook naar films vanuit het perspectief van de bestuurder kijken, maar zetten de film stop bij het begin van het zich ontwikkelende gevaar. Proefpersonen moesten vervolgens mondeling aangeven wat er vervolgens zou kunnen gebeuren. Ervaren automobilisten konden significant vaker benoemen welk gevaar zich aan het ontwikkelen

was dan beginners. Dit was alleen het geval wanneer, nadat de film was stilgezet, het beeld van het scherm verdween, maar niet als het bevroren beeld op het scherm bleef staan. Bij de methode van Borowsky, Oron-Gilad & Parmet (2009) werd proefpersonen gevraagd om de korte filmpjes die ze gezien hadden te sorteren. Bij beginnende automobilisten bleek de veroorzaker van het gevaar het dominante selectie criterium te zijn (bijvoorbeeld gevaren veroorzaakt door voetgangers of gevaren veroorzaakt door automobilisten); bij ervaren automobilisten was het dominante selectie criterium de wegomgeving (bijvoorbeeld gevaren op kruispunten of gevaren op autosnelwegen).

Naast films worden ook rijssimulators gebruikt om gevaarherkenning te meten (zie bijvoorbeeld Crundall et al., 2012; Fisher, Pollatsek & Pradhan, 2006; Pollatsek et al., 2006; Pradhan et al., 2009; Regan, Triggs & Godley, 2000; Vlakveld et al., 2011; Wang, Zhang & Salvendy, 2010). Vooral op het Human Performance Laboratory aan de Universiteit van Massachusetts Amherst (UMASS) is veel onderzoek verricht naar het meten van gevaarherkenning in een rijssimulator. Terwijl de proefpersonen in de simulator rijden wordt de richting waarin ze van moment tot moment kijken, geregistreerd door een eyetracker. In de scenario's van de ritten zijn zowel zichtbare als verborgen potentiële gevaren opgenomen, die niet manifest worden. Dat wil zeggen dat het potentiële gevaren blijven en dat ze dus niet steeds dreigender worden, zoals wel het geval is bij de methode van McKenna & Crick. In afdeksituaties komen dus bijvoorbeeld niet plotseling auto's op botskoers ergens achter vandaan, en voetgangers die, gelet op de omstandigheden, zouden kunnen oversteken, doen dat uiteindelijk niet. Gemeten wordt of proefpersonen hun blik op die potentiële gevaren richten. Het gaat er dus niet om hoe snel proefpersonen reageren op een potentieel gevaar dat steeds dreigender wordt, maar of ze reageren op een potentieel gevaar dat een potentieel gevaar blijft. Uit onderzoeken van dit 'lab' is gebleken dat ervaren automobilisten vaker potentiële gevaren weten te detecteren dan beginners (Pradhan et al., 2005) en dat beginners die een korte training in gevaarherkenning hebben doorlopen (op een pc of in een eenvoudige rijssimulator) significant meer potentiële gevaren weten te detecteren dan ongetrainde beginners (bijvoorbeeld Fisher, Pollatsek & Pradhan, 2006; Pradhan et al., 2009; Vlakveld et al., 2011).

Foto's worden bijna niet gebruikt om gevaarherkenning te meten. De reden is dat gevaren zich ontwikkelen en op foto's niet te zien is hoe de verkeerssituatie op de foto is ontstaan. Daarnaast kan op een foto niet ingeschat worden hoe snel verkeersdeelnemers zich voortbewegen. Nederland is het enige land waarin een gevaarherkenningstest met foto's is ontwikkeld (Vlakveld, 2011). Van deze Nederlandse methode wordt verslag gedaan in *Paragraaf 2.5*.

2.4. **Gevaarherkenningstests in het rijexamen**

Lang niet alle methoden die voor wetenschappelijke doeleinden gebruikt worden om gevaarherkenning te meten, zijn geschikt om te gebruiken in het rijexamen (theorie of praktijk). Voor de wetenschap geldt dat een test valide en betrouwbaar moeten zijn. Een test is valide als de test ook daadwerkelijk meet wat beoogd wordt (in dit geval gevaarherkenning). Als automobilisten die een ongeval hebben gehad de test slechter maken dan automobilisten die geen ongeval hebben gehad, is dit geen bewijs, maar wel een indicatie

dat de gevaarherkenningstest valide is. Doordat ervaren automobilisten een veel lager ongevalsrisico hebben dan beginnende automobilisten, is het ook een indicatie van validiteit wanneer ervaren automobilisten de gevaarherkenningstest beter maken dan beginnende automobilisten. Met betrouwbaarheid wordt de nauwkeurigheid van de meetmethode bedoeld. Als bijvoorbeeld twee beoordelaars een oordeel vellen over iemands vermogen gevaren te herkennen in het verkeer, komen ze dan tot precies hetzelfde resultaat? Ook is het voor de wetenschap van belang dat wanneer meerdere toetsitems gebruikt worden om iets te meten, die items in voldoende mate met elkaar samenhangen. Hiernaar verwijst de term 'interne consistentie': als items sterk met elkaar samenhangen, is er sprake van een hoge interne consistentie.

De twee wetenschappelijke criteria, validiteit en betrouwbaarheid, gelden ook voor tests in het rijexamen. De derde eis van interne consistentie is voor een test in het rijexamen zelfs zwaarder dan voor een test voor wetenschappelijk onderzoek. Dit komt doordat het in de wetenschap vaak gaat over het vergelijken van groepen, terwijl het in een test op het rijexamen gaat om een individuele beoordeling. Naast de genoemde criteria moet een test in het rijexamen echter aan nog een aantal criteria voldoen. Zo mag de test niet fraudegevoelig zijn en dienen direct na afloop van de test de resultaten beschikbaar te zijn. De test mag ook niet te lang duren en het moet niet erg omslachtig zijn om de test af te nemen. Daarnaast moet een test een duidelijke cesuur hebben. Dit wil zeggen dat vooraf bekend moet zijn bij welke score de kandidaat nog net slaagt en bij welke score de kandidaat zakt.

Men zou kunnen beoordelen hoe goed kandidaten zijn in gevaarherkenning tijdens een examenrit. Om te kunnen meten of kandidaten in staat zijn potentiële gevaren te herkennen, dienen er in de examenrit specifieke situaties voor te komen. Vertoont de kandidaat bepaald gedrag in die specifieke situaties? Gedacht kan bijvoorbeeld worden aan een bepaalde afdeksituatie (verborgen potentieel gevaar) of aan fietsers in de regen met hun capuchon op, die zonder achterom te kijken plotseling zouden kunnen oversteken (zichtbaar potentieel gevaar). Laat de kandidaat in die specifieke situaties op een bepaald moment het gas los en kijkt hij op een bepaald moment in die situaties in de juiste richting? Ook zou men de kandidaat tijdens de rit hardop kunnen laten spreken. Noemt hij de potentiële gevaren? Niet in alle examenritten komen echter potentieel gevaarlijke situaties voor. Hierdoor is het meten van gevaarherkenning tijdens een examenrit niet valide. Als er al potentieel gevaarlijke situaties in examenritten voorkomen, zullen dat meestal niet dezelfde potentiële gevaren zijn. En als ze wel hetzelfde zijn, zullen ze verschillen in hun mate van urgentie. Dit maakt het voor examinatoren lastig om te beoordelen hoe goed kandidaten in gevaarherkenning zijn, ook al komen er in alle ritten potentieel gevaarlijke situaties voor. Dat de beoordeling van gevaarherkenning tijdens examenritten moeilijk is, maakt deze methode niet valide en onbetrouwbaar.

Alleen in Nieuw-Zeeland is getracht om gevaarherkenning tijdens examenritten te toetsen. De examenritten zijn in Nieuw-Zeeland in hoge mate gestandaardiseerd. Dat wil zeggen dat kandidaten vaste routes rijden. Hun gedrag wordt bij acht specifieke punten in die route beoordeeld. Het verkeer mag dan voor alle kandidaten op die punten anders zijn, maar de omgeving en de handeling (bijvoorbeeld op een bepaald kruispunt linksaf slaan) is wel

voor alle kandidaten gelijk. Op de acht vaste punten in de testroute wordt de kandidaten gevraagd hardop te spreken en mondeling aan te geven waarom zij een bepaalde handeling verrichten. De examiner bepaalt op basis van wat de kandidaten zeggen en hoe zij handelen in die acht specifieke situaties, hoe goed ze zijn in gevaarherkenning (NZ Transport Agency, 2012). Voor zover bekend is er (nog) geen onderzoek verricht naar de validiteit en betrouwbaarheid van deze toetsmethode. Naast de vermoedelijk toch geringe validiteit en betrouwbaarheid van deze methode, bestaat het gevaar dat leerlingen de vaste routes gaan trainen. Hierdoor komt de validiteit van het gehele praktijkexamen in gevaar.

Met behulp van rijssimulatoren kunnen de gevaarherkenningscapaciteiten van rijexamenkandidaten in principe wel betrouwbaar gemeten worden (zie *Paragraaf 2.3*). Ook is uit onderzoek gebleken dat er een duidelijk verband is tussen gevaarherkenning in een rijssimulator en gevaarherkenning in het echte verkeer (Pollatsek et al., 2011). Daarmee is het testen van gevaarherkenning in een rijssimulator ook een valide methode. Dit is zeker het geval als tijdens de testrit in de rijssimulator de oogbewegingen van de kandidaat worden geregistreerd (zie *Paragraaf 2.3*). In een rijssimulator kunnen kandidaten precies in de gewenste potentieel gevaarlijke situaties gebracht worden en het gedrag van kandidaten in die situaties (bijvoorbeeld oogbewegingen bij snelheidsaanpassingen) kan in een rijssimulator objectief en precies gemeten worden.

Toch worden tot op heden nergens in de wereld rijssimulatoren voor het rijexamen gebruikt. Het examen zou daardoor te duur worden en het is ook niet praktisch. Rijssimulatoren zijn relatief kostbaar en oogbewegingen zijn lastig te registreren in een rijssimulator. Daarnaast vergt de analyse van oogbewegingen veel tijd en zullen kandidaten na een rit in een rijssimulator niet direct weten of zij geslaagd zijn voor het onderdeel gevaarherkenning. Door de snelle ontwikkelingen op het gebied van simulatoren en het registreren van oogbewegingen, zal het op termijn waarschijnlijk wel kosteneffectief worden om het rijexamen (deels) in een rijssimulator te laten plaatsvinden. In de luchtvaart gebeurt dit al op grote schaal. Een bijkomend probleem van het gebruik van simulatoren is dat een deel van de kandidaten tijdens de rit in een simulator misselijk zal worden. Aangezien de meeste rijexamenkandidaten jong zijn, zal simulatorziekte echter waarschijnlijk weinig voorkomen (Stoner, Fisher & Mollenhauer, 2011).

In het Verenigd Koninkrijk en in Australië wordt gevaarherkenning op het rijexamen getoetst met behulp van filmbeelden. In het Verenigd Koninkrijk is de gevaarherkenningstoets in 2002 vast onderdeel van het rijexamen geworden en in Australië is dat al in het midden van de jaren negentig van de vorige eeuw gebeurd. De methode waarmee op basis van filmbeelden gevaarherkenning op het rijexamen gemeten wordt, verschilt in Australië van staat tot staat. In de staat Victoria, in het zuidoosten van Australië, krijgen kandidaten korte films te zien die zijn opgenomen vanuit het perspectief van een bestuurder. Voorafgaand aan elke film krijgt de kandidaat schriftelijk op het computerscherm te zien welke manoeuvre de camera-auto (de auto van waaruit gefilmd is) moet gaan uitvoeren in de film. Dit kan bijvoorbeeld 'remmen' zijn, maar ook 'inhalen' of 'linksaf slaan'. Het is de taak van de kandidaat om het veiligste moment aan te geven waarop de manoeuvre zou kunnen worden uitgevoerd. Dit doet hij door met een vinger het scherm aan te raken (touchscreen). Gemeten wordt in hoeverre het moment van de

kandidaat afwijkt van het ideale moment. Uit onderzoek is naar voren gekomen dat er een zwak maar significant verband is tussen de hoogte van de scores en latere ongevalsbetrokkenheid. Doordat de cesuur laag was en bijna iedereen voor de test slaagde, kon worden nagegaan of kandidaten met een lage score voor gevaarherkenning na het behalen van het rijbewijs per gereden afstand meer ongevallen hadden dan kandidaten met een hoge score. Dit bleek voor ernstige ongevallen zo te zijn (Congdon, 1999). Uit hetzelfde onderzoek kwam echter ook naar voren dat de interne consistentie van deze testmethode laag was.

De methode met filmbeelden die in 2002 in het Verenigd Koninkrijk is ingevoerd, is de eerder besproken methode van McKenna & Crick (zie *Paragraaf 2.3*). Direct voorafgaand aan de invoering van de toets heeft een grootschalig onderzoek plaatsgevonden naar de validiteit en trainbaarheid van deze testmethode (Grayson & Sexton, 2002). De validiteit viel tegen. Beginnende automobilisten die in hun korte carrière als automobilist bij een ernstig ongeval betrokken waren geweest, maakten de test niet significant slechter dan beginnende automobilisten die niet bij een ongeval betrokken waren geweest. Er werden drie tests ontwikkeld, zodat niet iedere kandidaat op het theorie-examen dezelfde test zou krijgen. Uit het onderzoek bleek dat alle drie de tests ongeveer even zwaar waren. Ook bleek dat proefpersonen door training significant hoger gingen scoren op de gevaarherkenningstoets. Later is in een grote cohortstudie nagegaan of door de invoering van de gevaarherkenningstest in het Verenigd Koninkrijk de ongevals-betrokkenheid van jonge beginnende bestuurders is afgenomen (Wells et al., 2008). De onderzoekers concluderen dat door de invoering van de test het ongevalsrisico in het eerste jaar van het rijbewijsbezit met 3% is gedaald. Aangetekend dient te worden dat het onderzoekstechnisch bijzonder moeilijk is om op basis van studies zonder controlegroep oorzakelijke effecten vast te stellen (Elvik, 2002).

Behalve dat de validiteit van de testmethode uit het Verenigd Koninkrijk niet overtuigend is aangetoond, zijn er enkele praktische bezwaren tegen deze methode. Zoals reeds in *Paragraaf 2.2* is vermeld, weet men bij deze testmethode niet waarom een kandidaat op de knop heeft gedrukt. Een tweede bezwaar is dat alleen de reactietijd als afhankelijke variabele genomen wordt. In de filmpjes wordt het gevaar steeds manifester en op het laatst valt bijna niet meer aan een ongeval te ontkomen. Kandidaten maken de test niet slecht doordat ze niet drukken, maar doordat ze laat drukken. Doordat ze tot in een vergevorderd stadium de ontwikkeling van een gevaar krijgen te zien, lekt al snel uit waar het in de films om gaat. Gewapend met deze voorkennis kunnen latere kandidaten de test beter maken dan de kandidaten van het eerste uur. Een derde bezwaar is dat het met deze methode niet goed mogelijk is om verborgen potentiële gevaren te meten. Als er bijvoorbeeld in afdeksituaties uiteindelijk geen voertuig of voetganger op botskoers ergens achter vandaan komt, wordt het gevaar nooit acuut en blijft het dus verborgen. Kandidaten die het verborgen gevaar niet herkennen, zullen dan in het geheel niet drukken en er kan dus geen reactietijd gemeten worden.

Geconcludeerd kan worden dat er wereldwijd nog geen gevaarherkennings-toetsen in het rijexamen worden gebruikt die én betrouwbaar én valide zijn en die bovendien niet fraudegevoelig zijn.

2.5. Het testen van gevaarherkenning in Nederland

Bij de gevaarherkenningstoets die in Nederland is ontwikkeld, is voor foto's gekozen omdat het met de software die destijds bij het theorie-examen werd gebruikt, niet mogelijk was om met bewegende beelden te werken. Een andere reden was dat het theorie-examen klassikaal wordt afgenomen. Hierdoor kan er niet interactief gewerkt worden en kunnen kandidaten niet ieder voor zich een bepaalde plek van een touchscreen aanraken of met de muis een bepaalde plek aanklikken. Bij individuele afname van een gevaarherkenningstoets op een pc kan dit wel.

Zoals reeds vermeld is de fototest in 2009 onderdeel geworden van het theorie-examen voor het rijbewijs B. Deze toets bestaat uit 25 foto's die elk gedurende 8 seconden in beeld zijn. Onder in beeld van elke foto is een deel van het dashboard te zien waarop de snelheid die wordt gereden staat aangegeven. Op sommige van deze foto's is een acuut gevaar zichtbaar. Meestal is dit een verkeersdeelnemer met wie men zal botsen als alles zo doorgaat. Op andere foto's is een potentieel gevaar aanwezig (zowel zichtbaar als verborgen) en op nog weer een ander deel is noch een direct gevaar noch een latent gevaar te zien. De goede respons op de foto's met een acuut gevaar is 'remmen'. De goede respons op foto's met een potentieel gevaar is 'gas loslaten' en de goede respons op foto's zonder acuut of potentieel gevaar is 'niets doen'. De respons moet gegeven worden binnen de tijd dat een foto in beeld is.

Uit een eerste onderzoek met deze testvorm bleek dat ervaren bestuurders significant meer responsen goed hadden dan rijsoepleerlingen aan het einde van hun opleiding. Ook bleek dat beginners met anderhalf jaar rijervaring die zelf een ongeval hadden gerapporteerd, de toets significant slechter maakten dan beginners met anderhalf jaar rijervaring die geen ongeval hadden gerapporteerd (Vlakveld, 2008). Zowel het verschil tussen leerlingen en ervaren bestuurders als het verschil tussen beginners met een (zelfgerapporteerd) ongeval en zonder een ongeval, is een indicatie voor de validiteit van de fototest. Uit een vervolgonderzoek met een andere set foto's bleek echter dat er geen significant verschil was tussen oudere rijsoepleerlingen (25 jaar of ouder) en ervaren automobilisten. Ook bleek bij deze tweede set foto's dat de interne consistentie feitelijk te laag was om kandidaten individueel te kunnen beoordelen. Er was bij deze set foto's nog wel een significant verschil tussen jonge rijsoepleerlingen (18 en 19 jaar) en ervaren automobilisten (Vlakveld, 2011).

Behalve dat de resultaten uit eerder onderzoek maar gedeeltelijk gerepliceerd konden worden bij een vervolgonderzoek, kleven er een aantal bezwaren aan het toetsen van gevaarherkenning met foto's. Deze bezwaren zijn:

- *Situaties waarbij snelheid moet worden ingeschat, kunnen niet getest worden.*
Op een foto is niet te zien hoe snel andere auto's rijden. Items waarbij snelheid van met name auto's een grote rol speelt zijn er daarom niet. Foto's van situaties op autosnelwegen komen om deze reden nauwelijks voor in de huidige gevaarherkenningstoets.

- Het is onduidelijk of het bij een lage score nu zit in een gebrek aan hogere ordevaardigheid of in de 'wilde haren'.*

Van de vier processen van gevaaranticipatie die genoemd worden in *Paragraaf 2.2*, meet de fototest proces 4. Er wordt immers gevraagd of kandidaten wel of niet iets willen doen aan hun snelheid. In *Paragraaf 2.2* is al gemeld dat proces 4 niet goed te meten is op een examen, omdat het deels om motivatie gaat. In de fototest wordt gevraagd wat kandidaten met betrekking tot de weergegeven snelheid zouden doen; niets, gas loslaten of remmen. Wanneer een kandidaat bijvoorbeeld als antwoord 'niets' geeft, terwijl het goede antwoord 'remmen' is, kan dit zijn omdat hij het gevaar niet heeft gezien. Kandidaten kunnen echter ook 'niets' als respons hebben gegeven, terwijl ze het gevaar wel hebben gezien, maar remmen niet nodig vinden. Zij vinden dan bijvoorbeeld dat een andere automobilist maar beter moet opletten. Het is dus niet duidelijk of iemand een toetsitem verkeerd heeft gemaakt uit onkunde (het gevaar niet zien) of uit onderschatting van het gevaar en/of overschatting van de eigen vaardigheden. Dit is vooral een probleem voor de rijopleider. Een rijopleider moet immers heel andere instructies geven aan leerlingen die de gevaren niet weten te herkennen dan aan leerlingen die de gevaren wel herkennen, maar die vervolgens de risico's onderschatten en/of hun eigen vaardigheden overschatten.
- Kandidaten kunnen zakken wanneer ze te voorzichtige antwoorden geven.*

De antwoorden 'remmen' en 'gas los' zijn fout wanneer het goede antwoord 'niets doen' is. Ook het antwoord 'remmen' is fout wanneer het goede antwoord 'gas los' is. De vraag is of het legitiem is kandidaten voor gevaarherkenning te laten zakken omdat ze (te) voorzichtig zijn.
- De fototoets noopt tot louter examentraining zonder dat gevaarherkenning getraind wordt.*

Rijinstructeurs zouden hun leerlingen bijvoorbeeld als volgt kunnen instrueren: "Als je op het examen een foto te zien krijgt waarbij je snel tegen een ander zult aanbotsen, kies dan altijd voor 'remmen'. Als je andere verkeersdeelnemers ziet waarmee je niet op botskoers ligt, maar die wel dingen kunnen doen waardoor ze met je op botskoers komen, kies dan altijd voor 'gas los'. Zie je geen verkeersdeelnemers die dingen kunnen doen waardoor je op korte termijn tegen ze aan zult botsen, kies dan voor 'niets'." Wanneer een kandidaat zich hier strikt aan houdt, zal hij niet alle 25 vragen goed hebben, maar hij zal wel voldoende vragen goed hebben om ruimschoots te slagen voor het onderdeel gevaarherkenning. Door het aanleren van een ezelsbruggetje dat weinig met gevaarherkenning te maken heeft, kunnen kandidaten dus makkelijk slagen voor het onderdeel gevaarherkenning.

Door de lage cesuur (kandidaten zijn al geslaagd als ze 13 van de 25 items goed hebben) en door vooral het laatste punt (de mogelijkheid om een goed resultaat te behalen door het aanleren van een trucje), is het niet waarschijnlijk dat de gevaarherkenningstoets met foto's in zijn huidige vorm veel heeft bijgedragen aan de verkeersveiligheid. Het is zaak dat er een gevaarherkenningstoets komt die alleen de eerste drie processen van gevaaranticipatie meet (en dus niet het vierde proces), en waarop kandidaten alleen hoog kunnen scoren als ze daadwerkelijk geleerd hebben potentiële

gevaren in het verkeer te herkennen en niet door het aanleren van een trucje.

2.6. Eerder Nederlands onderzoek met bewegende beelden

Zoals reeds vermeld, kleven er bezwaren aan de methode met films die in het Verenigd Koninkrijk op het theorie-examen wordt gebruikt. Resumerend zijn deze bezwaren:

- *De validiteit is niet overtuigend aangetoond.*
Hoewel uit de meeste onderzoeken is gebleken dat ervaren automobilisten de test beter maken dan beginnende automobilisten, zijn er ook onderzoeken waarin dat niet is gebleken. Ook is niet overtuigend gebleken dat beginners die een ongeval hebben gehad de test slechter maken dan beginners die geen ongeval hebben gehad.
- *Het is niet duidelijk of een kandidaat het gevaar heeft herkend.*
Als kandidaten op de knop drukken, is nog niet duidelijk waarom ze op de knop hebben gedrukt. Het kan zijn dat ze het gevaar 'gezien' hebben, maar het kan ook zijn dat ze drukken vanwege een andere ontwikkeling in de film.
- *De testmethode is niet geschikt voor het meten van potentieel verborgen gevaren.*
Het is bij deze testmethode essentieel dat gevaren zich ontwikkelen en steeds dreigender worden. Gebeurt dit niet, dan zou er wel eens geen respons kunnen zijn. Zonder respons kan geen reactietijd gemeten worden. Met deze methode is het dus niet mogelijk om te toetsen of kandidaten mogelijke gevaren herkennen die verborgen blijven en niet steeds dreigender worden.
- *De testmethode is fraudegevoelig.*
Doordat in de film de gevaren zich ontwikkelen en steeds dreigender worden, wordt al snel duidelijk waar het in de films om gaat. Wanneer deze kennis wordt doorverteld, kunnen latere kandidaten daar hun voordeel mee doen.

In de periode dat de Nederlandse fototest voor het theorie-examen rijbewijs B werd ontwikkeld, hebben het CBR en de SWOV ook al onderzoek gedaan naar een filmtest. Het uitgangspunt daarbij vormde de meetmethode van gevaarherkenning in een rijnsimulator met een eyetracker, zoals ontwikkeld in de Verenigde Staten aan de Universiteit van Massachusetts Amherst (zie de *Paragrafen 2.4 en 2.5*). Dat wil zeggen dat de potentiële gevaren in de films potentiële gevaren bleven en niet steeds dreigender werden. In deze filmtest keken proefpersonen naar korte animatiefilms. Het voordeel van animatiefilms boven opnamen in het werkelijke verkeer, is dat onderzoekers in animatiefilms, net als in een rijnsimulator, precies die situaties kunnen creëren die ze willen hebben. Nadat een animatiefilm is gemaakt, is het bovendien nog mogelijk de film vrij gemakkelijk te verbeteren. Films die in de werkelijkheid zijn opgenomen, moeten opnieuw gemaakt worden als onderzoekers die willen verbeteren. Een nadeel van het gebruik van animatiefilms is wel dat de omgeving wat steriel en minder gedetailleerd is dan in werkelijkheid. Door de grote populariteit van computerspellen en de

snelle ontwikkelingen op het gebied van games, beginnen geanimeerde beelden echter wel steeds meer de werkelijkheid te benaderen.

Net als bij de tests waarvan dit rapport verslag doet, duurde elke film bij deze eerste onderzoeken ongeveer 40 seconden. De animatiefilms waren 'gefilmd' vanuit het perspectief van een bestuurder. Voorafgaand aan elke film werd met een beeld 'van bovenaf' getoond welke manoeuvre de auto in de film uitvoerde. Gedurende een film werd het beeld enkele malen stilgezet. In die pauzes, die 5 seconden duurden, konden proefpersonen met hun muis op het bevroren beeld zichtbare en verborgen potentiële gevaren aanklikken. Het inlassen van de pauzes was noodzakelijk, omdat een potentieel gevaar aanwijzen meer tijd kost dan kijken naar een potentieel gevaar.

Uit de eerste proef bleek dat proefpersonen de filmtest overwegend slecht maakten en dat ervaren automobilisten de test niet significant beter maakten dan beginnende automobilisten (Vlakveld, 2008). Gedacht werd dat dit kwam doordat de proefpersonen niet goed voorbereid waren op het maken van de test en dat de potentiële gevaren in de films niet duidelijk genoeg waren. Er werd daarom een verbeterde filmtest ontwikkeld. De zichtbare en verborgen potentiële gevaren in de nieuwe films waren duidelijker, de resolutie van de nieuwe films was hoger en de films werden groter in beeld gebracht, de taak werd eenvoudiger gemaakt en proefpersonen konden zich beter voorbereiden op de taak. Er waren minder pauzes per film (drie) en de duur van een pauze was korter (3 seconden). Aan proefpersonen werd nog duidelijker dan bij de eerste versie van de filmtest verteld dat klikken op plekken waar geen potentieel gevaar was, hun score zou verminderen en dat niet in alle pauzes potentiële gevaren voorkwamen. Ook bij deze verbeterde filmtest bleek er echter geen significant verschil te zijn in het aantal aangeklikte potentiële gevaren tussen rijsschoolleerlingen en ervaren, in dit geval zelfs professionele bestuurders (rijinstructeurs en rijexaminatoren; Vlakveld, 2011). Wel bleek uit dit tweede experiment dat proefpersonen die vaak computergames speelden de test significant beter maakten dan proefpersonen die nooit of vrijwel nooit computergames speelden. Vermoed werd dat de pauzes hierbij een rol speelden. Het zou kunnen dat proefpersonen die voor de pauze nog in het geheel geen potentieel gevaar 'gezien' hebben, door de pauze op het idee worden gebracht dat er een potentieel gevaar zou kunnen zijn en dan naar dat potentiële gevaar gaan zoeken (McGowan & Banbury, 2004). Waarschijnlijk kunnen personen die vaak computergames spelen ook sneller met hun muis zaken aanklikken op een scherm dan personen die nooit of bijna nooit computergames spelen.

Om na te gaan of het probleem in de responsmethode zat (de pauzes en het in korte tijd aanklikken op een scherm), werden de films zonder pauzes aangeboden op een computerscherm waar een eyetracker was ingebouwd. Dit was het derde experiment. Van de proefpersonen die naar de films keken en zich moesten inbeelden dat zij de bestuurder waren, registreerde de eyetracker de oogbewegingen en de fixaties. Daarnaast werd direct na afloop van iedere film aan de proefpersonen gevraagd wat er had kunnen gebeuren (maar wat dus niet gebeurd was) waardoor er een grote kans op een ongeval zou zijn ontstaan. Het bleek dat ervaren automobilisten significant vaker in de richting van verborgen potentiële gevaren keken dan rijsschoolleerlingen. Na afloop benoemden ervaren bestuurders verborgen

potentiële gevaren ook significant vaker (Vlakveld, 2011). Uit dit resultaat kan afgeleid worden dat er inderdaad duidelijke verschillen zijn tussen ervaren automobilisten en rijsschoolleerlingen in gevaarherkenning bij het kijken naar animatiefilms. Die verschillen komen echter niet naar voren als potentiële gevaren aangewezen moeten worden tijdens pauzes gedurende de films. Er moet dus naar een andere responsmethode gezocht worden.

Eyetrackers kunnen (nog) niet gebruikt worden op het theorie-examen. Ook is het stellen van een open vraag na afloop van iedere film (wat had er in deze film voor gevaarlijks kunnen gebeuren, wat niet gebeurd is?) geen geschikte responsmethode voor een theorie-examen dat jaarlijks door vele duizenden leerlingen gemaakt wordt.

De conclusies die getrokken kunnen worden uit de Nederlandse experimenten om een beter alternatief te ontwikkelen voor de methode die in het Verenigd Koninkrijk gebruikt wordt, is dat films met onderbrekingen niet werken. Wel is gebleken dat ervaren bestuurders significant vaker potentiële gevaren na afloop noemen en ook significant vaker naar potentiële gevaren kijken.

Het onderzoek waar onderhavig rapport verslag van doet, heeft tot doel voor het theorie-examen een gevaarherkenningstoets te ontwikkelen met een responsmethode waarbij de films niet onderbroken worden. Ook mag de methode niet de beperkingen kennen van de methode die McKenna & Crick (1997) ontwikkelden.

3. Methode van onderzoek

In een test waarin vaardigheden getoetst worden – en in dit geval is die vaardigheid ‘gevaarherkenning’ – gaat het erom kandidaten in zulke omstandigheden te brengen dat uit hun reacties afgeleid kan worden dat ze een bepaalde vaardigheid beheersen. Die omstandigheden zijn de *stimuli* en de reacties zijn de *responsen*. Op basis van de literatuur en de ervaringen uit het eerdere onderzoek in Nederland zijn voor dit onderzoek twee nieuwe testvormen bedacht. De stimuli zijn in beide testvormen gelijk en bestaan uit korte animatiefilms waarin gevaren verborgen zijn. De twee responsmethoden verschillen. Dit hoofdstuk beschrijft de twee tests en de bijbehorende responsmethoden (*Paragraaf 3.2*). Hieraan voorafgaand worden in *Paragraaf 3.1* de animatiefilms besproken die bij de tests zijn gebruikt. De steekproeven, de apparatuur die gebruikt is, de procedure tijdens de testafname en de manier waarop de data zijn geanalyseerd, staan beschreven in de *Paragrafen 3.3 tot en met 3.7*.

3.1. De animatiefilms voor de twee responsmethoden

Bij de twee responsmethoden die geëvalueerd zijn, zijn dezelfde dertien animatiefilms gebruikt. In iedere film zit een dominant potentieel gevaar. Dit kan een zichtbaar potentieel gevaar of een verborgen potentieel gevaar zijn. In geen van de films wordt dit potentiële gevaar manifest. In de *Bijlage* staat beschreven wat het dominante potentiële gevaar per film is. De volgorde waarin de films vertoond zijn, is bij beide responsmethoden gelijk aan de volgorde waarin de films in de *Bijlage* beschreven zijn. Naast het dominante verborgen gevaar bevinden zich in elke film nog enkele nevengevaren waarvan het minder aannemelijk is dat ze manifest zullen worden. Alles is mogelijk, maar niet alles is even waarschijnlijk. Zonder die andere, minder duidelijke potentiële gevaren zou de test te gemakkelijk worden. Zo'n minder groot nevengevaar kan bijvoorbeeld een auto op botskoers zijn die nog tamelijk ver weg is en waarvan het heel duidelijk is dat die geen voorrang heeft. Als in de film te zien is dat de bestuurder van die auto in de richting van de auto kijkt van waaruit de film is ‘opgenomen’ (in het vervolg de camera-auto genoemd), is het niet erg waarschijnlijk dat die auto geen voorrang zal verlenen.

Tien van de dertien films zijn verbeterde versies van de films die gebruikt zijn bij eerdere experimenten (zie *Paragraaf 2.6*; Vlakveld, 2008; 2011). Die films zijn verbeterd door het dominante verborgen gevaar wat dominanter te maken en de nevengevaren in de films wat minder dominant te maken. In twee films is een oorspronkelijk nevengevaar omgewerkt tot het dominante gevaar en is het oorspronkelijke dominante gevaar eruit gehaald. Er is voor het testen van de twee responsmethoden één compleet nieuwe film ontwikkeld.

De scenario's van de oorspronkelijke films zijn ontwikkeld in nauwe samenwerking met een groep van acht rijexaminatoren. Aan die examinatoren is uitgelegd wat gevaarherkenning is en wat het verschil is tussen zichtbare potentiële gevaren en verborgen potentiële gevaren. Vervolgens is aan die examinatoren gevraagd voorbeelden uit de praktijk te noemen waarbij een kandidaat een potentieel zichtbaar gevaar of een potentieel verborgen

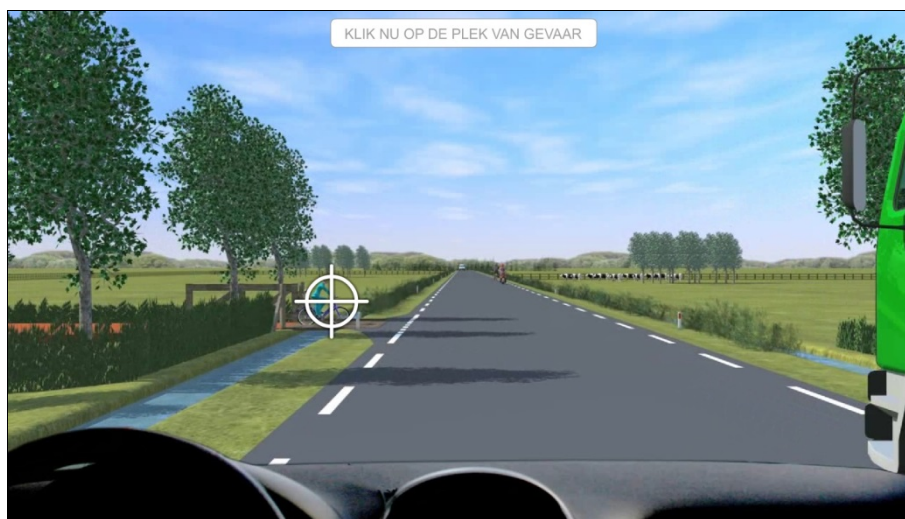
gevaar niet had opgemerkt. Van de meest exemplarische voorvallen zijn filmscenario's opgesteld en aan de hand van die scenario's zijn vervolgens animatiefilms gemaakt. Deze films zijn vervaardigd met het softwarepakket Cinema 4d. Het scenario voor de geheel nieuw vervaardigde film is op een andere wijze tot stand gekomen. Geen van de oorspronkelijke films bevatte een mogelijk ongeval met een fietser. Het aandeel ongevallen waar fietsers bij betrokken zijn, neemt in Nederland toe (Reurings et al., 2012). Het is bekend dat bij tweerichtingsfietspaden automobilisten fietsers uit de ongebruikelijke richting regelmatig niet opmerken (Räsänen & Summala, 1998; Schepers et al., 2011; Summala et al., 1996). Daarom is een film gemaakt met een verborgen potentieel gevaar waarbij de camera-auto afslaat en een tweerichtingsfietspad kruist. Dit is film 5 in de *Bijlage*.

3.2. De twee geëvalueerde responsmethoden

3.2.1. Responsmethode 1

Nadat ze een korte film hebben gezien, kunnen ervaren automobilisten vaker dan beginners in woorden zeggen wat het potentiële gevaar was in die film. Als ze een film voor de tweede keer zien, direct nadat ze de film voor de eerste keer hebben gezien, moeten ervaren automobilisten potentiële gevaren dus ook beter kunnen aanwijzen dan beginners. Op deze aanname is responsmethode 1 gebaseerd.

Bij responsmethode 1 krijgen proefpersonen een film eerst in zijn geheel te zien. Vooraf wordt op het scherm gemeld dat de proefpersoon zich bij het kijken naar de film moet voorstellen de bestuurder van de camera-auto te zijn. De proefpersoon start zelf een film door met de muis op de icoon 'start' te klikken. Op het scherm met de startknop staat ook een gestileerde plattegrond van de verkeerssituatie, waarin van bovenaf de camera-auto te zien is, met een korte mededeling van de manoeuvre die de camera-auto in de film gaat uitvoeren (bijvoorbeeld: "U slaat rechtsaf"). Nadat de proefpersoon een film gezien heeft, komt een scherm in beeld met de mededeling dat hij de film voor de tweede keer kan starten. Daarbij staat wel dat hij, voordat hij dat doet, voor zichzelf eerst moet bepalen op welk moment van de film hij het sterkste dacht: "Als dat nu maar niet gebeurt". Vervolgens start de proefpersoon de film voor de tweede keer en stopt hij de film door op de spatiebalk te drukken op het moment dat hij vooraf in gedachten genomen heeft. Het filmbeeld staat nu stil en de icoon van een vizier komt in beeld. De proefpersoon moet deze icoon met de muis naar de plek slepen waar zich volgens hem het zichtbare potentiële gevaar bevindt, of waar het verborgen potentiële gevaar voor het eerst zichtbaar zou kunnen worden. Door een muisklik wordt de icoon op die plek vastgezet. *Afbeelding 3.1* geeft het moment weer waarop het vizier naar het potentiële gevaar is gesleept, maar nog niet met de muis is geklikt om deze plek te bevestigen.



Afbeelding 3.1. Voorbeeld van een scherm dat gekozen is om het potentiële gevaar aan te wijzen. Het vizier is in de richting van dat potentiële gevaar gebracht, maar het gevaar is nog niet aangeklikt.

Drie seconden nadat de proefpersoon met de muis de plek heeft aangewezen en heeft geklikt, komt het startscherm van de volgende film in beeld.

3.2.2. Responsmethode 2

Een nadeel van responsmethode 1 is dat proefpersonen de mogelijke gevaren niet aangeven op het moment dat ze die herkennen. Achteraf moeten ze bedenken wat het grootste potentiële gevaar is geweest. Bij responsmethode 1 gaat het er dus vooral om met terugwerkende kracht te bepalen wat belangrijk was en wat niet. Vlakveld (2011) deed een test waarbij proefpersonen naar korte films met een verborgen potentieel gevaar keken, terwijl hun oogbewegingen werden geregistreerd en ze na afloop van iedere film de potentiële gevaren moesten benoemen. In 37,0% van de gevallen werden de verborgen potentiële gevaren gefixeerd en ook na afloop benoemd. In 30,2% van de gevallen fixeerden proefpersonen wel gevaren, maar benoemden ze die niet. In 5,7% van de gevallen werden gevaren niet gefixeerd maar wel benoemd en in 27,1% van de gevallen werden de potentiële verborgen gevaren noch gefixeerd noch benoemd. Er wordt dus meer gefixeerd dan dat er achteraf benoemd wordt. Bij verborgen potentiële gevaren is het onwaarschijnlijk dat de fixaties niet intentioneel zijn. Men zal niet vaak in richtingen kijken waar niets interessants te zien is. Waarschijnlijk kijkt iemand bij het passeren van een geparkeerde bus alleen vlak voor die bus langs als hij verwacht dat direct vóór de bus uitgestapte passagiers kunnen oversteken. Wanneer proefpersonen na de film niet de mogelijkheid noemen dat passagiers voorlangs kunnen lopen, terwijl ze wel voor die bus langs hebben gekeken, dan kan dat twee oorzaken hebben. Ze kunnen het inmiddels vergeten zijn, omdat het potentiële gevaar niet manifest is geworden. Ze kunnen ook uit routine voor de bus langs hebben gekeken, omdat in dergelijke situaties voor de bus langs kijken al in de mentale representaties (schemata) is opgenomen (Vlakveld, 2011). In dergelijke gevallen reageert iemand wel op potentiële gevaren, maar dringt het gevaar zelf niet meer tot het werkgeheugen door. Na afloop zal hij

routinegevaren niet snel benoemen, omdat hij die niet echt bewust heeft meegemaakt.

Bij responsmethode 2 is getracht ervoor te zorgen dat proefpersonen de potentiële gevaren markeren op het moment dat ze die potentiële gevaren ervaren, maar dat door dit markeren geen interrupties plaatsvinden en de film gewoon blijft doorlopen. Bij deze responsmethode krijgen proefpersonen elk filmpje slechts één keer te zien. Wanneer ze voelen dat er een potentieel gevaar is, drukken ze op de spatiebalk. Onder in beeld loopt een tijdbalk. Wanneer de proefpersoon op de spatiebalk heeft gedrukt, hoort hij een toon en komt op de tijdbalk een witte markering te staan op het tijdstip waarop hij heeft gedrukt, maar de film wordt niet onderbroken. Men kan maximaal vier keer per film op de spatiebalk drukken. Direct na afloop van elke film krijgt de proefpersoon de stilstaande beelden te zien van de momenten waarop hij heeft gedrukt. Als hij vier keer heeft gedrukt krijgt hij vier beelden te zien, bij drie keer drukken drie, enzovoort. Bij één keer drukken komt na afloop van de film direct dat ene moment groot in beeld. *Afbeelding 3.2* geeft het scherm weer dat na afloop van een film in beeld komt bij een proefpersoon die vier keer heeft gedrukt.



Afbeelding 3.2. Voorbeeld van een scherm dat direct na afloop in beeld komt bij een proefpersoon die tijdens het zien van de film vier keer op de spatiebalk heeft gedrukt om een potentieel gevaar te markeren.

Van de, in dit geval, vier momenten dient de proefpersoon het beeld uit te kiezen waarop hij meent dat het grootste potentiële gevaar te zien is, door op die afbeelding te klikken. Deze afbeelding komt daarna groot in beeld. Wanneer de proefpersoon tijdens de film maar één keer heeft gedrukt, komt dat moment dus al direct na de film groot in beeld. Vanaf het moment dat het gekozen beeld groot in beeld is, verloopt de test op dezelfde wijze als bij responsmethode 1. Dat wil zeggen dat er een icoon van een vizier in beeld komt en dat de proefpersoon die icoon naar de plek van het potentiële gevaar sleept en daarna met de muis klikt. Drie seconden nadat een proefpersoon met de muis de plek heeft aangewezen en heeft geklikt, komt het startscherm van de volgende film in beeld. Bij elk startscherm wordt, net als bij responsmethode 1, een plattegrond van de verkeerssituatie getoond en staat aangegeven welke manoeuvre de camera-auto in de film zal uitvoeren.

3.3. Deelnemers aan het onderzoek

De proefpersonen hebben óf de test met responsmethode 1 óf de test met responsmethode 2 gemaakt. Bij elk van de twee responsmethoden waren er twee groepen: een groep rijsschoolleerlingen en een groep professionals. Zowel de groep professionals die de test met responsmethode 1 heeft gemaakt als de groep professionals die de test met responsmethode 2 heeft gemaakt, bestond uit rijinstructeurs en rijexaminatoren. Vanuit twee examencentra van het CBR, namelijk Arnhem en Rijswijk, zijn rijsscholen in de omgeving van die centra benaderd of ze wilden deelnemen aan het onderzoek (zowel de leerlingen als de instructeurs). De rijexaminatoren die hebben deelgenomen aan het onderzoek zijn benaderd op de twee examencentra.

Deelnemers test responsmethode 1

- *Rijsschoolleerlingen*
N = 30. De jongste is 16,9 jaar en de oudste is 57,9 jaar. 23,3% is jonger dan 18 jaar. De gemiddelde leeftijd is 22,5 jaar ($SD = 9,8$). Gemiddeld hebben de leerlingen 28,7 uur rijles gehad op het moment dat ze de test maakten ($SD = 25,3$). Het minste aantal uren rijles is 0 en het meeste aantal uren rijles is 100. 53,3% is man.
- *Professionals*
N = 34. De jongste is 22,3 jaar en de oudste is 67,3 jaar. De gemiddelde leeftijd is 44,3 jaar ($SD = 10,6$). 82% is rijinstructeur en 18% is rijexaminator. Gemiddeld rijden de professionals 41.359 km per jaar in de auto als instructeur respectievelijk als examinator ($SD = 23.260$). 79,4% is man.

Deelnemers test responsmethode 2

- *Rijsschoolleerlingen*
N = 30. De jongste is 17,1 jaar en de oudste is 44,4 jaar. 6,7% is jonger dan 18 jaar. De gemiddelde leeftijd is 21,4 jaar ($SD = 6,4$). Gemiddeld hebben de leerlingen 35,6 uur rijles gehad ($SD = 22,5$). Het minste aantal uren rijles was 0 en het meeste aantal uren rijles was 113. 46,7% is man.
- *Professionals*
N = 30. De jongste is 22,5 jaar en de oudste is 63,9 jaar. De gemiddelde leeftijd is 44,9 jaar ($SD = 12,2$). 64% is rijinstructeur en 36% is rijexaminator. Gemiddeld rijden de professionals 54.187 km per jaar in de auto als instructeur respectievelijk als examinator ($SD = 58.512$). 90,0% is man.

Wat betreft de rijsschoolleerlingen zijn er tussen de steekproef die de test met responsmethode 1 heeft gemaakt en de steekproef die de test met responsmethode 2 heeft gemaakt, geen significante verschillen in leeftijd, geslacht en aantal uren rijles. Ook bij de professionals zijn er geen significante verschillen in leeftijd, geslacht, jaarkilometrage en instructie/examen. Ook zitten er niet significant meer rijexaminatoren in de steekproef van de test met responsmethode 1 dan in de steekproef van de test met

responsmethode 2. Om dit te testen is voor leeftijd, aantal uren rijles en aantal uren dat men rijinstructie geeft dan wel rijexamens afneemt, een t-toets voor twee onafhankelijke steekproeven gebruikt. Voor de categorische variabelen 'geslacht' en 'rijinstructeur of rijexaminator' is getoetst met de chi-kwadraattoets.

3.4. **Materiaal en apparatuur**

Zowel bij de test met responsmethode 1 als bij de test met responsmethode 2 zaten de proefpersonen voor een flatscreenmonitor met een beeld diagonaal van 21,5 inch en een breedte-hoogteverhouding van 16:9. De animatiefilms werden full screen vertoond en hadden een resolutie van 1.280 x 720 pixels. De beeldfrequentie was 25 beelden per seconde. De afstand tussen de ogen van de proefpersonen en het scherm bedroeg ongeveer 60 centimeter. Dit gaf de proefpersonen een beeldhoek van ongeveer 74° in het horizontale vlak.

Zoals reeds in *Paragraaf 3.4* beschreven, stopten de proefpersonen bij de test met responsmethode 1 de film door bij het voor de tweede keer zien van de film op de spatiebalk te drukken. Vervolgens klikten ze op het stilstaande beeld met hun muis de plek met het potentiële gevaar aan. De software van de test schreef de responsen (reactietijden en coördinaten) per proefpersoon in een klein bestand weg. Bij de test met responsmethode 2 werden van iedere proefpersoon per film de momenten waarop hij op de spatiebalk drukte om een potentieel gevaar aan te geven, in een klein bestand weggeschreven. In dit bestandje werden ook de coördinaten geregistreerd van de plek waar zich volgens de proefpersonen per film het grootste potentiële gevaar bevond. Dit was de plek die ze met de muis hadden aangeklikt op het beeld van het moment waarop volgens hen sprake was van het grootste potentiële gevaar.

Na afloop van de test werd de proefpersonen verzocht om een korte vragenlijst in te vullen waarin naar demografische gegevens werd gevraagd. In deze vragenlijst werd ook onder meer gevraagd hoe vaak ze met computers werkten, hoe vaak ze computergames speelden en of ze de test geschikt achtten als gevaarherkenningstest in het theorie-examen.

3.5. **Procedure**

Vanuit de examencentra van het CBR werd contact gelegd met rij scholen in de omgeving van die centra. De rij scholen kregen schriftelijk uitgelegd wat het onderzoek inhield, wat het doel was en hoe de privacy van de proefpersonen werd gewaarborgd. Wie wilde deelnemen aan het onderzoek diende een toestemmingsverklaringsformulier (informed consent) te ondertekenen. Indien de proefpersonen jonger dan achttien jaar waren, werd die verklaring ook ondertekend door een van de ouders of verzorgers. De rijexaminatoren die aan het onderzoek deelnamen, zijn geworven op de examencentra. Naast de toestemmingsverklaring dienden de deelnemers ook een geheimhoudingsverklaring te ondertekenen, omdat het materiaal van het onderzoek gebruikt kan gaan worden in het toekomstig theorie-examen. Experimentatoren van het CBR gingen met de apparatuur naar de rij scholen toe om daar de tests af te nemen. De rijexaminatoren werden getest op de examencentra.

Nadat de proefpersonen de twee formulieren ondertekend hadden, kregen ze een PowerPointpresentatie te zien. Deze presentatie werd automatisch afgespeeld op het scherm waarop ook de toets werd afgenomen. De proefpersonen hadden tijdens deze presentatie een koptelefoon op. Alle geschreven tekst in de PowerPoint kregen ze ook te horen. In de presentatie werd duidelijk gemaakt wat potentiële gevaren in het verkeer zijn en aan de hand van beelden werd het verschil tussen potentieel zichtbare gevaren en potentieel verborgen gevaren duidelijk gemaakt. In de PowerPoint werden verder instructies gegeven over wat men moest doen tijdens de test en werd verteld waarop men werd beoordeeld (het juiste moment – niet te vroeg en niet te laat – en de juiste plaats van het potentiële gevaar). Met een voorbeeldfilm werd in de PowerPoint nog eens extra geïllustreerd wat men moest doen. De presentatie duurde voor zowel de test met responsmethode 1 als de test met responsmethode 2 ongeveer 6,5 minuut.

Na de presentatie en voordat de proefpersonen met de test begonnen, kregen ze de gelegenheid om te oefenen. De twee proefopgaven bestonden uit een film met een dominant zichtbaar potentieel gevaar en een film met een dominant verborgen potentieel gevaar. Na iedere proefopgave kregen de proefpersonen te zien wat het juiste moment en de juiste plek was. Bij de test met responsmethode 1 duurde het oefenen ongeveer 6 minuten en bij de test met responsmethode 2 duurde het oefenen ongeveer 5 minuten.

Vervolgens werd de daadwerkelijke test afgenomen. Bij de test met responsmethode 1 duurde dit ongeveer 15 minuten en bij de test met responsmethode 2 ongeveer 11 minuten.

Na de test vulden de proefpersonen een korte vragenlijst in. Ze werden bedankt voor hun deelname en alleen de rijsschoolleerlingen ontvingen een waardebon van € 30,-.

De gehele procedure nam bij de test met responsmethode 1 ongeveer 40 minuten in beslag en bij de test met responsmethode 2 ongeveer 35 minuten.

3.6. Van respons naar score

Zowel bij de test met responsmethode 1 als bij de test met responsmethode 2 schreef de software per film de reactietijden (de tijd vanaf het begin van de film tot aan het moment dat de proefpersoon op de spatiebalk had gedrukt) en de coördinaten van de plek met het potentiële gevaar dat de proefpersoon met de muis had aangeklikt, in een bestand weg. Dit zijn nog geen scores. Om tot scores te komen is van iedere film vastgesteld op welk moment van de film het dominante potentiële gevaar voor het eerst te 'zien' was en op welk moment het te laat was om een botsing te voorkomen, als het dominante potentiële gevaar manifest zou zijn geworden. Dit *tijdvenster* van het dominante potentiële gevaar per film hebben drie experts van het CBR in gezamenlijk overleg vastgesteld. De drie experts hebben ook vastgesteld wat binnen het tijdvenster de plek van het dominante potentiële gevaar was. Dit kon een andere verkeersdeelnemer zijn die zich, gelet op de omstandigheden, onveilig kon gaan gedragen (zichtbare potentiële gevaren), of de plek die de richting aangaf waar een mogelijke verkeersdeelnemer op botskoers voor het eerst zichtbaar zou worden als deze ergens achter vandaan kwam (verborgen potentiële gevaren). Dit werden de

doelgebieden genoemd. Er zijn twee scoringsmethoden ontwikkeld die bij zowel de test met responsmethode 1 als de test met responsmethode 2 zijn toegepast. Scoringsmethode 1 houdt geen rekening met de reactietijden binnen een tijdvenster. Bij scoringsmethode 2 wordt de hoogte van de score deels bepaald door het moment waarop iemand binnen het tijdvenster reageert.

3.6.1. Scoringsmethode 1

Als in een film binnen het tijdvenster het doelgebied is aangeklikt, is de respons goed. Klikt men binnen het tijdvenster de verkeerde plek aan, dan is de respons fout. De respons is ook fout als de film buiten het tijdvenster wordt gestopt, ongeacht de plek die de proefpersoon aanwijst. Een goede score levert 1 punt op en een foute score levert 0 punten op. De eindscore over de dertien films wordt bepaald door het optellen van alle punten. De hoogst mogelijke score is dus 13.

3.6.2. Scoringsmethode 2

Bij deze methode wordt niet alleen gekeken of proefpersonen binnen het tijdvenster het juiste doelgebied hebben aangewezen, maar wordt ook rekening gehouden met het moment waarop ze binnen dat tijdvenster op de spatiebalk hebben gedrukt. Maximaal kan men per film 10 punten krijgen. Zit de proefpersoon binnen het tijdvenster en heeft hij de goede plek aangewezen, dan krijgt hij in ieder geval al 5 punten. Het aantal punten wordt hoger naarmate de proefpersoon meer aan het begin van het tijdvenster op de spatiebalk heeft gedrukt. Heeft hij helemaal aan het begin van het tijdvenster op de spatiebalk gedrukt, dan krijgt hij 10 punten; heeft hij helemaal aan het einde van het tijdvenster op de spatiebalk gedrukt, dan blijven het 5 punten. Het aantal punten loopt lineair af binnen het tijdvenster. De duur waarin het potentiële gevaar aanwezig is, verschilt per film. Des te langer het tijdvenster is, des te trager lopen de punten die men krijgt, af. Bij de meeste films zit er een duidelijk begin aan het dominante potentiële gevaar, maar bij de films 1, 11 en 12 (zie *Bijlage*) is het dominante gevaar al ver van te voren te zien, zonder dat het op dat moment al nodig is om door anticiperende maatregelen (in het oog houden, gas loslaten) de veiligheidsmarge te vergroten. In de introductie die iedere proefpersoon te zien en te horen kreeg, is gemeld dat hij niet te vroeg en niet te laat op de spatiebalk moest drukken en dat het 'juiste moment' het moment is dat hij in werkelijkheid anticiperende maatregelen zou nemen. Bijvoorbeeld in film 1 zijn de containers op de hoek van de T-splitsing al aan het begin van de film heel in de verte te zien. Deze containers schermen het zicht op mogelijk naderend verkeer af. In gevallen waarbij het potentiële gevaar al heel vroeg te zien is, maar het nog geen tijd is voor anticiperende maatregelen, ligt het moment binnen het tijdvenster waarop proefpersonen 10 punten scoren niet aan het begin van het tijdvenster, maar verderop in het tijdvenster. Dit is het moment waarop anticiperende maatregelen pas zin beginnen te krijgen. Indien proefpersonen voor dat moment op de spatiebalk hebben gedrukt, krijgen ze meer dan 5, maar minder dan 10 punten. Ook dit is lineair gedaan. Klikt iemand na dat moment dan loopt de score op dezelfde wijze af als bij de scores waarbij er een duidelijk moment is waarop het tijdvenster start. Maximaal kan men bij deze scoringsmethode 10 punten per film scoren. De eindscore is bepaald door de scores per film op te tellen en door 13 te delen. De maximale eindscore is dus 10 punten.

3.6.3. *Het automatisch bepalen van de score*

Voor eventueel toekomstig gebruik in het theorie-examen is het belangrijk dat de score direct bekend is na afloop van de test en dat beoordeling van examinatoren (heeft de kandidaat nu wel of niet het gevaar herkend?) geen rol speelt bij het bepalen van de score. Omwille van de betrouwbaarheid dient immers elke kandidaat volgens precies dezelfde criteria beoordeeld te worden en mag er geen ruimte zijn voor interpretatie. Hiertoe is software ontwikkeld die de data van de tests (reactietijden en coördinaten) omzet in een score volgens scoringsmethode 1 en volgens scoringsmethode 2.

Voor deze software is binnen het tijdvenster van het dominante potentiële gevaar van iedere film per filmbeeld – er gaan 25 filmbeelden in 1 seconde – vastgesteld wat het doelgebied is van het dominante potentiële gevaar. Doordat het om films gaat waarin de camera-auto en veel van de verkeersdeelnemers bewegen, verschilt het doelgebied enigszins van filmbeeld tot filmbeeld. De doelgebieden per filmbeeld werden gevormd door een of meer vierkanten waarvan de coördinaten werden bepaald. Een muisklik binnen die gebieden leverde bij scoringsmethode 1 de score 1 op. Bij scoringsmethode 2 was de hoogte van de score mede afhankelijk van het moment in de film waarop binnen het tijdvenster op de spatiebalk was gedrukt en of op dat filmbeeld het gebied binnen de coördinaten was aangeklikt.

3.7. **Design en analyse van de data**

De afhankelijke variabelen in de analyse waren de eindscores volgens scoringsmethode 1 en volgens scoringsmethode 2. De onafhankelijke variabelen waren de testmethode (de test met responsmethode 1 of de test met responsmethode 2) en de groep van proefpersonen (rijschoolleerlingen of professionals). Het ging om een tussenproefpersoonontwerp (between subjects design) waarbij de gemiddelden van de eindscores (volgens beide scoringsmethoden) van rijschoolleerlingen en professionals met elkaar zijn vergeleken. Ook zijn de scores per film tussen beide groepen met elkaar vergeleken. Nagegaan is of zowel de eindscores (volgens beide scoringsmethoden) als de scores per film (volgens beide scoringsmethoden) tussen rijschoolleerlingen en professionals significant van elkaar verschilden ($p < 0,05$).

Van de eindscores bij scoringsmethode 1 en scoringsmethode 2 is vooraf nagegaan of aan de voorwaarden van parametrisch testen werd voldaan. In alle gevallen werd hieraan voldaan. Bij het vergelijken van de scores per film was vooraf al duidelijk dat niet parametrisch getoetst kon worden. Bij scoringsmethode 1 is per film de score goed of fout (een categorische variabele dus) en bij scoringsmethode 2 is per film de score of 0, of een getal tussen 5 en 10. Bij de vergelijking tussen beide groepen per film bij scoringsmethode 1 is de Fisher's exact test gebruikt. Bij de vergelijking tussen beide groepen per film bij scoringsmethode 2 is de Mann-Whitney's U-test gebruikt.

Naast significantie is ook naar effectgrootte gekeken. De effectgrootte zegt in feite meer dan de significantie over de robuustheid van het verschil tussen twee groepen, omdat de significantie afhankelijk is van de grootte van de steekproef en de effectgrootte niet. Bij de parametrische toetsen is voor de effectgrootte de partiële η kwadraat (partiële η^2) gebruikt met partiële

$\eta^2 = 0,01$ als klein, partiële $\eta^2 = 0,06$ als gemiddeld en partiële $\eta^2 = 0,14$ als groot (Cohen, 1988). In geval van de Fisher's exact test is niet naar de effectgrootte gekeken, maar is de odds ratio vastgesteld om een indruk te krijgen van de robuustheid van het verschil tussen beide groepen. Bij gebruik van de Mann-Whitney's U-test is de effectgrootte uitgedrukt in de Pearson's correlatiecoëfficiënt (r) met $r = 0,10$ als klein (1% van de variantie die verklaard wordt), $r = 0,30$ als gemiddeld (9% van de variantie die verklaard wordt) en $r = 0,50$ als groot (25% van de variantie die verklaard wordt) (Cohen, 1988).

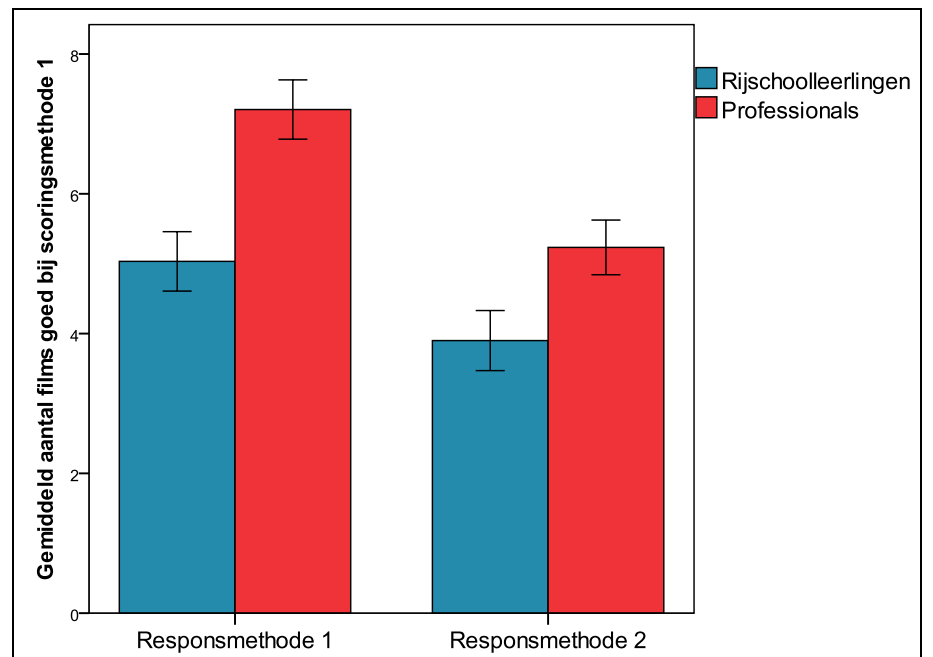
Ten slotte is tentatief naar de interne consistentie van beide tests gekeken. Vormen de films één schaal die tezamen één achterliggend construct meten (in dit geval gevaarherkenning)? Hiervoor is de Cronbach's alpha (α) genomen met $0,5 < \alpha \leq 0,6$ als een slechte interne consistentie, $0,6 < \alpha \leq 0,7$ als een twijfelachtige interne consistentie, $0,7 < \alpha \leq 0,8$ als een acceptabele interne consistentie en $0,8 < \alpha \leq 0,9$ als een goede interne consistentie. Er staat tentatief, omdat de Cronbach's alpha pas goed bepaald kan worden bij een grote steekproef. Bij kleine steekproeven, zoals in het onderhavige onderzoek, kunnen afwijkende scores van een enkele proefpersoon nog tot grote veranderingen in de waarde van de Cronbach's alpha leiden.

4. Resultaten

Het onderzoek heeft verschillende data opgeleverd, die in dit hoofdstuk worden gepresenteerd en geanalyseerd. Allereerst worden per responsmethode en scoringsmethode de scores van rijsschoolleerlingen en professionals vergeleken. Dat gebeurt in *Paragraaf 4.1* voor scoringsmethode 1 en in *Paragraaf 4.2* voor scoringsmethode 2. *Paragraaf 4.3* bevat de conclusies over de scoringsmethoden en responsmethoden. Ook kan specifiek per film worden gekeken hoe goed de verschillende groepen hebben gescoord met de verschillende scoringsmethoden. Dat gebeurt in *Paragraaf 4.4* voor scoringsmethode 1 en in *Paragraaf 4.5* voor scoringsmethode 2. *Paragraaf 4.6* bevat de conclusies hierover. De vragenlijsten die de proefpersonen hebben ingevuld worden geanalyseerd in *Paragraaf 4.7*. *Paragraaf 4.8* bespreekt hoe vaak de verschillende groepen (rijsschoolleerlingen en professionals) een gevaar hebben gemarkeerd bij responsmethode 2. Tot wordt in *Paragraaf 4.9* ingegaan op de interne consistentie van de tests en de vraag hoe die verbeterd zou kunnen worden. Om op het theorie-examen rijbewijs B gevaarherkenning met behulp van bewegende beelden te kunnen testen, is een hoge interne consistentie noodzakelijk.

4.1. Eindscores responsmethoden 1 en 2 bij scoringsmethode 1

Bij scoringsmethode 1 is een score op een film goed of fout. Omdat er dertien films zijn, is de maximale eindscore 13. Dit geldt zowel voor de test met responsmethode 1 als voor de test met responsmethode 2 (zie *Paragraaf 3.6.1*). *Afbeelding 4.1* geeft de gemiddelde eindscores weer van rijsschoolleerlingen en professionals bij de test met responsmethode 1 en de test met responsmethode 2.



Afbeelding 4.1. Gemiddeld aantal films goed (scoringsmethode 1) bij responsmethode 1 en responsmethode 2 van rijsschoolleerlingen en professionals. De foutenbalken geven +/- 1 maal de standaardfout (SE) weer.

Bij de test met responsmethode 1 hebben rijsschoolleerlingen gemiddeld 5,03 films goed ($SD = 2,33$) en hebben professionals gemiddeld 7,21 films goed ($SD = 2,47$). Dit verschil is statistisch significant en de effectgrootte is groot, $F(1,62) = 13,01$, $MSE = 5,78$, $p = 0,001$, partiële $\eta^2 = 0,17$. Bij de test met responsmethode 2 hebben rijsschoolleerlingen gemiddeld 3,90 films goed ($SD = 2,35$) en hebben professionals gemiddeld 5,23 films goed ($SD = 2,14$). Dit verschil is statistisch significant en de effectgrootte is gemiddeld, $F(1,58) = 5,26$, $MSE = 5,07$, $p = 0,025$, partiële $\eta^2 = 0,08$.

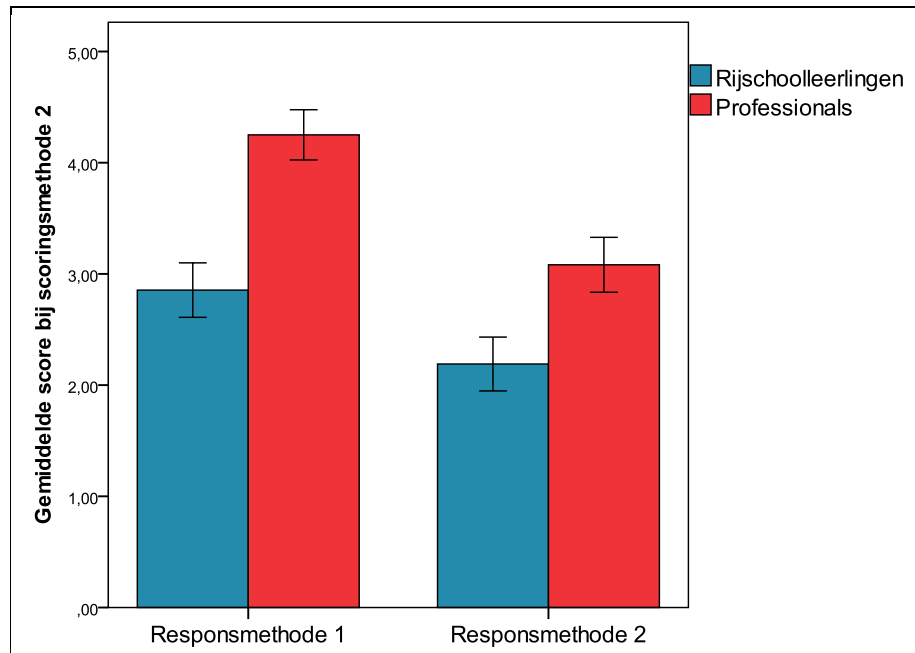
De eindscore van rijsschoolleerlingen is bij de test met responsmethode 2 marginaal significant lager dan bij de test met responsmethode 1 en de effectgrootte is klein, $F(1,58) = 3,52$, $MSE = 5,48$, $p = 0,066$, partiële $\eta^2 = 0,06$. De eindscore van professionals is bij de test met responsmethode 2 significant lager dan bij de test met responsmethode 1 met een grote effectgrootte, $F(1,62) = 11,48$, $MSE = 5,40$, $p = 0,001$, partiële $\eta^2 = 0,16$. Er is geen significant interactie-effect 'responsmethode X groep (rijsschoolleerlingen/professionals)', $F(1,120) = 1,00$, $MSE = 5,44$, $p = 0,319$.

Bij de test met responsmethode 1 hebben mannelijke rijsschoolleerlingen gemiddeld 4,9 films goed ($SD = 2,4$) en hebben vrouwelijke rijsschoolleerlingen gemiddeld 5,2 films goed ($SD = 2,3$). Het verschil tussen mannelijke en vrouwelijke rijsschoolleerlingen is niet significant, $F(1,28) = 0,15$, $MSE = 5,58$, $p = 0,70$. De gemiddelde score en de spreiding rond het gemiddelde is bij de test met responsmethode 2 precies gelijk voor mannelijke en vrouwelijke rijsschoolleerlingen ($M = 3,9$ ($SD = 2,4$)).

4.2. Eindscores responsmethoden 1 en 2 bij scoringsmethode 2

Ook bij de test met scoringsmethode 2 verhogen proefpersonen hun score als zij binnen het tijdvenster het juiste doelgebied aanklikken. De hoogte van de score wordt tevens bepaald door het moment waarop dat gebeurt binnen het tijdvenster. De score op een film is het hoogst (10 punten) als dat moment samenvalt met het meest geschikte moment om tijdig anticiperende maatregelen te nemen (zie *Paragraaf 3.6.2*). De maximale eindscore is 10 punten. *Afbeelding 4.2* geeft de gemiddelde eindscores weer van rijsschoolleerlingen en professionals bij de test met responsmethode 1 en de test met responsmethode 2.

Bij de test met responsmethode 1 scoren rijsschoolleerlingen gemiddeld 2,85 ($SD = 1,34$) en scoren professionals gemiddeld 4,25 ($SD = 1,32$). Dit verschil is statistisch zeer significant en de effectgrootte is groot, $F(1,62) = 17,60$, $MSE = 1,77$, $p = 0,000$, partiële $\eta^2 = 0,22$. Bij de test met responsmethode 2 scoren rijsschoolleerlingen gemiddeld 2,19 ($SD = 1,33$) en scoren professionals gemiddeld 3,08 ($SD = 1,35$). Dit verschil is statistisch significant en de effectgrootte is gemiddeld, $F(1,58) = 6,65$, $MSE = 1,80$, $p = 0,012$, partiële $\eta^2 = 0,10$.



Afbeelding 4.2. Gemiddelde eindscore berekend met scoringsmethode 2 bij responsmethode 1 en responsmethode 2 van rijschoolleerlingen en professionals. De foutenbalken geven +/- 1 maal de standaardfout (SE) weer.

De eindscore van rijschoolleerlingen is bij de test met responsmethode 2 marginaal significant lager dan bij de test met responsmethode 1 en de effectgrootte is gemiddeld, $F(1,58) = 3,72$, $MSE = 1,78$, $p = 0,059$, partiële $\eta^2 = 0,06$. De eindscore van professionals is bij de test met responsmethode 2 significant lager dan bij de test met responsmethode 1 en de effectgrootte is groot, $F(1,62) = 12,23$, $MSE = 1,78$, $p = 0,001$, partiële $\eta^2 = 0,17$. Er is geen significant interactie-effect 'responsmethode X groep (rijschoolleerlingen/professionals)', $F(1,120) = 1,10$, $MSE = 1,96$, $p = 0,296$.

Bij de test met responsmethode 1 scoren mannelijke rijschoolleerlingen gemiddeld 2,7 ($SD = 1,3$) en scoren vrouwelijke rijschoolleerlingen gemiddeld 3,0 ($SD = 1,4$). Het verschil tussen mannelijke en vrouwelijke rijschoolleerlingen is niet significant, $F(1,28) = 0,29$, $MSE = 1,84$, $p = 0,593$. Bij de test met responsmethode 2 scoren mannelijke rijschoolleerlingen gemiddeld 2,1 ($SD = 1,3$) en scoren vrouwelijke rijschoolleerlingen gemiddeld 2,3 ($SD = 1,4$). Ook dit verschil is niet significant, $F(1,28) = 0,10$, $MSE = 1,82$, $p = 0,754$.

4.3. Conclusie eindscores en responsmethoden

De professionals (rijinstructeurs en rijexaminatoren) maakten bij responsmethode 1 de test aanzienlijk beter dan de rijschoolleerlingen. Dit geldt zowel als de eindscores berekend werden door alleen te kijken of proefpersonen binnen het tijdvenster van het potentiële gevaar de juiste plek van dat potentiële gevaar hebben aangewezen (scoringsmethode 1) als wanneer bij de berekening ook het moment binnen dat tijdvenster werd meegeteld (scoringsmethode 2). Hoewel professionals ook significant beter

scoorden dan rijsschoolleerlingen bij de test met responsmethode 2, is dat verschil veel minder robuust dan bij responsmethode 1.

De verschillen in eindscore tussen rijsschoolleerlingen en professionals zijn iets groter bij scoringsmethode 2 dan bij scoringsmethode 1, maar het meenemen van de reactietijden in de eindscore (scoringsmethode 2) lijkt weinig toe te voegen.

Zowel professionals als rijsschoolleerlingen scoorden bij de test met responsmethode 2 lager dan bij de test met responsmethode 1. Professionals lijken nog wat meer moeite te hebben met responsmethode 2 dan rijsschoolleerlingen.

Er is nagenoeg geen significant verschil tussen mannelijke en vrouwelijke rijsschoolleerlingen in de hoogte van de eindscores.

4.4. De scores per film bij gebruik van scoringsmethode 1

Tabel 4.1 geeft per film de percentages rijsschoolleerlingen en professionals weer die de film goed hebben gemaakt bij de test met responsmethode 1 volgens scoringsmethode 1. Wat per film het dominante potentiële gevaar is, staat vermeld in de *Bijlage*.

Film	Type gevaar	Leerling	Professional	p Fisher's exact	Odds ratio (OR)
1	Verborgen	70,0%	76,5%	0,58	1,39
2	Zichtbaar	43,3%	70,6%	0,04*	3,14
3	Verborgen	23,3%	20,6%	1,00	0,85
4	Zichtbaar	63,3%	58,8%	0,80	0,83
5	Verborgen	20,0%	50,0%	0,02*	4,00
6	Zichtbaar	36,7%	67,6%	0,02*	3,61
7	Verborgen	23,3%	44,1%	0,11	2,59
8	Zichtbaar	16,7%	26,5%	0,38	1,80
9	Verborgen	13,3%	44,1%	0,01*	5,13
10	Zichtbaar	66,7%	73,5%	0,59	1,39
11	Verborgen	53,3%	64,7%	0,45	1,60
12	Zichtbaar	63,3%	82,4%	0,10	2,70
13	Verborgen	10,0%	41,2%	0,01*	6,30
* $p < 0,05$					

Tabel 4.1. *Percentages rijsschoolleerlingen en professionals die een film goed hebben gemaakt bij responsmethode 1 volgens scoringsmethode 1.*

In kolom 2 staat aangegeven of het dominante potentiële gevaar zichtbaar of verborgen is. Rijsschoolleerlingen hadden gemiddeld 48,3% van de zichtbare potentiële gevaren goed en 30,5% van de verborgen potentiële gevaren. Dat rijsschoolleerlingen meer moeite hadden met verborgen potentiële gevaren dan met zichtbare potentiële gevaren is significant, $t(29) = 3,62$, $p = 0,001$. Professionals hadden gemiddeld 63,2% van de zichtbare potentiële gevaren

goed en 48,7% van de verborgen potentiële gevaren goed. Ook professionals blijken significant meer moeite te hebben gehad met verborgen potentiële gevaren dan met zichtbare potentiële gevaren, $t(33) = 2,62$, $p = 0,013$.

In de kolommen 3 en 4 staan de percentages die rijsschoolleerlingen en professionals goed hadden. Van de dertien films zijn er elf waarbij het percentage 'goed' van professionals hoger is dan dat van rijsschoolleerlingen. Alleen bij de films 3 en 4 ligt het percentage 'goed' iets hoger bij rijsschoolleerlingen dan bij professionals. In kolom 5 staat de uitslag van de Fisher's exact test. De professionals hebben de films 2, 5, 6, 9 en 13 significant beter gemaakt dan de rijsschoolleerlingen.

Opvallend is het verschil in moeilijkheidsgraad van de films. De rijsschoolleerlingen hebben vooral de films 8, 9 en 13 zeer slecht gemaakt. De professionals hebben de films 3 en 8 naar verhouding slecht gemaakt.

In kolom 6 staan de odds ratio's (OR). Bij film 13 is de OR bijvoorbeeld 6,30. Dit wil zeggen dat de kans dat een professional deze film goed heeft gemaakt, 6,30 keer groter is dan de kans dat een rijsschoolleerling deze film goed heeft gemaakt.

Tabel 4.2 geeft per film de percentages rijsschoolleerlingen en professionals weer die de film goed hebben gemaakt bij de test met responsmethode 2 volgens scoringsmethode 1.

Film	Type gevaar	Leerling	Professional	p Fisher's Exact	Odds ratio (OR)
1	Verborgen	63,3%	60,0%	1,0	0,87
2	Zichtbaar	20,0%	40,0%	0,16	2,67
3	Verborgen	20,0%	20,0%	1,0	1,00
4	Zichtbaar	40,0%	26,7%	0,41	0,55
5	Verborgen	16,7%	23,3%	0,75	1,52
6	Zichtbaar	23,3%	43,3%	0,17	2,51
7	Verborgen	26,7%	36,7%	0,58	1,59
8	Zichtbaar	3,3%	6,7%	1,0	2,07
9	Verborgen	16,7%	30,0%	0,36	2,14
10	Zichtbaar	80,0%	83,3%	1,0	1,25
11	Verborgen	30,0%	70,0%	0,004**	5,44
12	Zichtbaar	40,0%	46,7%	0,80	1,31
13	Verborgen	10,0%	36,7%	0,04*	5,21

* $p < 0,05$; ** $p < 0,01$

Tabel 4.2. Percentages rijsschoolleerlingen en professionals die een film goed hebben gemaakt bij responsmethode 2 volgens scoringsmethode 1.

In tegenstelling tot de test met responsmethode 1 merkten rijsschoolleerlingen bij responsmethode 2 de verborgen gevaren niet significant

slechter op dan de zichtbare potentiële gevaren, $t(29) = 1,45$, $p = 0,159$. Dit geldt ook voor professionals, $t(29) = 0,29$, $p = 0,773$.

Van de dertien films zijn er tien waarbij het percentage 'goed' van de professionals hoger ligt dan het percentage 'goed' van de rijschoolleerlingen. Bij film 3 ligt het percentage van beide groepen precies gelijk. Bij maar twee films ligt het percentage 'goed' van professionals significant hoger. Dit zijn de films 11 en 13. Opvallend is dat de films 2, 5, 6 en 9, die de professionals bij de test met responsmethode 1 significant beter maakten, bij de test met responsmethode 2 in het geheel niet significant beter werden gemaakt door de professionals. Daarentegen maakten professionals film 11 bij de test met responsmethode 1 nauwelijks beter dan de rijschoolleerlingen, terwijl zij deze film bij de test met responsmethode 2 wel significant beter maakten.

Ook bij de test met responsmethode 2 blijken er grote verschillen in moeilijkheidsgraad te zijn tussen de films. Rijschoolleerlingen maakten de films 5, 8 en 13 zeer slecht. De professionals maakten alleen film 8 zeer slecht.

4.5. De scores per film bij gebruik van scoringsmethode 2

Tabel 4.3 geeft de gemiddelde score voor rijschoolleerlingen en professionals weer van de test met responsmethode 1 volgens scoringsmethode 2.

Film	Type gevaar	Leerling	Professional	Mann-Whitney's U	p	Effect-grootte r
1	Verborgen	5,4	5,9	536,5	0,72	0,05
2	Zichtbaar	3,5	5,2	610,5	0,16	0,18
3	Verborgen	1,9	1,5	483,5	0,62	-0,06
4	Zichtbaar	4,7	4,7	535	0,73	0,04
5	Verborgen	1,7	4,2	662	0,02*	0,30
6	Zichtbaar	2,5	5,1	703,5	0,00**	0,34
7	Verborgen	1,4	3,0	641	0,04*	0,26
8	Zichtbaar	1,4	2,4	569,5	0,27	0,14
9	Verborgen	1,1	3,4	665,5	0,01*	0,32
10	Zichtbaar	4,7	5,4	596	0,24	0,15
11	Verborgen	3,7	4,7	586,5	0,29	0,13
12	Zichtbaar	4,4	6,4	692,5	0,01*	0,31
13	Verborgen	0,8	3,2	669	0,00**	0,34

* $p < 0,05$, ** $p < 0,01$

Tabel 4.3. Gemiddelde score van de rijschoolleerlingen en professionals per film bij responsmethode 1 volgens scoringsmethode 2.

Omdat de scores niet normaal verdeeld zijn (zie Paragraaf 3.7) is de Mann-Whitney's U-test gebruikt om na te gaan of de scores op een film tussen rijschoolleerlingen en professionals significant van elkaar verschillen. Omdat gebruik is gemaakt van een non-parametrische test is de effectgrootte niet

uitgedrukt in de partiële η^2 maar in de Pearson's correlatiecoëfficiënt (r) (zie kolom 7).

Rijschoolleerlingen scoorden gemiddeld significant slechter op verborgen potentiële gevaren dan op zichtbare potentiële gevaren, $t(29) = 3,14$, $p = 0,004$. Dit geldt ook voor de professionals, $t(33) = 2,62$, $p = 0,013$.

Van de dertien films zijn er elf waarbij professionals gemiddeld hoger scoorden dan rijsschoolleerlingen. Op film 4 scoorden beide groepen gelijk en bij film 3 ligt de gemiddelde score van rijsschoolleerlingen iets boven die van de professionals. Professionals maakten de films 5, 6, 7, 9, 12 en 13 significant beter dan rijsschoolleerlingen. In tegenstelling tot bij scoringsmethode 1 wordt bij scoringsmethode 2 film 2 niet significant beter gemaakt, maar film 12 juist weer wel. Als per film gekeken wordt, zijn de resultaten volgens scoringsmethode 2 toch enigszins anders dan de resultaten volgens scoringsmethode 1. Rekening houden met de reactietijden binnen het tijdvenster lijkt dus iets uit te maken.

Ook bij scoringsmethode 2 zijn de verschillen per film groot en is dus de ene film aanmerkelijk moeilijker dan de andere. Voor rijsschoolleerlingen waren vooral de films 7, 8, 9 en 13 moeilijk. Voor professionals was film 3 moeilijk.

Tabel 4.4 geeft per film de gemiddelde scores van rijsschoolleerlingen en professionals weer van de test met responsmethode 2 volgens scoringsmethode 2.

Film	Type gevaar	Leerling	Professional	Mann-Whitney's U	p	Effect-grootte r
1	Verborgen	5,3	5,0	426,5	0,72	-0,05
2	Zichtbaar	1,5	2,8	534	0,13	0,20
3	Verborgen	1,5	1,5	448,5	0,98	-0,04
4	Zichtbaar	2,6	2,2	420	0,60	-0,07
5	Verborgen	1,2	1,8	485	0,46	0,10
6	Zichtbaar	1,5	3,2	556,5	0,06	0,24
7	Verborgen	1,6	2,5	515	0,24	0,15
8	Zichtbaar	0,3	0,5	464	0,58	0,07
9	Verborgen	1,1	2,2	515,5	0,19	0,17
10	Zichtbaar	6,0	6,9	566,5	0,08	0,22
11	Verborgen	2,3	4,9	613,5	0,01*	0,33
12	Zichtbaar	2,9	3,8	511	0,32	0,13
13	Verborgen	0,8	3,1	572,5	0,02*	0,32

* $p < 0,05$

Tabel 4.4. Gemiddelde score van de rijsschoolleerlingen en professionals per film bij responsmethode 2 volgens scoringsmethode 2.

In tegenstelling tot de test met responsmethode 1 hadden zowel rijsschoolleerlingen ($t(29) = 1,15$, $p = 0,262$) als professionals ($t(29) = 0,57$, $p = 0,572$) bij de test met responsmethode 2 en gebruikmaking van scoringsmethode 2

niet significant meer moeite met het detecteren van verborgen potentiële gevaren dan van zichtbare potentiële gevaren.

Van de dertien films zijn er tien waarbij professionals gemiddeld hoger scoorden dan rijsschoolleerlingen. Op film 3 scoorden beide groepen gelijk en bij de films 1 en 4 liggen de gemiddelde score van rijsschoolleerlingen iets boven die van professionals. Professionals maakten alleen de films 11 en 13 significant beter dan rijsschoolleerlingen.

Film 8 werd opvallend slecht gemaakt door zowel rijsschoolleerlingen als professionals. De films 3 en 5 gaven ook tamelijk veel problemen voor zowel rijsschoolleerlingen als professionals. Verder maakten rijsschoolleerlingen de films 9 en 13 slecht.

4.6. Conclusies van de scores per film

Bij de test met responsmethode 1 zijn er aanzienlijk meer films die professionals significant beter hebben gemaakt dan bij de test met responsmethode 2. Dit geldt zowel voor scoringsmethode 1 als scoringsmethode 2. Toch laat zelfs bij de test met responsmethode 1 iets minder dan de helft van de films bij beide scoringsmethoden een significant verschil zien tussen rijsschoolleerlingen en professionals. Er is zelfs een enkele film die beter, maar niet significant beter, werd gemaakt door de rijsschoolleerlingen.

De moeilijkheidsgraad fluctueert tamelijk sterk van film tot film. Vooral enkele films bij de test met responsmethode 2 laten een uitzonderlijk lage score zien. Dit geldt zowel voor rijsschoolleerlingen als voor professionals.

4.7. De vragenlijsten

Op één variabele na blijken de achtergrondvariabelen uit de vragenlijst van rijsschoolleerlingen geen significante invloed te hebben gehad op de hoogte van de eindscore. Dit geldt zowel voor de test met responsmethode 1 als voor de test met responsmethode 2 en voor beide scoringsmethoden. De variabelen die geen significante invloed blijken te hebben, zijn leeftijd, opleidingsniveau, aantal uren rijinstructie en hoe vaak proefpersonen een computer gebruiken. Dat het aantal uren rijinstructie geen invloed lijkt te hebben, doet op het eerste gezicht vreemd aan. Het is te verwachten dat hoe meer uren rijles proefpersonen hebben gehad, hoe hoger ze zullen scoren. Toch is dit minder vreemd dan het lijkt. Onder alle deelnemende rijsschoolleerlingen zijn er namelijk vrijwel geen die minder dan vijf klokuren rijles hebben gehad. Bij de test met responsmethode 1 hebben twee van de dertig rijsschoolleerlingen minder dan vijf klokuren rijles gehad en bij de test met responsmethode 2 geldt dit voor één van de dertig rijsschoolleerlingen. Daar staat tegenover dat zowel in de steekproef van responsmethode 1 als in de steekproef van responsmethode 2 vijf rijsschoolleerlingen zitten met meer dan vijftig klokuren rijles. In beide steekproeven komt zelfs een rijsschoolleerling voor met meer dan honderd klokuren rijles. Het kan zijn dat deze proefpersonen zo veel uren rijles hebben gehad, juist omdat de rijinstructeur vond dat ze slecht waren in gevaarherkenning en daarom nog niet klaar waren om rijexamen te doen. In een longitudinaal onderzoek uit het Verenigd Koninkrijk is gebleken dat automobilisten die pas na zeer veel rijlessen waren geslaagd, een hoog ongevalsrisico hadden na het behalen van het rijbewijs (Maycock & Forsyth, 1997).

De enige achtergrondvariabele die wel invloed heeft op de eindscore, is de mate waarin rijsschoolleerlingen computergames spelen, althans bij de test met responsmethode 2. In de vragenlijst konden proefpersonen aankruisen of ze (1) 'dagelijks', (2) 'twee tot drie keer per week', (3) 'gemiddeld niet meer dan één keer per week', (4) 'enkele keren per jaar' of (5) 'nooit' games speelden op de computer of smartphone. De categorieën 1 en 2 zijn samengevoegd tot de groep van zware gamers en de categorieën 3, 4 en 5 zijn samengevoegd tot de groep van lichte gamers. In de steekproef die de test met responsmethode 1 had gemaakt, zaten zestien zware gamers en dertien lichte gamers (één proefpersoon had deze vraag niet beantwoord). In de steekproef van rijsschoolleerlingen die de test met responsmethode 2 hadden gemaakt, zaten veertien zware gamers en zestien lichte gamers. Het verschil tussen zware en lichte gamers op de eindscore berekend met scoringsmethode 1 bij de test met responsmethode 1, is niet significant, $F(1,27) = 0,48$, $MSE = 5,71$, $p = 0,493$. Het verschil is ook niet significant bij gebruik van scoringsmethode 2, $F(1,27) = 0,240$, $MSE = 1,91$, $p = 0,622$. Bij de test met responsmethode 2 maakten zware gamers de test significant beter dan lichte gamers. Dit geldt zowel bij gebruik van scoringsmethode 1 ($F(1,28) = 5,85$, $MSE = 4,75$, $p = 0,022$, partiële $\eta^2 = 0,17$) als bij gebruik van scoringsmethode 2 ($F(1,28) = 6,41$, $MSE = 1,48$, $p = 0,017$, partiële $\eta^2 = 0,19$). Men heeft dus baat van ervaring met games bij het maken van de test met responsmethode 2, maar niet bij het maken van de test met responsmethode 1.

In de vragenlijst is ook gevraagd wat proefpersonen vonden van de introductie en of ze de test geschikt achtten om gevaarherkenning te meten op het theorie-examen. *Tabel 4.5* geeft de antwoorden in percentages weer op de vraag over de hoeveelheid informatie in de introductie.

Respons- methode	Groep	Om goed te kunnen onthouden vond ik de hoeveelheid informatie in de introductie...				
		... veel te veel	... te veel	... precies goed	... te weinig	... veel te weinig
Respons- methode 1	Rijschoolleerlingen	0	24,1	75,9	0	0
	Professionals	3,0	18,2	78,8	0	0
Respons- methode 2	Rijschoolleerlingen	0	43,3	53,3	3,3	0
	Professionals	8,3	37,5	54,2	0	0

Tabel 4.5. Antwoorden (in percentages) op de vraag naar de hoeveelheid informatie in de introductie.

Hoewel de meeste proefpersonen de hoeveelheid informatie precies goed vonden, vond een substantieel deel dat er te veel informatie werd geboden om te kunnen onthouden.

Tabel 4.6 geeft de antwoorden in percentages weer op de vraag of de test geschikt is voor in het theorie-examen.

Respons- methode	Groep	Hoe goed kan volgens u met deze test gevaarherkenning op het theorie-examen gemeten worden?				
		Heel goed	Goed	Niet goed en niet slecht	Slecht	Heel slecht
Respons- methode 1	Rijschoolleerlingen	17,6	73,5	8,8	0	0
	Professionals	17,6	73,5	8,8	0	0
Respons- methode 2	Rijschoolleerlingen	13,3	60,0	23,3	3,3	0
	Professionals	20,8	62,5	16,7	0	0

Tabel 4.6. Antwoorden (in percentages) op de vraag over de geschiktheid van de test in het theorie-examen.

Veruit de meeste rijsschoolleerlingen en professionals vonden de test die zij gedaan hadden (met responsmethode 1 of 2) een geschikte methode voor in het theorie-examen.

4.8. Het markeren van gevaren bij responsmethode 2

Bij responsmethode 2 kan de proefpersoon tijdens het zien van de film tot viermaal toe op de spatiebalk drukken om een potentieel gevaar te markeren. Na afloop van de film kiest hij dan van de keren dat hij op de spatiebalk heeft gedrukt, het moment met het grootste potentiële gevaar uit. Met de momenten die hij niet als het grootste potentiële gevaar aanmerkt, wordt bij zowel scoringsmethode 1 als scoringsmethode 2 niets gedaan, omdat anders geen zuivere vergelijking gemaakt kan worden met responsmethode 1.

Rijschoolleerlingen drukten per film gemiddeld 1,8 keer op de spatiebalk ($SD = 0,6$). Professionals drukten per film gemiddeld 1,7 keer op de spatiebalk ($SD = 0,6$). Tabel 4.7 geeft weer hoeveel procent van de rijsschoolleerlingen en professionals bij het zien van een film respectievelijk 0, 1, 2, 3 of 4 maal op de spatiebalk drukten om een potentieel gevaar te markeren.

Groep	0x	1x	2x	3x	4x
Rijschoolleerlingen	11,5	26,6	34,3	25,3	2,2
Professionals	12,4	28,1	35,5	21,4	2,7

Tabel 4.7. Percentage van de rijsschoolleerlingen en professionals dat bij responsmethode 2 respectievelijk 0, 1, 2, 3 of 4 maal per film op de spatiebalk heeft gedrukt om een potentieel gevaar te markeren.

De verschillen tussen beginners en professionals zijn niet significant. De overwegend lage scores bij responsmethode 2 worden dus deels veroorzaakt doordat proefpersonen bij het zien van een film in het geheel niet op de spatiebalk hebben gedrukt om een potentieel gevaar te markeren.

4.9. De interne consistentie van de tests

In *Tabel 4.8* staan de waarden van de interne consistentie (Cronbach's α) van de test met responsmethode 1 en de test met responsmethode 2 bij berekening van de scores met zowel scoringsmethode 1 als scoringsmethode 2.

Responsmethode	Scoringsmethode	Cronbach's α
Responsmethode 1	Scoringsmethode 1	0,68
	Scoringsmethode 2	0,62
Responsmethode 2	Scoringsmethode 1	0,58
	Scoringsmethode 2	0,61

Tabel 4.8. De interne consistentie van de toetsitems bij de test met responsmethoden 1 en 2 en scoringsmethoden 1 en 2.

In alle gevallen is de interne consistentie te laag om van een duidelijke schaal te kunnen spreken. De α is zeker te laag wanneer men bedenkt dat het uiteindelijke doel is om individuele kandidaten te beoordelen op hun vermogen gevaren te herkennen in het verkeer. Het doel van het onderhavig onderzoek was niet om beide tests psychometrisch te toetsen. Daarvoor zijn de twee steekproeven veel te klein. Bij een kleine steekproef kan door één afwijkende proefpersoon de α al snel veel van waarde veranderen. De α kan verhoogd worden door meer toetsitems (films) in de test op te nemen en/of door ervoor te zorgen dat alle toetsitems ongeveer even zwaar zijn. Aangezien het maken van de test bij responsmethode 1 al ongeveer een kwartier in beslag neemt, is het niet wenselijk dat er meer toetsitems komen. De toetsitems verschillen onderling sterk in zwaarte en op de meeste toetsitems scoorden proefpersonen laag (zie *Paragraaf 4.4*). De α zou verbeterd kunnen worden door in een aantal films het dominante potentiële gevaar prominenter te maken en de nevengevaren juist wat minder prominent. Dit kan al bereikt worden met enkele subtiele aanpassingen.

5. Conclusies en aanbevelingen

5.1. Conclusies

In dit rapport is het onderzoek besproken naar het gebruik van bewegende beelden om gevaarherkenning in het verkeer te toetsen. Het ging erom een valide, betrouwbare en fraudebestendige toetsmethode te vinden die geschikt is om onderdeel te worden van het toekomstige theorie-examen. De gevaarherkenningstest met foto's die momenteel deel uitmaakt van het theorie-examen, voldoet om verschillende redenen niet. Het is bij deze test niet duidelijk wat precies getest wordt. Op foto's is niet te zien met welke snelheid andere verkeersdeelnemers zich voortbewegen. Kandidaten kunnen zakken voor de fototest wanneer ze (te) voorzichtig zijn. En door het aanleren van een ezelsbruggetje dat nauwelijks iets te maken heeft met gevaarherkenning, kunnen kandidaten gemakkelijk slagen voor dit onderdeel.

Het doel van het onderhavige onderzoek was om een responsmethode te vinden voor een gevaarherkenningstoets met bewegende beelden. Die toets zou een duidelijk verschil in eindscores moeten geven tussen rijsschoolleerlingen en professionals (rijinstructeurs en rijexaminatoren). Duidelijk wil in dit geval zeggen: statistisch significant met een grote effectgrootte. De toets zou ook een responsmethode moeten hebben die waarschijnlijk minder fraudegevoelig is dan de gevaarherkenningstest met bewegende beelden die sinds 2002 in het Verenigd Koninkrijk gebruikt wordt. Ten slotte zou het in aanzet mogelijk moeten zijn om een uiteindelijke test te ontwikkelen met een hoge interne consistentie.

Op basis van de literatuur en de ervaringen uit het eerdere onderzoek in Nederland zijn twee responsmethoden ontwikkeld. Bij responsmethode 1 krijgt de proefpersoon telkens korte animatiefilms te zien met in elke film een dominant potentieel gevaar en enkele minder belangrijke nevengevaren. Direct na afloop van een film moet hij voor zichzelf vaststellen wat het grootste potentiële gevaar in die film was. Vervolgens start hij de film voor de tweede keer en stopt die dan door op de spatiebalk te drukken op het moment dat hij in gedachten genomen heeft. Daarna wijst hij op het stilstaande beeld het potentiële gevaar aan door er met de muis op te klikken. Bij responsmethode 2 krijgt de proefpersoon dezelfde animatiefilms te zien. Hij krijgt elk filmpje nu echter maar één keer te zien. Als hij tijdens het zien van een film voelt dat het gevaarlijk zou kunnen worden, drukt hij op de spatiebalk. Hij hoort een toon en er komt op de tijdbalk onder in beeld een witte markering te staan op het tijdstip waarop hij heeft gedrukt. De film blijft echter gewoon doorlopen. Men kan maximaal vier keer per film op de spatiebalk drukken. Direct na afloop van elke film krijgt de proefpersoon de stilstaande beelden te zien van de momenten waarop hij heeft gedrukt. Van die momenten dient hij het beeld uit te kiezen waarop het grootste potentiële gevaar te zien is. Vervolgens dient hij op dat beeld, dat groot op het scherm komt te staan, met de muis dat potentiële gevaar aan te klikken.

Doordat de gevaren zich tijdens de films niet tot weinig ontwikkelen, zullen de test met responsmethode 1 en de test met responsmethode 2 waarschijnlijk minder fraudegevoelig zijn dan de methode die in het Verenigd

Koninkrijk gebruikt wordt. Daarnaast is het bij die laatste methode niet goed mogelijk om met verborgen potentiële gevaren te werken en bij de twee nieuw ontwikkelde responsmethoden wel. Juist in het herkennen van verborgen potentiële gevaren blijken beginnende automobilisten slecht te zijn (Crundall et al., 2012; Vlakveld, 2011).

Er zijn twee scoringsmethoden ontwikkeld die voor beide responsmethoden hetzelfde waren. Bij scoringsmethode 1 was een score op een film goed (1 punt) als binnen het tijdvenster (de tijdsspanne waarin een gevaar te zien is) van het dominante potentiële gevaar het juiste potentiële gevaar (het doelgebied) was aangewezen. In alle andere gevallen kreeg de proefpersoon 0 punten. Bij scoringsmethode 2 werd de hoogte van de score niet alleen bepaald door het aanklikken van een doelgebied, maar ook door het moment binnen het tijdvenster dat de proefpersoon had gekozen. Hoe dichter dat moment lag bij het ideale moment om te anticiperen, hoe hoger de score was.

Bij beide responsmethoden scoorden professionals significant beter dan rijsschoolleerlingen. Het verschil was echter aanmerkelijk groter en robuuster bij responsmethode 1 dan bij responsmethode 2. Wanneer scoringsmethode 2 werd gebruikt, waren de verschillen tussen beide groepen iets groter dan wanneer scoringsmethode 1 werd gebruikt.

Als naar de scores per film gekeken wordt, zijn er veel meer films bij de test met responsmethode 1 die een significant verschil tussen beide groepen laten zien dan bij de test met responsmethode 2. Opvallend is wel dat de moeilijkheidsgraad per film sterk fluctueert en dat op een aantal films zeer slecht gescoord werd. Waarschijnlijk is daardoor de interne consistentie (de Cronbach's α) aan de lage kant.

Rijsschoolleerlingen die veelvuldig computergames spelen, maakten de test met responsmethode 2 significant beter dan rijsschoolleerlingen die weinig tot nooit computergames spelen. De mate waarin rijsschoolleerlingen computergames spelen had geen invloed op de hoogte van de eindscores bij de test met responsmethode 1.

De lage eindscores bij de test met responsmethode 2 kunnen deels verklaard worden uit het feit dat rijsschoolleerlingen in 11,5% van de gevallen bij het zien van een film in het geheel niet op de spatiebalk drukten om een potentieel gevaar te markeren. Bij de professionals was dit 12,4%. Het zou kunnen dat naast het vermogen gevaren te herkennen, spelstrategie invloed heeft gehad op de hoogte van de score bij responsmethode 2. Wellicht zijn mensen bijvoorbeeld minder geneigd om aan het begin van een film al op de spatiebalk te drukken, omdat ze bang zijn de vier mogelijkheden te snel te verspelen. Aan het einde van de film komen ze er dan achter dat ze nog mogelijkheden om te drukken over hebben. Een ander verschil is dat de proefpersoon bij de test met responsmethode 1, nadat hij de film in zijn geheel heeft gezien, kan overdenken wat nu het grootste potentiële gevaar was. Bij responsmethode 2 moet men echter direct op de spatiebalk drukken wanneer het potentiële gevaar zich voordoet. In film 3 bijvoorbeeld loopt een vrouw met winkelwagen over een voetgangersoversteekplaats. Uiteindelijk is dit geen gevaar, omdat de camera-auto vaart vermindert. Achteraf weet de proefpersoon dat, maar als hij de film voor de eerste keer ziet, nog niet. Die overstekende vrouw lijkt op dat moment een groot gevaar.

Het grotere en robuustere verschil tussen professionals en rijsschoolleerlingen bij de test met responsmethode 1, het grotere aantal films dat bij de test met responsmethode 1 significant beter wordt gemaakt door professionals dan door rijsschoolleerlingen en de ongevoeligheid van de test met responsmethode 1 voor ervaring met computergames, maken de test met responsmethode 1 geschikter om door te ontwikkelen tot een test op het theorie-examen dan de test met responsmethode 2.

5.2. Aanbevelingen

Met het onderhavige onderzoek is de eerste stap gezet op weg naar een gevaarherkenningsstoets met bewegende beelden voor het theorie-examen. Er moeten echter nog veel stappen gezet worden voordat de nieuwe test geïmplementeerd kan worden. De SWOV beveelt het CBR aan de volgende stappen te zetten:

- Op basis van de uitkomsten van dit onderzoek wordt aanbevolen door te gaan met de verdere ontwikkeling van de test met responsmethode 1 en te stoppen met de verdere ontwikkeling van de test met responsmethode 2.
- Als eerste zal aangetoond moeten worden dat een test met voldoende interne consistentie (Cronbach's α) ontwikkeld kan worden, waardoor het legitiem wordt om individuele kandidaten te beoordelen. Daartoe zullen enkele films enigszins aangepast moeten worden en dient een grote steekproef rijsschoolleerlingen de test met responsmethode 1 te maken.
- De software die is gebruikt om de eindscores te berekenen, is ontwikkeld binnen het softwarepakket SPSS. Dit pakket wordt niet gebruikt door het CBR. De ontwikkelde software zal dus moeten worden omgezet naar een pakket dat het CBR wel gebruikt.
- Dit onderzoek heeft alleen nog maar aangetoond dat professionals als groep de test met responsmethode 1 duidelijk beter maken dan de groep rijsschoolleerlingen. Het is nog niet aangetoond dat rijsschoolleerlingen zich door training, waarbij het niet louter gaat om training in het doen van de test, kunnen verbeteren. Het is belangrijk om dit aan te tonen, omdat het rijexamen een rijvaardigheidstest is. In tegenstelling tot rijgeschiktheid wordt aangenomen dat rijvaardigheid aan te leren is. Het is al wel aangetoond dat gevaarherkenning te trainen is (o.a. Vlakveld, 2011), maar er is nog geen onderzoek verricht waaruit blijkt dat men hogere scores op de test met responsmethode 1 weet te behalen door een generieke gevaarherkenningstraining. Onderzoek hiernaar wordt aanbevolen.
- Om fraude tegen te gaan en te voorkomen dat kandidaten die voor het onderdeel gevaarherkenning zijn gezakt, bij de tweede keer niet dezelfde test hoeven te maken, zullen meerdere tests van gelijke zwaarte met ongeveer dezelfde interne consistentie ontwikkeld moeten worden.
- Van die tests dient vervolgens de cesuur bepaald te worden. Bij welke score is een kandidaat nog juist geslaagd? Hiervoor is het waarschijnlijk ook noodzakelijk dat naast een grote steekproef rijsschoolleerlingen een grote steekproef professionals de tests maakt.

Literatuur

- Borowsky, A., Oron-Gilad, T. & Parmet, Y. (2009). *Age and skill differences in classifying hazardous traffic scenes*. In: *Transportation Research Part F*, vol. 12, p. 277–287.
- CBR (2005). *Van goed naar beter; Voorstel scenario vernieuwd rijexamen*. Centraal Bureau Rijvaardigheidsbewijzen, Rijswijk.
- Chapman, P. & Underwood, G. (1998). *Visual search of driving situations; Danger and experience*. In: *Perception*, vol. 27, nr. 8, p. 951-964.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd ed. Academic Press, New York.
- Congdon, P. (1999). *VicRoads hazard perception test: can it predict accidents?* CR 99-1. Australian Council for Educational Research, Camberwell.
- Craen, S. de (2010). *The X-factor; A longitudinal study of calibration in young novice drivers*. SWOV-Dissertatiereeks. SWOV, Leidschendam.
- Crone, E.A. & Dahl, R.E. (2012). *Understanding adolescence as a period of social-affective engagement and goal flexibility*. In: *Nature Reviews Neuroscience*, vol. 13, nr. 9, p. 636-650.
- Crundall, D., Chapman, P., Trawley, S., Collins, L., et al. (2012). *Some hazards are more attractive than others; Drivers of varying experience respond differently to different types of hazard*. In: *Accident Analysis & Prevention*, vol. 45, p. 600-609.
- Crundall, D., Underwood, G. & Chapman, P. (2002). *Attending to the peripheral world while driving*. In: *Applied Cognitive Psychology*, vol. 16, nr. 4, p. 459-475.
- Curry, A.E., Hafetz, J., Kallan, M.J., Winston, F.K., et al. (2011). *Prevalence of teen driver errors leading to serious motor vehicle crashes*. In: *Accident Analysis & Prevention*, vol. 43, nr. 4, p. 1285-1290.
- Darby, P., Murray, W. & Raeside, R. (2009). *Applying online fleet driver assessment to help identify, target and reduce occupational road safety risks*. In: *Safety Science*, vol. 47, nr. 3, p. 436-442.
- Elvik, R. (2002). *The importance of confounding in observational before-and-after studies of road safety measures*. In: *Accident Analysis & Prevention*, vol. 34, nr. 5, p. 631-635.
- Fisher, D.L., Pollatsek, A.P. & Pradhan, A. (2006). *Can novice drivers be trained to scan for information that will reduce their likelihood of a crash?* In: *Injury Prevention*, vol. 12, Suppl 1, p. i25–i29.

Grayson, G.B., Maycock, G., Groeger, J.A., Hammond, A.M., et al. (2003). *Risk, hazard perception and perceived control*. TRL560. Transport Research Laboratory (TRL), Wokingham, Berkshire, UK.

Grayson, G.B. & Sexton, B.F. (2002). *The development of hazard perception testing*. TRL558. Transport Research Laboratory (TRL), Wokingham, Berkshire, UK.

Horswill, M.S. & McKenna, F.P. (2004). *Drivers' hazard perception ability; Situation awareness on the road*. In: Banbury, S. & Tremblay, S. (red.), *A cognitive approach to situation awareness*. Ashgate, Aldershot, UK, p. 155-175.

Jackson, L., Chapman, P. & Crundall, D. (2009). *What happens next? Predicting other road users' behaviour as a function of driving experience and processing time*. In: *Ergonomics*, vol. 52, nr. 2, p. 154-164.

Maycock, G. & Forsyth, E. (1997). *Cohort study of learner and novice drivers Part 4; Novice driver accidents in relation to methods of learning to drive, performance in the driving test and self assessed driving ability behaviour*. TRL Report 275. Transport Research Laboratory TRL, Crowthorne, Berkshire, UK.

McGowan, A.M. & Banbury, S.P. (2004). *Evaluating interruption-based techniques using embedded measures of driver anticipation*. In: Banbury, S.P. & Tremblay, S. (red.), *A cognitive approach to situation awareness*. Ashgate, Aldershot, UK, p. 176-192.

McKenna, F.P. & Crick, J. (1997). *Developments in hazard perception; Prepared for road safety division of DETR*. TRL Report 297. Transport Research Laboratory (TRL), Crowthorne, Berkshire, UK.

McKenna, F.P. & Horswill, M.S. (1999). *Hazard perception and its relevance for driver licensing*. In: *IATSS Research*, vol. 23, nr. 1, p. 36-41.

McKenna, F.P., Horswill, M.S. & Alexander, J.L. (2006). *Does anticipation training affect drivers' risk taking?* In: *Journal of Experimental Psychology: Applied*, vol. 12, nr. 1, p. 1-10.

McKnight, A.J. & McKnight, A.S. (2003). *Young novice drivers: careless or clueless?* In: *Accident Analysis & Prevention*, vol. 35, nr. 6, p. 921-925.

Mourant, R.R. & Rockwell, T.H. (1972). *Strategies of visual search by novice and experienced drivers*. In: *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 14, p. 325-335.

NZ Transport Agency (2012). *Full licence test guide*. Geraadpleegd 25-12-2012 op <http://www.nzta.govt.nz/licence/photo/docs/full-test-guide.pdf>.

OECD (2006). *Young drivers: the road to safety*. (ITRD E130375). Organisation for Economic Co-operation and Development OECD, Paris

- Pelz, D.C. & Krupat, E. (1974). *Caution profile and driving record of undergraduate males*. In: Accident Analysis & Prevention, vol. 6, nr. 1, p. 45-58.
- Pollatsek, A., Narayanaan, V., Pradhan, A. & Fisher, D.L. (2006). *Using eye movements to evaluate a PC-based risk awareness and perception training program on a driving simulator*. In: Human Factors, vol. 48, nr. 3, p. 447-464.
- Pollatsek, A., Vlakveld, W.P., Kappé, B., Pradhan, A.K., et al. (2011). *Driving simulators as training and evaluation tools: novice drivers*. In: Fisher, D.L., Rizzo, M., Caird, J.K. & Lee, J.D. (red.), Handbook of driving simulation for engineers, medicine, and psychology. CRC Press, Taylor & Francis Group, Boca Raton, USA, p. 30-31-30-18.
- Pradhan, A.K., Hammel, K.R., DeRamus, R., Pollatsek, A., et al. (2005). *Using eye movements to evaluate effects of driver age on risk perception in a driving simulator*. In: Human Factors, vol. 47, nr. 4, p. 840–852.
- Pradhan, A.K., Pollatsek, A., Knodler, M. & Fisher, D.L. (2009). *Can younger drivers be trained to scan for information that will reduce their risk in roadway traffic scenarios that are hard to identify as hazardous?* In: Ergonomics, vol. 52, nr. 6, p. 657-673.
- Quimby, A.R. & Watts, G.R. (1981). *Human factors and driving performance*. TRRL Laboratory Report 1004. Transport and Road Research Laboratory (TRRL), Crowthorne, Berkshire, UK.
- Räsänen, M. & Summala, H. (1998). *Attention and expectation problems in bicycle-car collisions: an in-depth study*. In: Accident Analysis & Prevention, vol. 30, nr. 5, p. 657-666.
- Regan, M.A., Triggs, J.T. & Godley, S.T. (2000). *Simulator-based evaluation of the DriveSmart novice driver CD-Rom training product*. Paper gepresenteerd op Road Safety Research, Policing and Education Conference, 2000, Brisbane, Queensland, Australia.
- Reurings, M.C.B., Vlakveld, W.P., Twisk, D.A.M., Dijkstra, A., et al. (2012). *Van fietsongeval naar maatregelen: kennis en hiaten*. R-2012-8. SWOV, Leidschendam.
- Sagberg, F. & Bjørnskau, T. (2006). *Hazard perception and driving experience among novice drivers*. In: Accident Analysis & Prevention, vol. 38, nr. 2, p. 407-414.
- Schepers, J.P., Kroeze, P.A., Smeets, W. & Wüst, J.C. (2011). *Road factors and bicycle-motor vehicle crashes at unsignalized priority intersections*. In: Accident Analysis & Prevention, vol. 43, nr. 3, p. 853-861.
- Scialfa, C.T., Deschênes, M.C., Fencs, J., Boone, J., et al. (2011). *A hazard perception test for novice drivers*. In: Accident Analysis & Prevention, vol. 43, nr. 1, p. 204-208.
- Sexton, B.F. (2000). *Development of hazard perception testing*. In: Novice drivers conference. 1 and 2 June, Bristol, UK.

Soliday, S.M. (1974). *Relationship between age and hazard perception in automobile drivers*. In: *Perceptual and Motor Skills*, vol. 39, nr. 1, Part 1, p. 335-338.

Stoner, A.S., Fisher, D.L. & Mollenhauer, M. (2011). *Simulator and scenario factors influencing simulator sickness*. In: Fisher, D.L., Rizzo, M., Caird, J.K. & Lee, J.D. (red.), *Handbook of driving simulation for engineering, medicine, and psychology*. CRC Press, Boca Raton, FL, p. 14-11 - 14-24.

Summala, H., Pasanen, E., Räsänen, M. & Sievänen, J. (1996). *Bicycle accidents and drivers' visual search at left and right turns*. In: *Accident Analysis & Prevention*, vol. 28, nr. 2, p. 147-153.

SWOV (2012). *Jonge beginnende automobilisten*. SWOV-Factsheet, november 2012. SWOV, Leidschendam.

Vlakveld, W.P. (2008). *Toetsen en trainen van gevaarherkenning*. D-2008-02. SWOV, Leidschendam.

Vlakveld, W.P. (2011). *Hazard anticipation of young novice drivers*. SWOV-Dissertatiereeks. SWOV, Leidschendam.

Vlakveld, W.P., Romoser, M., Mehranian, H., Diete, F., et al. (2011). *Do crashes and near crashes in simulator-based training enhance novice drivers' visual search for latent hazards?* In: *Transportation Research Record*, vol. 2265, p. 154-160.

Wallis, T.S.A. & Horswill, M.S. (2007). *Using fuzzy signal detection theory to determine why experienced and trained drivers respond faster than novices in a hazard perception test*. In: *Accident Analysis & Prevention*, vol. 39, nr. 6, p. 1177-1185.

Wang, Y., Zhang, W. & Salvendy, G. (2010). *Effects of a simulation-based training intervention on novice drivers' hazard handling performance*. In: *Traffic Injury Prevention*, vol. 11, nr. 1, p. 16-24.

Wells, P., Tong, S., Sexton, B.F., Grayson, G.B., et al. (2008). *Cohort II: A study of learner and new drivers; Volume 1 – main report*. Department for Transport, London.

Wetton, M.A., Horswill, M.S., Hatherly, C., Wood, J.M., et al. (2010). *The development and validation of two complementary measures of drivers' hazard perception ability*. In: *Accident Analysis & Prevention*, vol. 42, nr. 4, p. 1232-1239.

Per film is aangegeven wat het dominante potentiële gevaar is. Potentiële gevaren kunnen verborgen of zichtbaar zijn.

- **Verborgen potentieel gevaar**

Bij verborgen potentiële gevaren gaat het om andere verkeersdeelnemers die er zouden kunnen zijn, maar die de bestuurder niet kan zien doordat het zicht op die mogelijke verkeersdeelnemers is afgeschermd. Proefpersonen hebben een verborgen potentieel gevaar herkend als ze een film hebben gestopt op een moment dat die mogelijke verkeersdeelnemer ergens achter vandaan kan komen en er nog voldoende tijd is om door acties een ongeval te voorkomen. Naast het juiste moment moet de proefpersoon ook de goede richting aangeven waar de mogelijke verkeersdeelnemer voor het eerst te zien is als hij erachter vandaan zou komen. Die richting is in de afbeeldingen van deze bijlage met een witte pijl aangegeven.

- **Zichtbaar potentieel gevaar**

Zichtbare potentiële gevaren zijn andere verkeersdeelnemers die te zien zijn en die, gelet op de verkeerssituatie, zich onveilig zouden kunnen gaan gedragen. Proefpersonen hebben een zichtbaar potentieel gevaar herkend als ze de film hebben gestopt op het moment dat die verkeersdeelnemer nog niet een acuut gevaar vormt en ze de verkeersdeelnemer hebben aangewezen. De verkeersdeelnemers die zich onveilig zouden kunnen gaan gedragen worden op de onderstaande afbeeldingen aangewezen door een witte pijl.

In deze bijlage wordt het perspectief van waaruit gefilmd is aangeduid met 'camera-auto'.

Film 1 (verborgen potentieel gevaar)



De camera-auto rijdt rechtdoor en nadert een T-kruising. De camera-auto heeft voorrang, maar de bestuurder kan niet zien of er verkeer vanuit de zijweg nadert, omdat het zicht erop wordt afgeschermd door de twee gele containers. Eerder in de film is een busje vanuit de zijweg rechtsaf geslagen. Dit busje in de verte is op te vatten als een vooraankondiging van het mogelijke gevaar.

Film 2 (zichtbaar potentieel gevaar)



De camera-auto nadert een kruispunt en wil rechtdoor. Voor hem rijdt een auto die rechtsaf wil slaan. Het verkeerslicht staat op groen. Als de bestuurder van de auto die rechtsaf wil slaan de vrouw die over zou kunnen steken, niet tijdig heeft gezien, moet die auto hard remmen. Wanneer de camera-auto niet voldoende afstand houdt zal de camera-auto in dat geval op de voorligger botsen. Het gevaar is dus de voorligger. In dit geval is ook goed gerekend als de vrouw was aangewezen.

Film 3 (verborgen potentieel gevaar)



De camera-auto nadert een voetgangersoversteekplaats. Vlak voor het zebrapad staat een blauw bestelbusje stil. De vrouw met de boodschappenwagen loopt zo snel dat ze zal zijn overgestoken op het moment dat de camera-auto bij het zebrapad is. Het gevaar zijn voetgangers die na de vrouw oversteken (bijvoorbeeld een kind dat achter de vrouw aanloopt) die de bestuurder van de camera-auto niet kan zien door het geparkeerde busje.

Film 4 (zichtbaar potentieel gevaar)



De camera-auto voegt uit. Links naast hem rijdt een bus. De bus rijdt langzamer dan de camera-auto. Een groen bestelbusje heeft zojuist de bus ingehaald en geeft richting aan. Het zou kunnen dat het bestelbusje nog op het allerlaatst uitvoegt waardoor de camera-auto wordt gesneden. Het busje gaat uiteindelijk alleen maar één rijstrook naar rechts.

Film 5 (verborgen potentieel gevaar)



De camera-auto slaat rechtsaf en moet een tweerichtingsfietspad over. Een bus steekt langzaam het kruispunt over. Daardoor heeft de bestuurder van de camera-auto lange tijd niet kunnen zien of er fietsers van links naderen. Het gevaar zijn mogelijke fietsers van links op wie de bus het zicht ontnemt.

Film 6 (zichtbaar potentieel gevaar)



De camera-auto rijdt tamelijk dicht op de voorligger. De bestuurder van de camera-auto weet dat er een bromfietser voor het bestelbusje rijdt. Deze bromfietser was aan het begin van de film in beeld. Wanneer de bromfietser plotseling oversteeft naar het tweerichtingsfietspad links (let op het bord), moet het bestelbusje hard remmen. Hierdoor kan de camera-auto op het bestelbusje botsen.

Film 7 (verborgen potentieel gevaar)



De camera-auto rijdt rechtdoor over het kruispunt heen. Voor de camera-auto slaat een bus langzaam linksaf. Hierdoor heeft de bestuurder van de camera-auto lange tijd geen zicht gehad op mogelijk tegemoetkomend verkeer (dat op het kruispunt bijvoorbeeld linksaf wil slaan) of voetgangers die achter de bus van links naar rechts willen oversteken. Het gevaar zijn de mogelijke verkeersdeelnemers op botskoers achter de bus.

Film 8 (zichtbaar potentieel gevaar)



De camera-auto passeert een rij geparkeerde auto's. De brommer rijdt even hard als de camera-auto. Een van de geparkeerde auto's heeft een open dak en er zit een bestuurder in. Het voorwiel van die geparkeerde auto draait naar buiten. Deze auto zou direct na het passeren van de brommer weg kunnen rijden en in botsing kunnen komen met de camera-auto.

Film 9 (verborgen potentieel gevaar)



De camera-auto slaat linksaf. De vrachtauto wil ook linksaf slaan, maar wacht. Verkeer rechts van de vrachtauto dat door de vrachtauto niet te zien is, zou kunnen naderen en het kruispunt kunnen oversteken. De camera-auto zou met dit mogelijke verkeer in botsing kunnen komen.

Film 10 (zichtbaar potentieel gevaar)



De camera-auto rijdt rechtdoor. De groene auto staat stil met de alarmlichten aan. Achter de groene auto nadert snel een motorfiets. De motorfietser zou kunnen proberen nog snel de parkeerde auto te passeren terwijl de camera-auto nadert.

Film 11 (verborgen potentieel gevaar)



De camera-auto haalt een bus in die stilstaat bij een bushalte. Een stukje voor de bus ligt een voetgangersoversteekplaats. De fietser links is zojuist gestopt. Uitgestapte passagiers zouden vlak voor de bus langs kunnen oversteken, zonder de moeite nemen om over het even verderop gelegen zebrapad te lopen.

Film 12 (zichtbaar potentieel gevaar)



De camera-auto rijdt rechtdoor. De naderende auto links is geen probleem (voldoende ruimte voor twee auto's). De bromfietser loopt snel op de fietser in en de camera-auto loopt weer op de bromfietser in. De bromfietser zou zonder achterom te kijken naar links kunnen gaan om de fietser in te halen en daarmee op botskoers met de camera-auto kunnen komen.

Film 13 (verborgen potentieel gevaar)



De camera-auto rijdt rechtdoor. Het rode busje is net gestopt. De bestuurder van de camera-auto heeft door de groene bus, en eerder in de film door een vrachtauto met oplegger die voor het rode busje linksaf is geslagen, al geruime tijd geen zicht gehad op mogelijk verkeer van rechts. Hoewel de camera-auto voorrang heeft, zou verkeer van rechts de weg op kunnen draaien.



Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV
SWOV Institute for Road Safety Research

Postbus 93113
2509 AC Den Haag
Bezuidenhoutseweg 62
2594 AW Den Haag
T 070 - 317 33 33
F 070 - 320 12 61
E info@swov.nl
I www.swov.nl

PO Box 93113
2509 AC The Hague, The Netherlands
Bezuidenhoutseweg 62
2594 AW The Hague, The Netherlands
T 070 - 317 33 33
F 070 - 320 12 61
E info@swov.nl
I www.swov.nl