

Situatiebewustzijn in het verkeer

Een onderzoek naar de validiteit en betrouwbaarheid van de Toets Situatiebewustzijn

Guus Lambert



Situatiebewustzijn in het verkeer

**Een onderzoek naar de validiteit en betrouwbaarheid van de
Toets Situatiebewustzijn**

Guus Lambert

**Begeleiders: Erik Roelofs, Cito
Frans Prins, Universiteit Utrecht
Liesbeth Baartman, Universiteit Utrecht**



**Cito
Arnhem, juni 2008**

Uit deze uitgave mogen zonder toestemming van de auteur geen aanhalingen worden overgenomen.

Samenvatting

Ondanks het feit dat jaarlijks het aantal verkeersslachtoffers daalt, blijft de overheid het belangrijk vinden de veiligheid op de weg verder te verbeteren. Niet alleen door wegsignalering en verkeersborden, maar ook door het verbeteren van verkeersinzicht en het vergroten van zelfstandig en verantwoordelijk rijgedrag. Per 1 januari 2008 is dan ook een vernieuwd rijexamen ingevoerd met daarin aandacht voor hogere orde vaardigheden, zoals gevaarherkenning in het verkeer. Sitatiebewustzijn van de bestuurder is belangrijk voor een veilige en verantwoorde verkeersdeelname. Het samenwerkingsverband CitoDrive ontwikkelt instrumenten voor het meten van hogere orde vaardigheden inzake deelname aan het verkeer, waaronder de Toets Sitatiebewustzijn. In dit onderzoek is onderzocht of de Toets Sitatiebewustzijn in voldoende mate valide en betrouwbaar is. De afnamecondities zijn bekeken. Daarnaast is de interne consistentie van de toets onderzocht, het antwoordmodel bekeken, de samenhang met externe criteria zoals ongevalbetrokkenheid is onderzocht en ten slotte is de samenhang met persoonlijke achtergrondvariabelen, zoals leeftijd bekeken. Uit de resultaten bleek dat de afnamecondities niet optimaal waren, vooral de kwaliteit van de filmpjes werd als onvoldoende ervaren. Op bepaalde aspecten was er samenhang met externe criteria en bepaalde persoonlijke achtergrondvariabelen hingen samen met de toetsscores. Op basis van de resultaten en conclusies worden enkele aanbevelingen gedaan ter verbetering van de toets en worden enkele implicaties voor vervolgonderzoek gegeven. Dit verslag is tevens een onderzoeksrapport van het project 'multimediale toetsen rijvaardigheid'.

Voorwoord

Met deze masterthesis sluit ik mijn opleiding Onderwijskunde aan de Universiteit Utrecht af. Daarmee komt een einde aan twee interessante, leerzame jaren waarin duidelijk is geworden wat de onderwijskunde inhoudt en waar een wetenschappelijk onderwijskundige zich zoal mee bezig kan houden.

In dit voorwoord wil ik met nadruk aangeven dat ik, na enkele basiscolleges methodeleer en statistiek, pas tijdens het schrijven van deze afstudeerscriptie goed wetenschappelijk onderzoek heb leren doen. Daarvoor wil ik in het bijzonder Dr. Erik C. Roelofs bedanken. Zonder zijn leerzame begeleiding, heldere feedback en gedrevenheid had dit eindresultaat niet tot stand kunnen komen.

Daarnaast wil ik Cito bedanken voor het bieden van de mogelijkheid tot het doen van een afstudeeronderzoek. Ook de respondenten die meegewerkt hebben aan dit onderzoek wil ik van harte bedanken, zonder hen was dit afstudeeronderzoek niet geslaagd.

Tevens wil ik Dr. Frans Prins en Dr. Liesbeth Baartman bedanken voor hun begeleiding, hulp en feedback vanuit de universiteit.

Ten slotte wil ik mijn ouders, broer en mijn vrienden bedanken voor hun steun en interesse tijdens deze laatste fase van mijn studie. Zonder hun steun en interesse had dit afstudeeronderzoek niet tot stand kunnen komen.

Ik hoop dat deze masterthesis een bijdrage levert aan de ontwikkeling van meetinstrumenten voor verkeersassessment. Ook hoop ik dat deze thesis als voorbeeld kan dienen bij de ontwikkeling van nieuwe, valide en betrouwbare, multimediale assessmentmiddelen.

Utrecht, juni 2008

Guus Lambert

Afstudeeronderzoek Master Onderwijskundig Ontwerp en Advisering
Faculteit Sociale Wetenschappen, Universiteit Utrecht
Cito, Arnhem

Inhoudsopgave

Samenvatting

Voorwoord

1.	Context	5
1.1	Inleiding	5
1.2	De ontwikkeling van instrumenten	6
2.	Het meten van Situatiebewustzijn	8
2.1	Wat is situatiewustzijn?	8
2.1.1	Situatiewustzijn: perceptie, begrip en projectie	9
2.1.2	Beslissingen nemen	10
2.1.3	Prestaties	10
2.1.4	Taak- en systeemfactoren	10
2.1.5	Individuele factoren	11
2.2	Het meten van situatiewustzijn	12
2.3	Bestaande meetinstrumenten van situatiewustzijn	13
2.3.1	Meetinstrumenten Gugerty (1997)	13
2.3.2	Meetinstrumenten McKenna, Horswill & Alexander (2006)	15
2.3.3	Meetinstrumenten en aannames Endsley (2000)	17
2.4	Validiteit	18
2.5	Betrouwbaarheid	23
2.6	Onderzoeksvragen en verwachtingen	24
3.	Methode	26
3.1	Subjecten	26
3.2	Instrumenten	27
3.2.1	Toets Situatiewustzijn	27
3.2.2	Zelfevaluatie Rijgedrag	31
3.2.3	Vragenlijst over de Toets Situatiewustzijn	34
3.3	Design en dataverzameling	34
3.4	Analyseplan	35
4.	Resultaten	38
4.1	Betrouwbaarheid van de Toets Situatiewustzijn en haar subschalen	38
4.2	Realisatie van de afnamecondities	41

4.3	Analyse van antwoordgedrag en antwoordmodel bij de vier opgaventypen	44
4.3.1	Opgaventype 'recollection'	45
4.3.2	Opgaventype 'action selection'	48
4.3.3	Opgaventype 'hazard perception'	53
4.3.4	Opgaventype 'timed response'	53
4.4	Samenhang tussen de toetsscores en externe criteria	59
4.4.1	Samenhang met Zelfevaluatie Rijgedrag.....	59
4.4.2	Samenhang met ongevalbetrokkenheid	60
4.5	Relatie tussen toetsresultaten en achtergrondkenmerken.....	60
5.	Conclusie en discussie	62
5.1	Onderzoeksvragen.....	62
5.2	Conclusies	62
5.2.1	Taakdomein	62
5.2.2	Interne consistentie in relatie tot afnamecondities en antwoordmodel	63
5.2.3	Samenhang Toets Situatiebewustzijn met externe criteria	64
5.2.4	Samenhang Toets Situatiebewustzijn met persoonlijke achtergrondvariabelen	64
5.3	Beperkingen onderzoek en toets	65
5.4	Verbetervoorstellen Toets Situatiebewustzijn	66
5.4.1	Algemene verbetervoorstellen Toets Situatiebewustzijn	66
5.4.2	Verbetervoorstellen opgaventype 'recollection'	68
5.4.3	Verbetervoorstellen opgaventype 'action selection'	68
5.4.4	Verbetervoorstellen opgaventype 'hazard perception'	68
5.4.5	Verbetervoorstellen opgaventype 'timed response'	69
5.5	Implicaties voor vervolgonderzoek.....	69
6.	Literatuur	71
	Bijlage A Gegevens over subjecten.....	73
	Bijlage B Draaiboek filmen videofragmenten	75
	Bijlage C Vragenlijst.....	79
	Bijlage D Itemkarakteristieken gereviseerde toets	85
	Bijlage E Statistieken Toets Situatiebewustzijn.....	87
	E1 Samenhang tussen de waardering van afnamecondities en de toetsscores	87
	E2 Samenhang tussen achtergrondvariabelen en toetsscores	95
	Bijlage F Grafieken Timed Response	97

1. Context

1.1 Inleiding

Uit de Veiligheidsbalans 2007, een gezamenlijke uitgave van de Inspectie Verkeer en Waterstaat en Rijkswaterstaat, blijkt dat het aantal slachtoffers in het verkeer in 2006 opnieuw is afgenomen. Ook werd in februari 2007 in verschillende media gemeld dat Nederland de veiligste wegen van Europa heeft. De afname van het aantal verkeersdoden en ziekenhuisslachtoffers is sinds 1972 jaarlijks een feit. Desondanks waren er in 2006 nog altijd 811 verkeersdoden en 16.750 ziekenhuisslachtoffers te betreuren in het verkeer. Het merendeel van de ziekenhuisslachtoffers dat in het ziekenhuis terechtkomt, behoort tot de verkeersgroep van (brom- en snor)fietzers. De meeste doden in het verkeer vallen onder bestuurders van personenauto's (Inspectie Verkeer en Waterstaat & Rijkswaterstaat, 2007). Een punt van aandacht is dat één op de vier jongeren binnen vier jaar na het behalen van het rijbewijs betrokken raakt bij een verkeersongeluk. In vergelijking met andere weggebruikers zijn jongeren relatief vaak betrokken bij verkeersongelukken. Ook bestuurders van leaseauto's zijn vaker betrokken bij ongevallen dan personen die in een eigen auto rijden.

Ondanks het feit dat jaarlijks het aantal verkeersslachtoffers daalt, blijft de overheid het belangrijk vinden de veiligheid op de weg verder te verbeteren. Niet alleen door wegsignalering en verkeersborden, maar ook door het verbeteren van verkeersinzicht en het vergroten van zelfstandig en verantwoordelijk rijgedrag. Uit onderzoek van Hatakka, Keskinen, Gregersen, Glad & Hernetkoski (2002) blijkt namelijk dat het bezit van de juiste psychomotorische vaardigheden en psychologische functies niet voldoende is om veilig en verantwoord deel te kunnen nemen aan het verkeer. Naast vaardigheden als het besturen van een voertuig en het beheersen van bijvoorbeeld kijktechnieken, spelen volgens hen hogere orde vaardigheden een belangrijke rol. Voorbeelden van deze vaardigheden zijn gevaarherkenning, inzicht in verkeerssituaties en het tijdig anticiperen op mogelijk gevaar. Deze hogere orde vaardigheden vereisen een integratie van verschillende andere vaardigheden. Naast genoemde vaardigheden hebben ook persoonlijke motieven, achtergronden en attitudes grote invloed op het rijgedrag en de betrokkenheid bij ongevallen (Rothengatter, 1997). Het ontwikkelen van hogere orde vaardigheden dient te geschieden aan de hand van langdurende verkeerseducatie, dat ook na het behalen van het rijbewijs dient te worden voortgezet (Roelofs, Van Onna, Nägele, Mesken, Kuiken & Cozijnsen, 2007).

Bovenstaande inzichten hebben ertoe geleid dat het ministerie van Verkeer en Waterstaat vernieuwing van het rijexamen nodig achtte. In het vernieuwde theorie- en praktijkexamen komen nieuwe onderdelen naar voren, waarmee beoogd wordt bovengenoemde hogere orde vaardigheden te toetsen. Als nieuw onderdeel wordt gevaarherkenning getoetst. Per 1 januari 2008 is met goedkeuring van het ministerie het vernieuwde rijexamen van het Centraal Bureau Rijvaardigheidsbewijzen (CBR) ingevoerd (CBR, 2007). Verwacht mag worden dat de rijopleidingen hun opleidingen aanpassen om aan de eisen van het vernieuwde theorie- en praktijkexamen te kunnen voldoen. Tevens wordt verwacht dat post-rijopleidingen hun trainingen en toetsvormen bijstellen om bij te dragen aan een cultuur van permanent lerende autobestuurders.

1.2 De ontwikkeling van instrumenten

In lijn met de in de vorige paragraaf beschreven ontwikkelingen ontwikkelt Cito in samenwerking met DHV en TNO Human Factors instrumenten voor het meten van hogere orde vaardigheden inzake deelname aan het verkeer. Samen vormen de genoemde organisaties het samenwerkingsverband CitoDrive. CitoDrive ontwikkelt een webportal (CitoDrive Portal) waar de instrumenten op worden aangeboden. Het doel van deze instrumenten is het toetsen of verkeersdeelnemers de kritische competenties beheersen die nodig zijn voor veilige en verantwoorde verkeersdeelname. Onder kritische competenties worden onder andere gevaarherkenning, inzicht in verkeerssituaties en het tijdig anticiperen op mogelijk gevaar verstaan, maar ook sociaal, doorstromingbevorderend, gecontroleerd en milieubewust rijgedrag zijn hierbij van belang. De belangrijkste doelgroep bestaat uit bestuurders die hun rijbewijs al hebben behaald en zich dagelijks voor hun werk op de weg bevinden.

In de periode van januari 2007 tot juni 2008 zijn vier instrumenten ontwikkeld: een instrument voor zelfevaluatie van rijvaardigheid, een Assessment Bestuurdersrisico, de Toets Situatiewustzijn en een Performance Assessment Rijvaardigheid (PAR). In Tabel 1.1 worden de instrumenten kort gekarakteriseerd.

Tabel 1.1

Beschrijving Instrumenten CitoDrive

Instrument	Omschrijving
Zelfevaluatie Rijgedrag	Hiermee wordt beoogd de kwaliteit van het rijgedrag in kaart te brengen op een vijftal kwaliteitsaspecten, namelijk: veilig, sociaal, doorstroming bevorderend, gecontroleerd en milieubewust en energiezuinig rijgedrag.
Assessment Bestuurdersrisico	Webbased vragenlijst met als doel het bepalen van risicogroepen (laag, midden en hoog) en het besluiten welke vorm van rijvaardigheidstraining zinvol is.
Toets Situatiewustzijn	Webbased digitale videotest met als doel het bepalen of bestuurders in staat zijn om adequaat te reageren op veranderingen in hun omgeving. De toets kent vier typen opgaven: het opmerken van factoren, het beslissen tot handelen, de aandacht richten en tijdig actie ondernemen.
Performance Assessment Rijvaardigheid	Webbased formulieren waarmee de kwaliteit van het praktische rijgedrag van bestuurders in kaart wordt gebracht. De formulieren worden ingezet tijdens een rijvaardigheidstraining en verschaffen inzicht in de sterke en zwakke punten van de kandidaat.

Voordat de genoemde instrumenten definitief aangeboden worden op de markt, zal voor elk instrument een vooronderzoek moeten worden uitgevoerd om te toetsen of de instrumenten in voldoende mate betrouwbaar en valide zijn. Het onderzoek in deze thesis heeft betrekking op de Toets Situatiewustzijn. Centraal in dit onderzoek staat de volgende algemene vraagstelling: *In hoeverre meet de ontwikkelde toets op een betrouwbare en valide wijze situatiewustzijn bij automobilisten?* Een uitgebreide beschrijving van de Toets Situatiewustzijn wordt gegeven in paragraaf 3.2 *Instrumenten*.

2. Het meten van Situatiebewustzijn

In dit onderzoek staat het meten van kwaliteiten van een autobestuurder centraal. Aan de hand van de toets kan bepaald worden op welke aandachtsgebieden training zinvol wordt geacht. De functie van de Toets Situatiebewustzijn is *plaatsing* voor deelname aan trainingen in rijvaardigheid. Plaatting is één van de mogelijke functies van toetsen (zie Moelands & Sanders, 1996).

Tijdens de constructie van de Toets Situatiebewustzijn, is via literatuurstudie onderzocht hoe situatiewustzijn gemeten kan worden. Het begrip situatiewustzijn wordt hieronder uiteengezet, grotendeels aan de hand van het model van Endsley (1995b). Daarbij wordt beschreven hoe het construct werd gemeten in eerdere onderzoeken. In de paragrafen 2.4 en 2.5 wordt ingegaan op de validiteit en betrouwbaarheid van toetsen in het algemeen en van toetsen voor situatiewustzijn in het bijzonder.

2.1 Wat is situatiewustzijn?

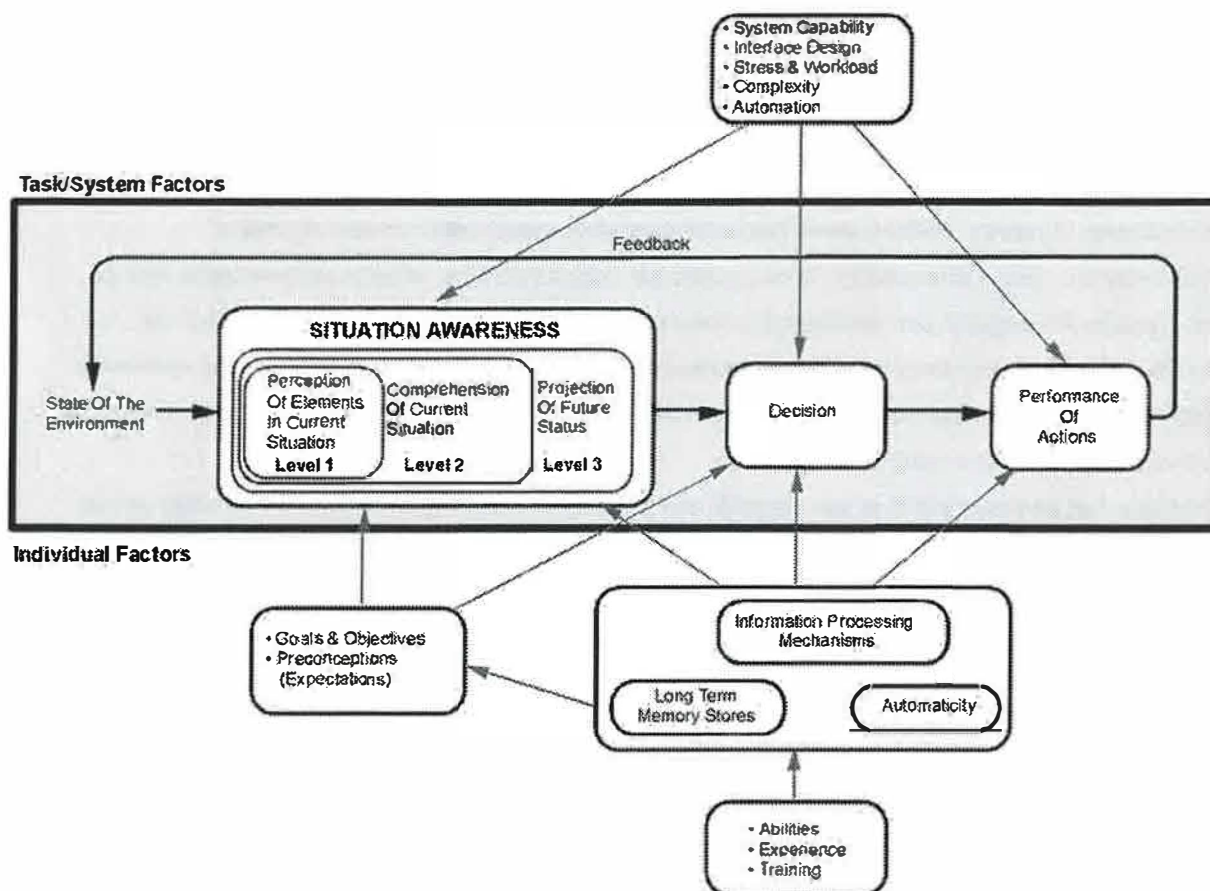
Het construct situatiewustzijn kent zijn oorsprong in de luchtvaart (Roelofs et al., 2008), maar is de laatste decennia ook toegepast op andere domeinen. Het verbeteren van situatiewustzijn is een belangrijk doel geworden van ontwikkelaars van toetsen en trainingen voor onder andere piloten, operators en weggebruikers (Endsley, 2000). Ook vanuit militair onderzoek is veel aandacht voor het begrip (Stanner & French, 2005). Situatiewustzijn vormt inmiddels een belangrijk onderdeel binnen cognitief onderzoek (O'Brien & O'Hare, 2007).

De oorzaak van de interesse in situatiewustzijn kan volgens Endsley (2000) verklaard worden door de vele technische ontwikkelingen die in de afgelopen decennia hebben plaatsgevonden. De mens wordt blootgesteld aan steeds meer prikkels in vergelijking met enkele decennia geleden. Ondanks de vele prikkels is het in verschillende domeinen, zoals ook in het verkeer, essentieel dat men zich goed bewust is en blijft van de situatie om zich heen.

In de literatuur komen verschillende definities van situatiewustzijn naar voren. Een veel gebruikte en algemeen geaccepteerde definitie van situatiewustzijn is die van Endsley (1988): 'The perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.' (p. 97).

Situatiewustzijn speelt een grote rol in het wegverkeer. Situaties in het verkeer veranderen continu en anticipatie door de autobestuurder is letterlijk van levensbelang. Daarbij speelt gevaarherkenning een grote rol (Horswill & McKenna, 2004). Het tijdig anticiperen op de potentiële gevaren voorkomt risicovolle situaties of verkeersongelukken. Het herkennen van potentiële gevaren in het verkeer is daarom kenmerkend voor de mate van situatiewustzijn.

Het model van situatiewustzijn van Endsley (1995b) is weergegeven in Figuur 2.1. Het model wordt hieronder toegelicht.



Figuur 2.1 Model Situation Awareness (Endsley, 1995b)

Endsley (2000) benoemt in het model verschillende niveaus en factoren die van invloed zijn op situatiebewustzijn. Onderscheid wordt gemaakt tussen situatiebewustzijn zelf en de factoren die situatiebewustzijn beïnvloeden. Deze niveaus en factoren worden hieronder toegelicht. De kern van het model bestaat uit situatiebewustzijn, het nemen van beslissingen en het daadwerkelijk handelen.

2.1.1 Situatiebewustzijn: perceptie, begrip en projectie

Situatiebewustzijn bestaat volgens Endsley (2000) uit drie niveaus: perceptie, begrip en projectie. Fundamenteel bij situatiebewustzijn is de perceptie van de belangrijke elementen in een situatie. Het is bijvoorbeeld belangrijk dat een autobestuurder de medeweggebruikers, verkeerstekens en alle andere omgevingsfactoren waarneemt. Bij het waarnemen is aandacht een belangrijk element (Endsley, 1995b; Durso & Gronlund, 1999).

Het afzonderlijk waarnemen van losse elementen is echter niet voldoende voor het besturen van een auto. Het is belangrijk dat de bestuurder de waargenomen elementen in een situatie begrijpt en de relevante elementen integreert tot één geheel, zodat hij daarnaar kan handelen. De bestuurder beslist welke elementen relevant zijn voor het veilig uitvoeren van rijtaken en welke niet, zodat rijtaken veilig worden uitgevoerd.

Het veilig besturen van een auto is echter niet mogelijk zonder het derde en tevens hoogste niveau van situatiebewustzijn: projectie. Projectie is het voorspellen van mogelijke reacties van medeweggebruikers. Op basis van die voorspellingen en de doelen die de bestuurder voor ogen heeft, zal de autobestuurder een beslissing nemen over de beste wijze van handelen. Die beslissing

zal zich meestal uiten in de handelingen van de bestuurder. Dit proces van waarnemen, begrijpen, projecteren, beslissen en handelen, moet in hoog tempo plaatsvinden. Tijd en snelheid spelen dus een grote rol bij situatiebewustzijn, vooral bij niveau 2, begrip van de situatie, en niveau 3, projectie.

2.1.2 Beslissingen nemen

In het model van Endsley (1995b) wordt het beslisproces losgekoppeld van het eigenlijke situatiebewustzijn. Situatiewustzijn wordt gezien als de persoonlijke, interne representatie van de situatie, waarop vervolgens een beslissing gebaseerd wordt. Endsley (2000) stelt daarnaast dat iemand zich volledig bewust kan zijn van de verkeerssituatie, maar vervolgens alsnog een verkeerde beslissing kan nemen, omdat de persoon in kwestie niet beschikt over de essentiële kennis of ervaring over de ideale handeling in een situatie.

Ondanks het onderscheid is er wel degelijk een positieve relatie tussen situatiebewustzijn en het nemen van beslissingen, zo blijkt uit onderzoek van Stanners en French (2005). Endsley (2000) erkent dat situatiebewustzijn en het nemen van beslissingen zeer sterk samenhangen, maar ziet het niet als één proces. Volgens hem ligt situatiebewustzijn ten grondslag aan beslissen en handelen en worden de begrippen daarom conceptueel gescheiden.

2.1.3 Prestaties

Volgens Endsley (1995b) is er naast een link tussen situatiebewustzijn en beslissen, tevens een link tussen situatiebewustzijn en prestatie. Over het algemeen is te verwachten dat de prestatie slecht is, wanneer het situatiebewustzijn laag is. Prestatie kan echter ook laag zijn wanneer men onbekend is met de benodigde handeling of wanneer tijd of andere beperkende factoren het onmogelijk maken om juist te handelen. Een slechte prestatie staat dus niet gelijk aan een lager situatiebewustzijn en daarom is presteren conceptueel gescheiden van situatiebewustzijn en beslissen.

Situatiebewustzijn, beslissing en prestatie worden volgens het model beïnvloed door verschillende factoren, zowel taak- en systeemfactoren, als individuele factoren. Hieronder worden deze factoren kort beschreven.

2.1.4 Taak- en systeemfactoren

Verschillende taak- en systeemfactoren beïnvloeden het situatiebewustzijn, de beslissing en de prestatie. Endsley (1995b) onderscheidt de volgende factoren:

- 'System capability': in hoeverre informeert het systeem (bijvoorbeeld een navigatiesysteem of radar) de bestuurder over de benodigde informatie in de omgeving? Deze factor geldt vooral bij het besturen van vliegtuigen, niet zozeer bij het besturen van auto's.
- 'Interface design': op welke manier wordt de informatie gepresenteerd?
- 'Stress' & 'Workload': de rol van fysieke stressfactoren (bijvoorbeeld geluid, hitte, medicijnen, drank, vermoeidheid) en sociaalpsychologische stressfactoren (bijvoorbeeld angst, onzekerheid, tijdsdruk, e.d.). Hierdoor kunnen bepaalde elementen in een verkeerssituatie bewust of onbewust genegeerd worden. Een hoge mentale werkdruk is een belangrijke stressfactor. Wanneer de capaciteit van het werkgeheugen de informatiestroom niet aankan, zal het situatiebewustzijn afnemen.

- 'Complexity': de complexiteit van een verkeerssituatie en weinig ervaring met dat soort situaties kunnen ervoor zorgen dat het situatiebewustzijn afneemt.
- 'Automation': situatiebewustzijn wordt aangetast, wanneer handelingen worden overgenomen door automatische systemen. Dit geldt vooral voor bijvoorbeeld piloten die met de automatische piloot vliegen. Wanneer het fout gaat, moeten piloten in staat zijn de problemen manueel op te lossen.

2.1.5 Individuele factoren

Naast taak- en systeemfactoren, worden het situatiebewustzijn, de beslissing en de prestatie beïnvloed door individuele factoren, zoals de hoeveelheid training en ervaring van een individu. In het model van Endsley (1995b) worden de volgende individuele factoren genoemd:

- Informatieverwerking: hoe snel verwerkt iemand de juiste informatie?
- Opslag informatie in lange termijn geheugen: hoe slaat iemand informatie op en hoe wordt de opgeslagen informatie gebruikt?
- Automatisering: de mate van ervaring en/of training bepaalt de mate waarin automatisch gereageerd wordt op een bekende situatie en bepaalt daarmee de kwaliteit en snelheid van het proces van beslissen en handelen.
- Doelen: de doelen die iemand heeft beïnvloeden de mate van situatiebewustzijn en de aard van beslissingen. Iemand verzamelt de benodigde informatie die nodig is om het doel te bereiken. Het situatiebewustzijn zal bijvoorbeeld anders zijn wanneer iemand tot doel heeft om tussen twee voertuigen te parkeren dan wanneer het doel is om op een leeg parkeerveld te parkeren.
- Verwachtingen: de verwachtingen die iemand heeft beïnvloeden ook in welke mate iemand zich bewust is van een situatie en de aard van de beslissingen. Verschil in verwachtingen kan ontstaan door verschil in ervaring: een ervaren autobestuurder die eerder soortgelijke verkeerssituaties heeft meegemaakt verwacht andere zaken dan een beginnende bestuurder die nog nooit een soortgelijke verkeerssituatie heeft meegemaakt. De ervaren bestuurder kan bijvoorbeeld de reacties van medeweggebruikers beter schatten.

Bovenstaande individuele factoren worden beïnvloed door training, ervaring en de mogelijkheden van een individu. Daarbij wordt het verschil tussen beginners – zoals leerling-bestuurders – en gevorderden – zoals ervaren leaserijders – duidelijk. Durso en Gronlund (1999) menen dat experts, door hun ervaring, in complexe situaties veel tijd besteden aan het toetsen van de situatie. Zij schatten situaties efficiënter en op een meer diagnostische manier dan beginners. De expert houdt zich vooral bezig met de aard van de situatie en koppelt daar vrijwel automatisch een actie aan (automatisering), terwijl beginners minder sterk de situatie analyseren en veelal gericht zijn op de mogelijkheden tot handelen (Randel & Pugh, 1996). Uit onderzoek van Randel & Pugh (1996) blijkt tevens dat experts gevaren (in het betreffende onderzoek was dat de positie van vijandelijke troepen) beter herinneren dan beginners. Dit is het gevolg van eerdere ervaringen, feedback op eerder handelen en de daardoor opgebouwde geheugenstructuren. Endsley (1995b) noemt dit in zijn model 'Long Term Memory Stores'. Daarnaast verschillen volgens Endsley (1995b) de verwachtingen en doelstellingen van

beginners met die van experts, waardoor het situatiebewustzijn en de aard van de beslissing verschillen. Ook dit verschil wordt bepaald door ervaring.

Per situatie kunnen zowel taak- en systeemfactoren als individuele factoren verschillen. Door dit verschil in factoren zal ook het situatiebewustzijn van een bestuurder per situatie verschillen. Het is echter in het belang van de verkeersveiligheid dat autobestuurders situatiebewust zijn in alle verkeerssituaties en (in hoog tempo) adequaat kunnen handelen.

2.2 Het meten van situatiebewustzijn

Door situatiebewustzijn te meten en ook trainen kan situatiebewustzijn bij verkeersdeelnemers worden bevorderd, waardoor ook de verkeersveiligheid in het algemeen wordt bevorderd. Wanneer het gaat om het meten van situatiebewustzijn, moeten verschillende aannames worden gedaan. Endsley (2000) doet een aantal aannames, welke in deze paragraaf worden besproken en in de volgende paragraaf naast eerder ontwikkelde instrumenten voor het meten van situatiebewustzijn worden gelegd.

Het concept situatiebewustzijn bevat volgens Endsley (2000) alleen iemands bewustzijn over de dynamische situatie waarin hij zich bevindt en niet de onderliggende processen die ervoor zorgen dat die staat van bewustzijn bereikt wordt. Een goede informatieverzamelaar hoeft namelijk niet altijd in hoge mate situatiebewust te zijn. Dat is bijvoorbeeld het geval wanneer een doorgaans goede informatieverzamelaar door gebrek aan ervaring verkeerde informatie verzamelt. De informatieverzameling kan dan goed zijn, men is zich echter niet in dezelfde goede mate bewust van de situatie. Het meten van dit soort onderliggende processen is dus een indirecte meting van situatiebewustzijn. Het geeft informatie over de manier waarop iemand informatie verzamelt, niet over situatiebewustzijn.

Zoals in de vorige paragraaf werd beschreven maakt Endsley (2000) een conceptueel onderscheid tussen situatiebewustzijn en het nemen van beslissingen. Hetzelfde geldt voor gedrag en prestaties. Aangenomen kan worden dat de kwaliteit van situatiebewustzijn doorwerkt in de genomen beslissingen en het getoonde rijgedrag. Een sterkere aanname is dat uit gedrag afgeleid kan worden of iemand zich bewust is van de situatie in plaats van uit onderliggende processen.

Een derde aanname van Endsley (2000) is dat de kwaliteit van situatiebewustzijn afhangt van de hoeveelheid aandacht die een bestuurder beschikbaar heeft in een situatie. Wanneer de aandacht wordt verstoord, gaat dit ten koste van de kwaliteit van situatiebewustzijn (Durso & Gronlund, 1999). Metingen van situatiebewustzijn moeten niet gepaard gaan met afleiding door irrelevante zaken, waardoor het proces van situatiebewustzijn zelf wordt verstoord.

Een vierde aanname bij het meten van situatiebewustzijn is dat het begrip moeilijk te meten is. Vooral bij geautomatiseerde processen kan het voorkomen dat men zich bewust is van de situatie en de juiste beslissingen neemt, maar niet in staat is de genomen beslissingen te beargumenteren. Daarnaast wordt de informatie slechts tijdelijk in het werkgeheugen opgeslagen. De informatie is na bepaalde tijd alleen nog beschikbaar als het is opgeslagen in het lange termijn geheugen en de informatie is dan minder gedetailleerd. Het is daarom van belang dat metingen direct plaatsvinden.

Echter, het proces van meten tijdens de situatie kan ongewenst de situatie zélf beïnvloeden. Dit zou zoveel mogelijk voorkomen moeten worden.

Endsley (2000) noemt voorts dat een relatie tussen situatiebewustzijn en werkbelasting (workload) in bepaalde situaties mogelijk is. Een toename van de complexiteit in verkeerssituaties (veel medeweggebruikers, werkzaamheden, snelheid van handelen en weersomstandigheden) leidt tot een hogere mentale werklast. Een hogere werklast kan invloed hebben op situatiebewustzijn, maar werklast en situatiebewustzijn dienen volgens Endsley (2000) onafhankelijk van elkaar te worden gemeten. Het is niet zo dat bij een lage werklast automatisch een hoog situatiebewustzijn aanwezig is. Door bijvoorbeeld een lage motivatie of onoplettendheid kan het situatiebewustzijn laag zijn.

Ten slotte kan de mate van situatiebewustzijn niet in direct meetbare, absolute waarden worden uitgedrukt. Wanneer de mate van situatiebewustzijn toeneemt, neemt slechts de *kans* op het nemen van de juiste beslissingen en prestaties toe. Dit hoeft niet daadwerkelijk het geval te zijn, bijvoorbeeld wanneer iemand in een complexe verkeerssituatie komt, maar niet over de vaardigheden beschikt om in die situatie juist te handelen of wanneer iemand bij toeval de juiste handelingen verricht zonder dat hij zich bewust is van de situatie. Van betere prestaties kan dus niet direct afgeleid worden dat het situatiebewustzijn is toegenomen.

Als bij instrumenten voor het meten van situatiebewustzijn rekening wordt gehouden met bovenstaande aannames, is de verwachting dat een betere constructrepresentatie van situatiebewustzijn wordt bereikt.

2.3 Bestaande meetinstrumenten van situatiebewustzijn

Om de mogelijkheden waarop situatiebewustzijn in het verkeer kan worden gemeten na te gaan, worden in deze paragraaf een aantal onderzoeken besproken waarin verschillende typen meetinstrumenten zijn onderzocht. Allereerst wordt gekeken naar meetinstrumenten uit onderzoek van Gugerty (1997), daarna staan de meetinstrumenten uit het onderzoek van McKenna, Horswill & Alexander (2006) centraal. Een relatie wordt gelegd met de aannames van Endsley (2000).

2.3.1 Meetinstrumenten Gugerty (1997)

Gugerty (1997) onderscheidt verschillende typen metingen: herinneringsmetingen, deelprestatiemetingen (zoals het herkennen van mogelijke gevaren) en metingen met betrekking tot de algemene prestatie (de algemene prestatie in de auto waarmee een ongeluk wordt voorkomen). Een herinnering vindt intern plaats en is niet tastbaar zoals een (deel)prestatie en wordt daarom als een apart type meting genoemd.

Herinneringsmetingen kunnen volgens Gugerty (1997) met behulp van een simulator worden uitgevoerd. Een kandidaat bevindt zich in een gesimuleerde, realistische verkeerssituatie welke 18 tot 35 seconden duurt. Vervolgens wordt het scherm wit en wordt een herinneringsvraag gesteld, bijvoorbeeld over de positie van medeweggebruikers. Hiermee wordt getoetst of de bestuurder tot in detail op de hoogte is van de situatie om hem heen. Gugerty (1997) stelt echter dat met alleen herinneringsitems situatiebewustzijn niet in kaart gebracht kan worden. Bovendien vinden in de rijpraktijk vaak goed geoefende, automatische processen plaats, waarbij het niet altijd nodig is dat

informatie kort daarna nog wordt herinnerd. Dit automatisme komt tevens als individuele invloedsfactor naar voren in het model van Endsley (1995b). Het is om die reden belangrijk te zoeken naar meetmethoden die niet afhankelijk zijn van expliciete herinnering. Deze methoden noemt Gugerty (1997) (deel)prestatiemetingen van situatiebewustzijn.

Het meten van (deel)prestatiemetingen kan eveneens geschieden met behulp van een simulatie. Een voorbeeld uit het onderzoek van Gugerty (1997) is een simulatie waarbij de bestuurder een plotseling opdoemend gevaar moet ontwijken en daarbij moet voorkomen dat hij een eventuele auto naast hem raakt. Daarbij bestuurt de deelnemer de auto zelf met behulp van pijltjes op het toetsenbord. Uit deze taak blijkt of een bestuurder de auto wel of niet gezien had. Indirect wordt gemeten of de deelnemer zicht bewust is van weggebruiker in zijn dode hoek. Met dit soort deelprestatiemetingen beoogt Gugerty (1997) hoofdzakelijk situatiebewustzijn zélf te meten, niet zozeer de beslissing en/of handeling die daaruit voortkomt. Gugerty maakt net als Endsley een conceptueel onderscheid tussen situatiebewustzijn en handeling.

Naast herinneringsmetingen en deelprestatiemetingen gebruikt Gugerty (1997) ook meetinstrumenten die de algemene prestatie meten. Deze metingen bevatten zowel situatiebewustzijn als het beslissings- en reactieproces. Een voorbeeld van dit soort metingen is 'crash avoidance', waarbij de hoeveelheid veilig ontweken gevaren wordt gemeten.

Gugerty (1997) maakt dus onderscheid tussen herinneringsmetingen, deelprestatiemetingen en metingen die de algemene prestatie meten. In Tabel 2.1 zijn Gugerty's definities van de meetinstrumenten weergegeven.

Tabel 2.1

Meetinstrumenten Gugerty (1997)

Meetinstrument	Beschrijving Gugerty (1997)
1. Herinneringsmetingen:	
Recall	Percentage of actual cars recalled.
2. Deelprestatiemetingen:	
Hazard detection	Sensitivity at responding to hazards and not responding on catch trials.
Blocking car detection	Frequency of avoiding some blocking cars on hazard trials.
3. Algemene prestatie (inclusief situatiebewustzijn):	
Crash avoidance	Frequency of avoiding all hazard and blocking cars on hazard trials.

2.3.2 Meetinstrumenten McKenna, Horswill & Alexander (2006)

Ook McKenna, Horswill & Alexander (2006) hebben in hun onderzoek verschillende maten voor situatiebewustzijn gehanteerd. Zij hebben zich in het bijzonder gericht op het aspect gevaarherkenning van situatiebewustzijn. Dit aspect werd daarbij gerelateerd aan de mate waarin men risico's neemt in verkeerssituaties. In het onderzoek komen geen expliciete herinneringsmetingen of algemene prestatiemetingen voor, zoals bij Gugerty (1997). In Tabel 2.2 worden de meetinstrumenten uit het onderzoek weergegeven. Deze meetinstrumenten kunnen ondergebracht worden bij de categorie 'deelprestatiemetingen'.

Tabel 2.2

Meetinstrumenten McKenna, Horswill & Alexander (2006)

Meetinstrument gevaarherkenning	Gericht op
Video speed test	Het herkennen van potentieel gevaarlijke situaties afgeleid van het al dan niet matigen van snelheid.
Close following video test	Het voorkomen van gevaarlijke situaties door het al dan niet behouden van een acceptabele volgafstand.
Gap acceptance video test	Het voorkomen van gevaarlijke situaties door het al dan niet invoegen wanneer de ruimte daarvoor acceptabel is.
Hazard perception video test	Het louter herkennen van potentieel gevaarlijke situaties.

Met een aantal computer-based instrumenten werd in het onderzoek van McKenna, Horswill en Alexander beoogd te meten of het trainen van gevaarherkenning effect had op het nemen van risico's. Gevaarherkenning is essentieel in het verkeer en komt overeen met het onderdeel 'projectie' uit het model voor situatiebewustzijn van Endsley (1995b). Kan een voorspelling worden gemaakt van de situatie waardoor mogelijk gevaarlijke situaties kunnen worden voorkomen? Een set van vier meetinstrumenten uit het onderzoek van McKenna, Horswill & Alexander (2006) wordt hier besproken.

Het eerste instrument dat in het onderzoek werd getoetst is de 'video speed test', een prestatiemeting. In deze test kregen de deelnemers videofragmenten te zien van zeven verkeerssituaties, opgenomen vanuit het perspectief van de bestuurder. In elke situatie moesten de deelnemers aangeven of ze sneller of juist langzamer zouden rijden. De deelnemers noteerden op papier hoeveel sneller of langzamer ze gereden zouden hebben, bijvoorbeeld +10 (10 miles per uur sneller). Van de ingeschatte snelheid kan worden afgeleid of de deelnemer in hoge of lage mate situatiebewust is. Uit het onderzoek bleek dat deze test aan snelheid gerelateerde ongevallen kan voorspellen en onderscheid maakt tussen leeftijd en geslacht. McKenna, Horswill & Alexander (2006) geven dat deze resultaten consistent zijn met eerder door McKenna en Horswill geobserveerde onderzoeken.

Het tweede instrument dat getoetst werd is de 'close following test'. In deze toets zien de deelnemers een videofragment, wederom vanuit het oogpunt van de bestuurder. De bestuurder nadert de auto die voor hem rijdt. De deelnemers moesten met een druk op de knop aangeven wanneer de afstand tussen de auto's ongeveer net zo groot was als de afstand waarop ze in de werkelijkheid een voorgaande auto zouden volgen. De deelnemer moest een tweede keer op de knop drukken wanneer de afstand onacceptabel klein werd. Eerst kreeg men de gelegenheid om te oefenen. Vervolgens werd hetzelfde videofragment vier keer herhaald. De gemiddelde reactietijd van de deelnemer voor de eerste druk op de knop en de gemiddelde reactietijd van de tweede druk op de knop zijn als score gebruikt. Ook dit instrument meet een prestatie. Uit de reacties van de deelnemers blijkt of de deelnemer de situatie wel of niet overziet (situatiebewustzijn). Dit meetinstrument maakte onderscheid tussen de volgafstand van mannen en de volgafstand die vrouwen hanteren, wederom in overeenstemming met eerder door McKenna uitgevoerd onderzoek.

Het derde meetinstrument uit het onderzoek van McKenna, Horswill & Alexander (2006) was de 'gap acceptance video test'. Dit meetinstrument meet geen situatiebewustzijn, maar het nemen van risico's. Het wordt hier besproken, omdat het deel uitmaakte van de set meetinstrumenten die elke deelnemer moest doorlopen. De deelnemer kreeg een ruim vijf minuten durend videofragment te zien. De bestuurder staat op een T-splitsing en moet tussen het rijdende, rechtdoorgaande verkeer invoegen. Het videofragment is vanaf de zijkant van de bestuurder opgenomen, zodat het verkeer waar de bestuurder tussen moet voegen te zien was. Wanneer de deelnemer een acceptabele ruimte zag tussen twee auto's, waarbij het mogelijk wordt geacht in te voegen, drukte hij op een knop. Met dit instrument wordt het nemen van risico's gemeten. Het instrument bleek onderscheid te maken tussen bestuurders die nog nooit een ongeval hebben gehad en bestuurders die twee of meer ongelukken hebben gehad.

Het vierde meetinstrument dat in het onderzoek naar voren komt is de 'hazard perception test'. De deelnemers kregen een 3,5 minuut durend videofragment te zien vanuit het oogpunt van de bestuurder. In dit videofragment vinden acht (potentieel) gevaarlijke situaties plaats. Een gevaarlijke situatie wordt door McKenna, Horswill en Alexander (2006) gedefinieerd als een situatie waarin de mogelijkheid aanwezig is dat een ongeluk of bijna-ongeluk plaatsvindt. Het videofragment werd voorafgegaan door een schriftelijke instructie en een 1,5 minuut durend oefenfragment met een soortgelijke taak. Telkens wanneer de deelnemer iets waarnam wat mogelijk zou kunnen leiden tot een gevaarlijke situatie moest op de knop worden gedrukt. De reactietijd op de acht gevaren werd gemeten en aan het einde werd de gemiddelde reactietijd berekend. De reactietijd was het aantal seconden vanaf het ontstaan van de gevaarlijke situatie tot de reactie van de deelnemer. Wanneer de deelnemer niet op een gevaar reageerde werd de maximale reactietijd toegekend, namelijk de tijd waarop de gevaarlijke situatie niet meer in beeld was. Dit instrument blijkt onderscheid te maken tussen beginnende autobestuurders, ervaren autobestuurders en expert-autobestuurders. Tevens voorspelt het instrument betrokkenheid bij ongelukken.

De hierboven genoemde instrumenten zijn dus vooral gericht op gevaarherkenning en de reactie daarop. De mate waarin gevaar herkend wordt is een onderdeel van situatiebewustzijn en de reactie

op potentieel gevaar is een gevolg van situatiebewustzijn. Om die reden zijn meetinstrumenten die gevaarherkenning meten interessant als maat voor situatiebewustzijn.

Uit de resultaten van dit onderzoek blijkt dat risicovol rijgedrag te beïnvloeden is door anticipatietraining. Dit impliceert dat ook situatiebewustzijn meetbaar en trainbaar is: het bereiken van minder risicovol rijgedrag vereist een goed bewustzijn van de situatie.

2.3.3 Meetinstrumenten en aannames Endsley (2000)

In deze paragraaf wordt gekeken of de eerder besproken meetinstrumenten voor situatiebewustzijn van Gugerty (1997) en McKenna, Horswill en Alexander (2006) in de aannames passen die Endsley (2000) heeft opgesteld voor metingen van situatiebewustzijn (zie paragraaf 2.2). In Tabel 2.3 wordt samenvattend weergegeven of de onderzoeken van Gugerty en van McKenna, Horswill en Alexander wel of niet voldoen aan de aannames van Endsley.

Tabel 2.3

Aannames Endsley

Aannames Endsley (2000)	Voldoet Onderzoek Gugerty (1997)	Voldoet Onderzoek McKenna et al. (2006)
Een goed instrument bevat herinneringsmetingen en prestatietingen.	Ja, richt zich op zowel herinnerings- als prestatietingen.	Nee, richten zich alleen op prestatietingen.
Situatiebewustzijn kan beter afgeleid worden van prestatie dan van onderliggende processen, zoals een goede informatieverzameling.	Ja, situatiebewustzijn wordt afgeleid van prestatietingen.	Ja, situatiebewustzijn wordt afgeleid van prestatietingen
Kwaliteit situatiebewustzijn hangt van de aandacht, er mag geen sprake zijn van afleiding.	Nee, uitgebreide procedures, maar het aspect 'aandacht' niet expliciet beschreven.	Nee, uitgebreide procedures, af maar het aspect 'aandacht' niet expliciet beschreven.
Werkbelasting (workload) kan invloed hebben op situatiebewustzijn.	Ja, bij hoge taakcomplexiteit verandert de focus.	Nee, taakcomplexiteit niet expliciet onderzocht.
Rapportage van scores moet zo min mogelijk invloed hebben op de situatie, maar tegelijkertijd moet rapportage bij voorkeur meteen plaatsvinden.	Ja.	Ja.

De eerste aanname was dat het concept situatiebewustzijn het bewustzijn van een persoon bevat en niet de processen, zoals informatieverzameling, die ervoor zorgen dat het bewustzijn wordt bereikt. Deze metingen én de metingen van de processen van beslissen en handelen worden gezien als prestatiemetingen. Gugerty (1997) beschrijft dat herinneringsmetingen meer direct meten of iemand zich bewust is van de verkeerssituatie, maar benadrukt dat dit aangevuld moet worden met resultaten van (deel)prestatiemetingen om een gedegen uitspraak te kunnen doen over situatiebewustzijn. Dit is een juiste benadering, omdat louter goede herinnering niet hoeft te leiden tot een goed situatiebewustzijn. Zowel Endsley (2000) als Gugerty (1997) beamen dat een goed meetinstrument beide soorten metingen moet bevatten. McKenna, Horswill en Alexander (2006) richten zich in het besproken onderzoek alleen op prestatiemetingen.

Volgens Endsley (2000) speelt ook werklast een rol bij situatiebewustzijn. Gugerty (1997) stelt dat een hogere werkbelasting (workload) samenhangt met een verandering in focus. Verkeersdeelnemers richten hun aandacht dan vooral op de potentieel directe gevaren: de auto's voor en achter hen. Opvallend daarbij is, volgens Gugerty, dat de focus op de auto's in de dode hoek niet toeneemt. Van het gedrag van de voertuigen direct voor en achter leidt men af hoe de situatie zich ontwikkelt. In het onderzoek van McKenna, Horswill & Alexander (2006) wordt werklast niet expliciet onderzocht.

Ten slotte kan in het onderzoek van zowel Gugerty (1997) als dat van McKenna, Horswill en Alexander (2006) gesproken worden van directe meting van scores: tijdens de getoonde verkeerssituaties worden de handelingen van de deelnemers direct geregistreerd, of direct na een videofragment. In beide onderzoeken wordt daardoor voorkomen dat informatie onvolledig en minder gedetailleerd is, wat het geval is wanneer deelnemers achteraf worden bevraagd. Dit komt overeen met de eerder besproken aanname van Endsley (2000) over metingen van situatiebewustzijn.

Uit de besproken literatuur en onderzoeken, kan geconcludeerd worden dat situatiebewustzijn gemeten kan worden aan de hand van herinneringsmetingen en (deel)prestatiemetingen. De belangrijkste aspecten die in een meetinstrument voor situatiebewustzijn naar voren moeten komen zijn:

- Herinnering: Wat heeft de bestuurder zojuist gezien?
- Gevaarherkenning: Welke gevaren herkent de bestuurder?
- Actieselectie: Welke actie wordt uitgevoerd? (beslissen/handelen)
- Tijdig reageren: Hoe snel wordt tot actie overgegaan om mogelijke (bijna)ongelukken te voorkomen? (beslissen/handelen)

In de volgende paragrafen wordt nader uiteengezet wat de voorwaarden zijn voor een valide en betrouwbare toets.

2.4 Validiteit

Validiteit wordt beschouwd als het belangrijkste kwaliteitskenmerk van assessments (Messick, 1989). Messick (1989) omschrijft validiteit als de mate waarin empirisch bewijs en theorieën de juistheid en gepastheid van interpretaties en acties, gebaseerd op testcores, ondersteunen. Algemeen gesteld is de vraag bij validiteit of daadwerkelijk met het instrument gemeten wordt, wat de onderzoeker wil meten. Het onderbouwen van validiteit is complexer dan het onderbouwen van geschatte

betrouwbaarheid of generaliseerbaarheid. De reden hiervan is dat de schatting van betrouwbaarheid en de schatting van generaliseerbaarheid algoritmisch te onderbouwen zijn en dat een schatting van de validiteit deels berust op menselijk oordelen (Crooks & Kane, 1996). Meestal zijn er meerdere soorten empirisch bewijs nodig om de validiteit van een test te kunnen onderbouwen (Kane, 2006). Volgens Messick (1989) is bewijs altijd incompleet, waardoor validatie in essentie inhoudt dat de meest aanvaardbare onderbouwing voor een validiteitsargument wordt gegeven.

Kane (2006) beschrijft drie belangrijke vormen van validiteit en de daarbij horende methode van validiteitsonderzoek, die in de loop der tijd tot ontwikkeling zijn gekomen. De drie volgens Kane (2006) belangrijkste modellen worden hieronder kort besproken.

Inhoudsvaliditeit

In validiteitsonderzoek volgens het model van inhoudsvaliditeit wordt nagegaan of de testscores van de getoetste onderdelen kunnen worden geïnterpreteerd als zijnde representatief voor het totale toetsdomein van dat construct. Met andere woorden: de toets moet een voldoende grote steekproef van het taakdomein bevatten. Inhoudsvaliditeit wordt als essentieel kwaliteitskenmerk gezien van onderwijstoetsen. Een beperking van validiteitsonderzoek volgens het model van inhoudsvaliditeit is dat de beslissing, of de toetsonderdelen representatief zijn of niet, vaak als subjectief wordt gezien. De ene docent kan meer waarde hechten aan een bepaald onderwerp binnen een kennisdomein dan een andere docent. Wanneer een toets inhoudsvalide is, wordt dat gezien als een basis voor validiteit, waarbij aanvullende bewijzen nodig zijn.

Criteriumvaliditeit

Het criteriummodel houdt in dat de validiteit wordt getoetst door een toets te koppelen aan een bepaald criterium, zoals prestatie in de werkelijkheid of prestatie op andere toetsen die hetzelfde beogen te meten. Voldoet de toets aan dat criterium, dan wordt de toets als valide beschouwd. Twee soorten criteriumvaliditeit zijn te onderscheiden:

- Concurrerende validiteit: hierbij wordt gekeken naar de correlatie tussen de toetsscores en scores die verwante constructen beogen te meten.
- Voorspellende validiteit, gebruikt toekomstige prestaties als criterium: in welke mate voorspelt de toets de prestatie van de deelnemer in de toekomst?

Indien er een duidelijk criterium voorhanden is, is het criteriummodel effectief, aldus Kane. Het geeft een objectief beeld van de validiteit. Een voorwaarde voor criteriumvaliditeit is dus dat een duidelijk criterium aanwezig is.

Constructvaliditeit

Omdat in de sociale wetenschappen constructen worden gemeten die niet direct observeerbaar zijn, voldoen bovenstaande modellen vaak niet bij de validering van toetsing. In het constructmodel wordt uitgegaan van een meer theoretisch begrip van validiteit. Volgens Messick (1989) bevat het model voor constructvaliditeit bijna alle vormen van validiteit (zowel voorspellende en concurrerende, als inhoudsvaliditeit). Constructvaliditeit omvat het bewijs en de theorieën die de geloofwaardigheid van

interpretatie van scores vergroot. Het bewijs bestaat uit concepten die zowel de toetsscore als de relatie met andere variabelen verklaren (Messick, 1989). Het externe aspect van constructvaliditeit houdt dan ook in dat de betekenis van toetsscores wordt verklaard door de mate van samenhang tussen de empirische relaties van de toetsscores met andere maten. Een construct moet bij verschillende meetmethoden dezelfde resultaten opleveren, zodat verschillende methoden of theorieën het construct onderbouwen. Messick (1989) noemt verschillende componenten waaruit constructvaliditeit bestaat, waaronder:

- Trait validiteit: constructen moeten niet gebonden worden aan één meetmethode, maar aan meerdere. De essentie van trait validiteit houdt dan ook in dat een maat een sterkere relatie moet hebben met verschillende methoden die hetzelfde construct toetsen dan met maten van verschillende constructen die met dezelfde methode zijn getoetst.
- Nomologische validiteit: de theorie van het te meten construct dient een basis te zijn voor het afleiden van empirisch te toetsen links tussen toetsscores en de maten van andere constructen. Dit onderbouwt de constructtheorie en de toetsinterpretatie.
- Nomothetic span: relateert toetsscores aan maten van andere constructen en criteria.
- Externe component: relateert toets- of constructscores aan externe variabelen in het algemeen. Dit komt overeen met het bovengenoemde criteriummodel: in dat model wordt ook de validiteit bepaald door vergelijking met een extern criterium.

Drie uitgangspunten (ontwikkelingen) van het model hebben volgens Kane (2006) inmiddels geleid tot een verschuiving van algemene principes van validering.

- Het verschuiven van de nadruk op de validering van de toets aan de hand van een criterium naar de nadruk op ontwikkeling en validering van een voorgestelde interpretatie van toetsscores.
- Het valideren van een toets vindt plaats op basis van uitgebreid onderzoek, in plaats van één empirische studie.
- Er is meer aandacht gekomen voor het evalueren van een score en het overwegen van alternatieve interpretaties.

Argument-based approach

Constructvaliditeit kan volgens Kane het best onderzocht worden door gebruik te maken van een op argumenten gebaseerde benadering, waarmee gestructureerd argumenten (bewijzen) kunnen worden verzameld voor het aantonen van validiteit van een toets. Het begrip validiteit is namelijk zo breed en complex dat er lastig mee te werken valt, waardoor het risico bestaat dat validiteit genegeerd wordt bij beheer van de kwaliteit van toetsing (Crooks & Kane, 1996). Deze *argument-based* benadering kent twee soorten argumenten: interpretatieve argumenten en validiteitsargumenten. Een interpretatief argument maakt aannemelijk hoe op basis van de prestatie op de toets een uitspraak gedaan kan worden over prestaties in het door de toets gerepresenteerde domein. Een validiteitsargument evalueert het interpretatief argument: zijn de aannames en conclusies aanvaardbaar? Normaal gesproken kunnen namelijk niet alle aannames als vanzelfsprekend worden beschouwd (Kane, 2006). Validering van het interpretatief argument vindt plaats tijdens de ontwikkelfase en in de evaluatiefase van de toets. In de ontwikkelfase worden de toets en het interpretatief argument bij de toets ontwikkeld. Gebruik wordt gemaakt van een iteratieve strategie: de aannames en conclusies van het

interpretatief argument worden constant geëvalueerd tijdens de toetsontwikkeling, zodat voor elke zwakte, die zou kunnen duiden op een lagere validiteit, meteen de toets en het interpretatief argument kan worden aangepast.

Tijdens de evaluatiefase wordt de toets beschouwd als een afgerond product. In deze fase wordt uitgebreid en kritisch gekeken of het interpretatief argument daadwerkelijk onderbouwt dat de toets een uitspraak doet over het in de toets gerepresenteerde domein.

In een poging de complexiteit van een uitspraak over validiteit te verminderen hebben Crooks en Kane (1996) een model opgesteld, waarin acht stappen van het toetsingproces worden beschreven, met in elke stap een set bedreigingen voor de validiteit. De acht stappen worden als een keten gezien: als één stap zwak is, is de hele keten zwak. De acht stappen worden in Tabel 2.4 gedefinieerd.

Tabel 2.4

Fasen van toetsingsproces (Crooks & Kane (1996))

Fasen/stappen toetsingsproces	Beschrijving fase:
Afnamefase	Deelnemers krijgen assessmenttaken voorgeschoteld. Duidelijk dient te zijn welke procedures gehanteerd worden bij de assessmenttaak.
Scoringsfase	Prestaties worden uitgedrukt in scores. Voor de validiteit van de toets is het belangrijk dat de scores nauwkeurig in kaart worden gebracht.
Aggregatiefase (combineren)	Individuele scores worden gegroepeerd in subschalen en eindscores. Combinaties van scores worden gemaakt en de juiste weging van scores/assessmenttaken wordt vastgesteld.
Generalisatiefase (domeinscores)	Gekeken wordt naar de representatie van het toetsdomein: Zijn de gebruikte assessmenttaken en condities representatief voor het geheel aan taken dat vereist wordt?
Extrapolatiefase (criterium domeinscores)	Gekeken wordt naar het taakdomein: Zijn de gebruikte assessmenttaken en condities voldoende realistisch? Repreenteert een hoge score een goede prestatie in de realiteit?
Evaluatiefase	Duidelijk wordt gemaakt wat de scores op het taakdomein betekenen: Wat is de waarde van de prestatie van de deelnemer?
Beslissingsfase	Beslist wordt welke acties worden ondernomen naar aanleiding van de beoordeling. Dit kan uiteenlopen van de beslissing om geen feedback te geven tot de beslissing dat een deelnemer niet mag doorgaan met de cursus.
Impact (acties)	Het belang van de impact van de assessmenttaken op de deelnemers.

Per fase kunnen bepaalde beslissingen genomen worden of gebeurtenissen plaatsvinden die de validiteit van de toets bedreigen. Crooks en Kane (1996) noemen per fase bedreigingen voor een goede validiteit. De bedreigingen worden per fase weergegeven in Tabel 2.5.

Tabel 2.5

Bedreigingen validiteit per fase van toetsingsproces

Fasen/stappen toetsingsproces	Bedreigingen validiteit
Afnamefase	- Lage motivatie: wanneer kandidaten niet gemotiveerd zijn kan hun prestatie minder representatief zijn voor hun werkelijke kennis en vaardigheden. - Faalangst: gevolg van hoge motivatie; de angst zorgt voor een prestatie die minder representatief is voor de werkelijk aanwezige kennis en vaardigheden. - Ongepaste toetscondities: de procedures zijn niet goed, waardoor lager gescoord wordt of juist hoger (door bijvoorbeeld ongepaste, directe hulp tijdens de toets). - Taak of antwoord niet juist gecommuniceerd: de taak wordt niet volledig begrepen. De deelnemer is niet in staat zijn antwoord juist te communiceren.
Scoringsfase	- Het scoringsproces haalt de belangrijke kwaliteiten niet uit de prestatie op de toets. - Niet alle factoren worden beoordeeld. - Overdreven nadruk op sommige criteria, vormen of stijlen van antwoorden. - Ontbreken van voldoende inter- of intrabeoordelaarsovereenstemming. - De manier van scoren is te analytisch: wanneer te veel aandacht uitgaat naar verschillende aspecten van de prestatie, kan de rijkheid van de <i>gehele</i> opdracht verloren gaan. - De manier van scoren is te holistisch: er wordt ongepast één cijfer gegeven voor het totaal.
Aggregatiefase (combineren)	- Groepering van scores té divers, zodat er minder goed een uitspraak kan worden gedaan over de totaalscore. - Ongepast gewicht geven aan verschillende aspecten van de prestatie.
Generalisatiefase (domeinscores)	- De condities waaronder de toets gemaakt wordt verschillen, niet gestandaardiseerd. - Inconsistentie in scoringscriteria voor verschillende toetsitems. Dit leidt tot een lage correlatie tussen toetsitems en dat is niet wenselijk voor de generaliseerbaarheid van de toets. Standaardisering van de criteria leidt echter tot een versmalling van het toetsdomein. - Hoe kleiner de steekproef, hoe lager de generaliseerbaarheid van de scores.

Extrapolatiefase (criterium domeinscores	- De assessmenttaken zijn te beperkt: het toetsdomein is dan minder representatief voor het taakdomein. Wanneer bijvoorbeeld alleen meerkeuzevragen gesteld worden en iemand goed scoort, wil dan niet zeggen dat de deelnemer ook goed scoort als dezelfde kennis en vaardigheden worden bevraagd in een ander toetsformat. - Delen van het taakdomein worden niet getoetst of er wordt te weinig gewicht aan gegeven: het toetsdomein is dan minder representatief voor het taakdomein (construct-underrepresentation).
Evaluatiefase	- Verkeerde interpretatie van de toetsresultaten, omdat men niet op de hoogte is van de informatie die de toets geeft en de beperkingen die de toets met zich meebrengt. - Een construct dat inadequaat een vaardigheid toetst en dus eigenlijk een andere vaardigheid meet. - Interpretatie en/of uitleg van de toets kunnen beïnvloed worden doordat de beoordelaar op de hoogte is van eerdere resultaten of andere eigenschappen van de deelnemers.
Beslissingsfase	- Ongepaste standaarden, zijn de juiste standaarden gehanteerd? - De verkeerde didactische beslissingen worden genomen op basis van de toetsresultaten, wat van directe invloed is op de deelnemer.
Impact (acties)	- Positieve gevolgen van de toets worden niet bereikt, zoals plaatsing, bruikbare feedback, e.d. - Serieuze negatieve gevolgen, door bijvoorbeeld oneerlijkheid tijdens de toets of door buitensporig veel stress.

Een kritiek op de benadering van Crooks en Kane (1996) zou kunnen zijn dat om economische reden geen tijd of geld beschikbaar is voor dermate uitgebreid validiteitsonderzoek naar toetsing. Een oplossing hiervoor is dat bepaalde fasen worden onderzocht in plaats van de volledige keten. Dit gaat echter wel ten koste van de volledigheid van validiteitstoetsing.

2.5 Betrouwbaarheid

De eerste fase van validiteitsbepaling in het model van Crooks en Kane, bevat het bepalen van de betrouwbaarheid van een toets. Betrouwbaarheid is dus een onderdeel van validiteit. De betrouwbaarheid geeft aan in welke mate het aannemelijk is dat de werkelijk behaalde score overeenkomt met de gemiddelde score uit een hypothetische verdeling. De hypothetische verdeling bestaat uit de mogelijke scores die behaald zouden worden bij herhaalde metingen van dezelfde toets. Simpler gesteld: wordt met de toets de volgende keer, onder precies dezelfde omstandigheden, hetzelfde resultaat bereikt? Dit komt ook naar voren in het besproken model van Crooks en Kane (1996): zowel in de scoringsfase (inter- en intrabeoordelaarsovereenstemming) en in de generalisatiefase speelt betrouwbaarheid een hoofdrol. Omdat betrouwbaarheid onderdeel uitmaakt van validiteit, bepaalt de mate van betrouwbaarheid volgens Crooks en Kane (1996) de bovenste grens van validiteit. Wanneer de betrouwbaarheid laag is, is de mate van validiteit beperkt.

Wanneer een toets echter in hoge mate betrouwbaar is, hoeft dat niet te betekenen dat de toets ook in hoge mate valide is. In de extrapolatiefase van het model van Kane kan bijvoorbeeld blijken dat delen van het taakdomein onvolledig getoetst zijn. De toets kan dan steeds dezelfde scores opleveren, maar valide is de toets niet.

Met behulp van bijvoorbeeld de klassieke testtheorie en de generaliseerbaarheidtheorie kan de standaard meetfout worden uitgerekend. De standaard meetfout geeft een indicatie voor de nauwkeurigheid waarmee één enkele score de verwachte waarde benadert van een hypothetische verdeling van mogelijke scores van dezelfde toetsdeelnemer (Haertel, 2006). Kortom: wat is de standaardafwijking van de daadwerkelijk behaalde score ten opzichte van het hypothetische gemiddelde? Hoe kleiner de afwijking, hoe betrouwbaarder de toets. Ter aanvulling op de standaard meetfout wordt vaak de precisie van testcores uitgedrukt in een betrouwbaarheidscoëfficiënt. De betrouwbaarheidscoëfficiënt wordt, in tegenstelling tot de standaard meetfout, niet gedefinieerd in één enkele score, maar voor alle toetsscores uit een steekproef van de toetspopulatie. De coëfficiënt geeft weer hoe hoog de correlatie is tussen de ware scores en de scores van de deelnemers. De betrouwbaarheidscoëfficiënt kent waarden tussen 0 en 1, waarbij 0 geen betrouwbaarheid geeft en 1 maximale betrouwbaarheid. Om te testen of de Toets Situatiebewustzijn betrouwbaar is, zal de standaard meetfout en de betrouwbaarheidscoëfficiënt worden berekend.

Nu duidelijk is wat situatiebewustzijn is, hoe het gemeten kan worden, wat de kenmerken van een valide toets zijn en hoe betrouwbaarheid kan worden berekend kunnen de onderzoeksvragen geformuleerd worden.

2.6 Onderzoeksvragen en verwachtingen

Aan de hand van de eerder besproken validiteitsketen van Crooks en Kane zijn de onderzoeksvragen opgesteld. Aangezien het onderzoek moet uitwijzen in welke mate een toets die beoogd situatiebewustzijn te meten valide en betrouwbaar is, zal de volgende onderzoeksvraag centraal staan in deze scriptie:

In hoeverre meet de ontwikkelde toets op een betrouwbare en valide wijze het construct situatiebewustzijn bij automobilisten?

Om te onderzoeken in welke mate de Toets Situatiebewustzijn valide en betrouwbaar het construct situatiebewustzijn meet, worden de volgende specifieke vraagstellingen onderzocht:

1. In hoeverre worden bij de Toets Situatiebewustzijn passende afdnamecondities geschapen, die de kandidaten gelegenheid geven hun kunnen te tonen?
(komt aan bod in de afdamefase van het validiteitsmodel van Crooks & Kane, 1996)
2. Hoe groot is de interne consistentie van de Toets Situatiebewustzijn?
(aspect uit de generalisatiefase van het validiteitsmodel van Crooks & Kane, 1996)

3. In hoeverre is het gekozen antwoordmodel bij de itemtypes passend bij het beoogde taakgedrag in de opgaven?
(vergelijk met de scoringsfase van het validiteitsmodel van Crooks & Kane, 1996)
4. In hoeverre hangen scores op de Toets Situatiebewustzijn samen met maten voor rijvaardigheid in het algemeen en relevante externe criteria, zoals ongevalbetrokkenheid?
(vergelijk met de extrapolatiefase van het model van Crooks & Kane, 1996)
5. In hoeverre hangt de mate van situatiebewustzijn samen met persoonlijke achtergrondvariabelen, die uit de literatuur als relevant naar voren komen?
(komt aan bod in de extrapolatiefase van het model van Crooks & Kane, 1996)

Onderzoek naar de samenhang van de toets met een authentiek criterium, namelijk het feitelijk rijgedrag, was voorzien, maar kon om praktische redenen niet gerealiseerd worden. In verband met vertraging in de oplevering van de toets kon niet tijdig data worden verzameld over de feitelijke rijvaardigheid. Als alternatief is onderzoek gedaan naar de samenhang tussen de toets en een zelfevaluatie voor rijvaardigheid. Verwacht wordt dat situatiebewustzijn samenhangt met het feitelijk gerapporteerde rijgedrag.

3. Methode

Het samenwerkingsverband CitoDrive ontwikkelt instrumenten voor het meten van aspecten van rijvaardigheid van autobestuurders in het verkeer. CitoDrive is bezig met de ontwikkeling van een webportal (CitoDrive Portal) waar de instrumenten op worden aangeboden. Het doel van deze instrumenten is het toetsen of verkeersdeelnemers de kritische competenties beheersen die nodig zijn voor veilige en verantwoorde deelname aan het verkeer.

Zoals eerder beschreven werden in de periode van januari 2007 tot juni 2008 vier instrumenten ontwikkeld: een instrument voor zelfevaluatie van rijvaardigheid, een Assessment Bestuurdersrisico, een Performance Assessment Rijvaardigheid (PAR) en de Toets Situatiebewustzijn. Voordat deze instrumenten definitief aangeboden worden op de markt, zal voor elk besproken instrument een vooronderzoek moeten worden uitgevoerd naar de betrouwbaarheid en validiteit. Na dit vooronderzoek kunnen aanpassingen worden verricht en worden de instrumenten op het webportal aangeboden. In dit vooronderzoek wordt de Toets Situatiebewustzijn onderzocht.

Aanvankelijk was het de bedoeling om de toets te valideren aan de hand van resultaten op het Performance Assessment Rijvaardigheid (PAR). In verband met vertraging in de oplevering van de Toets Situatiebewustzijn kon niet tijdig data worden verzameld over de feitelijke rijvaardigheid. Wel kon gebruik gemaakt worden van proefgegevens, verzameld in een onderzoek onder medewerkers van Cito.

3.1 Subjecten

In dit onderzoek is de Toets Situatiebewustzijn gemaakt door 81 medewerkers van Cito. De deelnemers van de toets variëren in de leeftijd van 23 tot 61 jaar, met een gemiddelde leeftijd van 40 jaar (SD = 10,8 jaar). Van de respondenten was 54 procent man en 46 procent vrouw. Als bestuurders rijden ze per jaar gemiddeld 12.734 km. Dit loopt uiteen van 0 km per jaar tot 58.000 km per jaar. De periode waarin deelnemers rijbewijs B hebben behaald loopt van 1965 tot 2003. Het hoogst afgeronde opleidingsniveau van de deelnemers varieert van basisschool tot universitaire opleiding. In Tabel 3.1 worden frequenties en percentages van de verschillende opleidingsniveaus weergegeven (zie tevens bijlage A).

Tabel 3.1
Opleidingsniveau

Opleiding	Frequentie	Percentage
Basisschool	1	1,2
VMBO (mavo/vbo)	2	2,5
Havo	7	8,6
VWO	5	6,2
MBO	13	16,0
HBO	22	27,2
WO	30	37,0
Overig	1	1,2
Totaal	81	100,0

3.2 Instrumenten

In deze paragraaf worden de gehanteerde instrumenten besproken.

3.2.1 Toets Sitatiebewustzijn

Met de Toets Sitatiebewustzijn wordt beoogd het situatiebewustzijn van een autobestuurder te meten. Het is een digitale toets die via Internet op elke computer met Windows kan worden afgenomen.

De toets laat een veertigtal videofragmenten zien van verkeerssituaties. In de constructiefase van de toets zijn door verkeersdeskundigen scenario's ontworpen van verkeerssituaties die aan bod zouden moeten komen in de videofragmenten. Met behulp van twee camera's in de auto zijn de videofragmenten opgenomen (zie bijlage B voor een draaiboek van de opnames). Hierbij is vooraf een domein van verkeersscenario's onderscheiden waarin situatiebewustzijn een rol speelt bij de uitvoering van verkeerstaken. Uitgangspunt van het domein was het gebied waarin een automobilist zich bevindt. Elk gebied bevat een aantal kenmerkende verkeerstaken van verschillende complexiteit. Uit het ruwe videomateriaal zijn fragmenten gesneden die elk een kritisch verkeersscenario vertegenwoordigen. Bij 59 scenario's zijn opgaven geconstrueerd van vier typen, die elke een aspect van situatiebewustzijn beogen te meten. De opgaventypen worden hieronder besproken. Na montage zijn de fragmenten verschillende malen bekeken op bruikbaarheid, opnieuw bijgesneden en aangepast zodat kritieke factoren in het scenario duidelijker te zien waren. Daarna zijn bijbehorende antwoorden geformuleerd.

De gemonteerde videofragmenten tonen de situatie vanuit het oogpunt van de bestuurder, met in de rechterbovenhoek een ingekaderde opname van de situatie achter de bestuurder. Dit kader fungeert als achteruitkijkspiegel (zie Figuur 3.1).



Figuur 3.1 Screenshot Toets Sitatiebewustzijn

Instructie

De toets wordt voorafgegaan door een algemene instructie. Voor elk type opgave wordt een voorbeelditem aangeboden. In totaal zijn er vijf voorbeelditems: het type 'recollection' kent twee verschillende opgaven, dus ook twee voorbeeldopgaven. Elke voorbeeldopgave wordt voorafgegaan door een aparte instructietekst. Na de algemene instructie volgen vier blokken met 10 opgaven, elk voorafgegaan door een korte instructie.

De opgaven vallen onder vier aspecten van situatiebewustzijn, namelijk 'recollection', 'timed response', 'action selection' en 'hazard perception'. Elk genoemd aspect heeft een ander type vraagvorm. Deze vraagvormen worden hieronder nader besproken.

Opgaventype 'recollection'

Het opgaventype 'recollection' bevat herinneringsvragen: worden risicofactoren herinnerd die van direct belang kunnen zijn voor de verkeerssituatie?

Eerst wordt een videofragment getoond. Door te klikken op een 'play'-knop start het videofragment. De duur van deze videofragmenten loopt uiteen van 10 tot 21 seconden. Wanneer het fragment is afgelopen blijft de eindsituatie in beeld en wordt de vraag gesteld. De set opgaven van het type 'recollection' kent twee soorten vragen. Het eerste type meet het opmerken van een teken of bord door naar de consequenties te vragen (zie Figuur 3.3). Het tweede type bevraagt of een te maken gedragskeuze veilig kan worden genomen in een complexe verkeerssituatie (Figuur 3.4).

De antwoordmogelijkheden zijn geformuleerd in meerkeuzevorm. Er is geen tijdslimiet voor het beantwoorden van de vragen, het videofragment wordt één keer getoond.

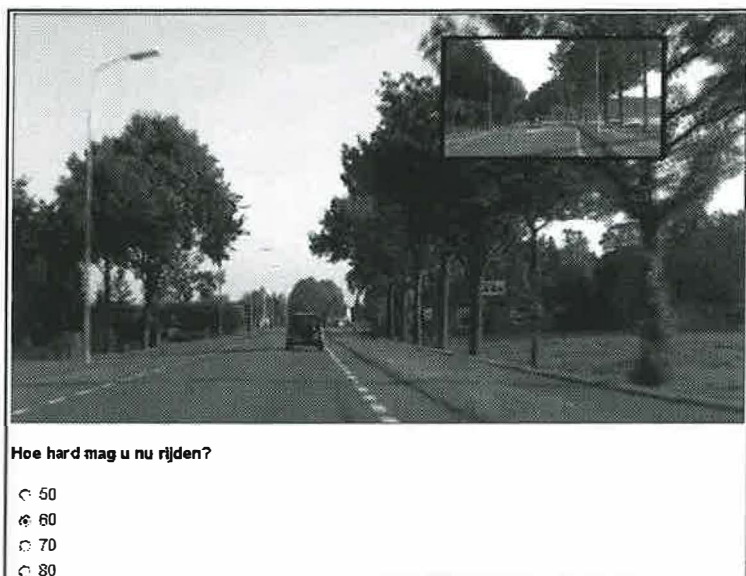
Voorbeeld 1:

In het videofragment rijdt u op een doorgaande weg (zie Figuur 3.2).



Figuur 3.2 Een verkeersbord wordt gepasseerd

Het beeld stopt en vervolgens komt de vraag in beeld (zie Figuur 3.3).



Figuur 3.3 Het beeld stopt en de opgave komt in beeld

De deelnemer klikt op een antwoordalternatief en gaat verder naar de volgende opgave door zelf op een navigatiepijl onderin beeld te klikken.

Voorbeeld 2:

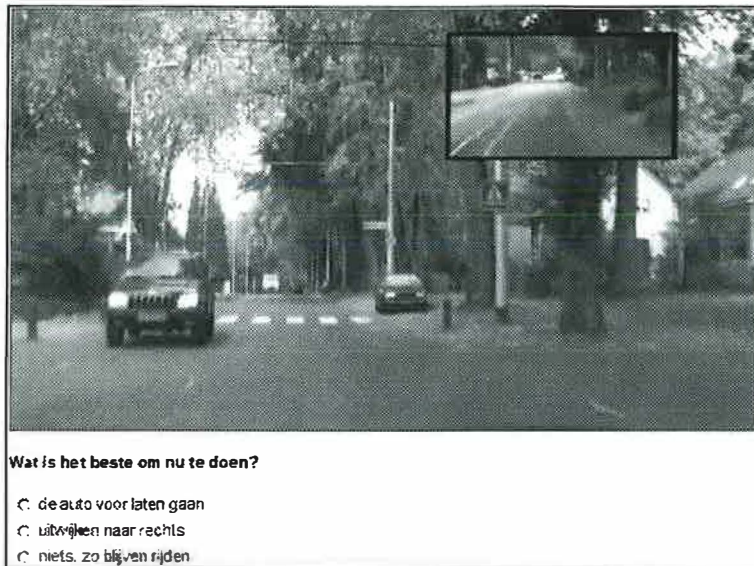
In Figuur 3.4 wordt ter illustratie een voorbeeld gegeven van de tweede vraagsoort bij het opgaventype 'recollection'.



Figuur 3.4 Het beeld is gestopt en de opgave staat in beeld

Opgaventype 'action selection'

Het opgaventype 'action selection' bevat opgaven waarbij de keuze dient te worden gemaakt voor de beste gevaarafwendende handeling op dat moment. De deelnemer klikt op de 'play'-knop. Een videofragment wordt getoond. De duur van de videofragmenten loopt uiteen van 5 tot 19 seconden. Het videofragment stopt en het laatste beeld blijft zichtbaar. Vervolgens wordt de vraag gesteld. De antwoordmogelijkheden zijn geformuleerd in meerkeuzevorm (zie Figuur 3.5).



Figuur 3.5 Het beeld is gestopt en de opgave komt in beeld

Opgaventype 'hazard perception'

Het opgaventype 'hazard perception' bevat opgaven waarbij het herkennen van (mogelijke) gevaren in de verkeerssituatie centraal staat.

De deelnemer klikt op de 'play'-knop en een videofragment wordt getoond. De duur van de videofragmenten loopt uiteen van 6 tot 12 seconden. Het fragment bevriest op het laatste beeld van het videofragment. Vervolgens wordt de vraag gesteld en komt een rood kruisje onderin beeld. De kandidaat beantwoordt de vraag door te klikken op een plaats in het bevroren videofragment. Vervolgens verschijnt het rode kruisje op die plaats, zie Figuur 3.6. Dit opgaventype staat bekend als 'hotspot'.



Figuur 3.6 Het filmpje is gestopt en het kruisje komt in beeld

Het is juist wanneer de deelnemer het kruisje plaatst op de man achter de bus die wil oversteken.

Opgaventype 'timed response'

In het opgaventype 'timed response' staat het *op het juiste moment* kiezen van een gevaarafwendende handelingsmanoeuvre centraal.

Eerst wordt een vraag gesteld. Bijvoorbeeld:

- *Wat is het beste moment om te remmen? Klik op dat moment op het filmpje.*
- *Wat is het beste moment om naar de linkerstrook te gaan om in te halen? Klik op dit moment op het filmpje.*

Wanneer de deelnemer op de 'play'-knop klikt wordt een videofragment getoond. De duur van deze videofragmenten loopt uiteen van 6 tot 25 seconden. De deelnemer antwoordt door éénmaal te klikken tijdens het afspelen van het videofragment en geeft daarmee zijn moment van handelen weer. Het videofragment blijft op het moment van klikken stilstaan (zie ter illustratie Figuur 3.7). In het betreffende videofragment rijdt de bestuurder door een bocht en ziet plots een vrachtwagen geparkeerd staan.



Figuur 3.7 Opgave 'timed response'

Bij dit type opgaven wordt de opgave dus beantwoord tijdens het videofragment. Er staat geen tijdsindicatie in beeld. Nadat het videofragment is afgelopen en de deelnemer nog niet geklikt heeft, is de opgave fout beantwoord. De deelnemer moet dan verder naar de volgende opgave door zelf op de pijl onderin beeld te klikken.

Voor alle typen opgaven, uitgezonderd het opgaventype 'timed response', geldt dat maar één goed antwoord mogelijk is. Elk blok bevat 10 opgaven, per blok kan dus 100 procent gescoord worden. Voor het opgaventype 'timed response' is een kritiek tijdsvenster gedefinieerd waarbinnen een antwoord goed gerekend wordt. De deelnemer ontvangt voor elk blok een percentage goed gescoord.

3.2.2 Zelfevaluatie Rijgedrag

De Zelfevaluatie Rijgedrag is een zelfbeoordeling van de eigen rijcompetentie, aan de hand van stellingen die zijn afgeleid uit criteria voor rijcompetentie, namelijk veiligheid, sociaal rijgedrag,

doorstromen, milieuverantwoord rijden en gecontroleerd rijden. Hieronder worden de criteria beschreven, zoals in de handleiding van CitoDrive.

Het criterium 'veiligheid' heeft te maken met de eigen objectieve snelheid, maar ook met de snelheid en afstand ten opzichte van anderen. De bestuurder moet zodanig rijden dat hij voldoende veiligheidsmarges heeft ingebouwd. Dit betekent dat de bestuurder de rit zodanig plant dat hij met een veilige snelheid op tijd de bestemming kan bereiken. Tevens geldt dat de bestuurder continu met een aangepaste snelheid rijdt: aangepast aan omstandigheden/taaksituaties maar ook aangepast aan eigen kunnen. Hij houdt voldoende afstand tot de voorganger. Hij schat risico's juist in, herkent gevaren op tijd en kiest ervoor om daarnaar te handelen.

Het criterium 'sociaal rijgedrag' betekent dat de bestuurder in het verkeer niet alleen oog heeft voor zijn eigen veiligheid en doorstroming, maar ook voor de belangen van anderen. Dit betekent dat hij zich sociaal gedraagt, dat hij niet vasthoudt aan eigen rechten en dat hij fouten van anderen door de vingers ziet. Hij voorkomt dat hij anderen verrast door zijn bedoelingen ruim van te voren duidelijk te maken en geeft anderen de ruimte om gemaakte fouten te herstellen. Hij kan zich in andere weggebruikers verplaatsen en kan inschatten of zijn eigen gedrag tot irritatie of nervositeit leidt bij anderen. Hij reageert niet op agressie van andere weggebruikers.

Het derde criterium is 'doorstromen'. Voor het strategisch niveau betekent dit dat de bestuurder wordt geacht besluitvaardig te rijden en vlot beslissingen te nemen, zodanig dat de doorstroming van de overige weggebruikers niet in gevaar komt. Dit betekent eventueel buiten de spits vertrekken of een andere route te kiezen in de spits. Op het tactische niveau dienen de verkeerstaken veilig maar ook doortastend en vlot te worden uitgevoerd. Bij het stoppen dient bijvoorbeeld rekening te worden gehouden met de plaats van stoppen: die wordt zodanig gekozen dat dit de minste vertraging oplevert voor ander verkeer. Hetzelfde geldt voor afslaan en, op het operationele niveau, het uitvoeren van manoeuvres. Manoeuvres worden met lage snelheid uitgevoerd, maar niet zo traag dat zij hinder opleveren voor andere weggebruikers.

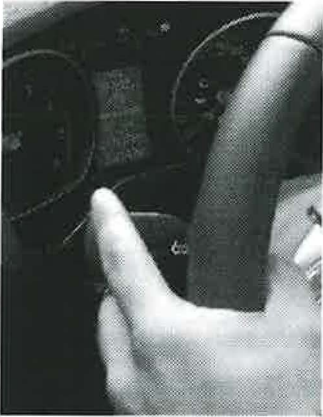
Het vierde criterium is 'milieuverantwoord rijden'. Op het strategische niveau kan al rekening worden gehouden met het milieu, bijvoorbeeld door routes die al zwaar belast zijn te mijden. Ook het rijden met de juiste bandenspanning hoort bij het strategische niveau. Op het tactische niveau gaat het om het vermijden van grote snelheidsverschillen: rustig optrekken en afremmen, tijdig opschakelen naar een hogere versnelling en gebruik maken van het zelfrollend vermogen van de auto. Wanneer langere tijd wordt stilgestaan wordt de motor afgezet. Op het operationele niveau kan gedacht worden aan het juiste gebruik van gas en koppeling in relatie tot de lage snelheid waarmee manoeuvres worden uitgevoerd.

Het laatste criterium waar stellingen over worden voorgelegd is 'controle'. Hierbij gaat het om de technische (voertuig)beheersing en bediening. Op het strategische niveau kan dan gedacht worden aan het gemakkelijk kunnen omgaan met navigatiesystemen, het kunnen plaatsen van sneeuwkettingen etc. Op het tactische niveau gaat het meer om de manier waarop de verkeerstaken technisch worden uitgevoerd: wordt er gecontroleerd en vloeiend gestuurd, gaat het afslaan of het stoppen en weggrijden soepel en zonder horten en stoten. Ook op het operationele niveau moeten alle

manoeuvres soepel en beheerst uitgevoerd kunnen worden. Het sturen, gas geven, ontkoppelen, koppelen en remmen moet op geautomatiseerd niveau plaatsvinden.

De Zelfevaluatie Rijgedrag is bedoeld voor ervaren (zakelijke) bestuurders. Met deze zelfevaluatie beoordelen bestuurders indirect het eigen rijgedrag. Het is een digitale toets. In deze toets worden 50 stellingen over autorijden in specifieke verkeerssituaties voorgelegd. De kandidaten geven aan in welke mate ze een voorgesteld gedrag vertonen. Hierbij antwoorden ze op een schaal lopend van (vrijwel) nooit tot (vrijwel) altijd. Net als de Toets Situatiewaardering wordt de toets via internet afgenomen, dus op elke gewenste computer kan de zelfevaluatie gemaakt worden (zie ter illustratie Figuur 3.8 en 3.9).

CitoDrive - Zelfevaluatie Vraag 3 van 45



Geeft u richting aan als u inhaalt?

Vrijwel altijd ☒ ☐ ☐ ☐ ☐ Vrijwel nooit

Figuur 3.8 Voorbeeld van een opgave uit de zelfevaluatie (de opgavetekst is in dit figuur vergroot ten behoeve van de leesbaarheid)

CitoDrive - Zelfevaluatie Vraag 8 van 45

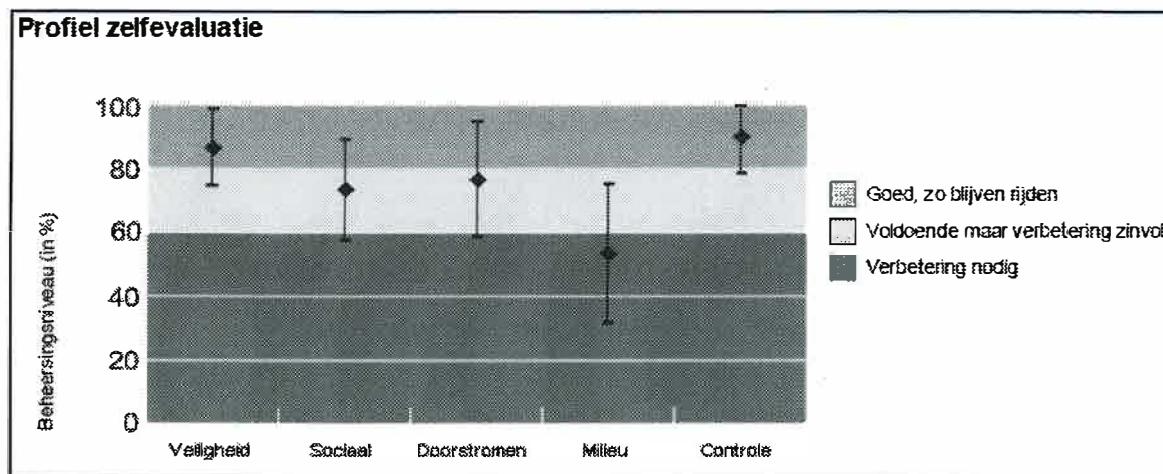


Rijdt u, als u wilt invoegen bij druk verkeer, zo lang mogelijk door op de vluchtstrook?

Vrijwel altijd ☐ ☒ ☐ ☐ ☐ Vrijwel nooit

Figuur 3.9 Voorbeeld van een opgave uit de zelfevaluatie (de opgavetekst is in dit figuur vergroot ten behoeve van de leesbaarheid)

In de portal van CitoDrive wordt een scoreraportage voor de kandidaat gegenereerd. Aan het einde van de zelfevaluatie krijgt de kandidaat dan ook een individuele rapportage met daarin de scores en een toelichting daarop. De deelnemer krijgt scores voor veilig rijden, sociaal rijden, doorstroming bevorderend rijden, milieubewust en energiezuinig rijden, en gecontroleerd rijden (zie Figuur 3.10). Daarnaast krijgt de deelnemer een toelichting op de resultaten.



Figuur 3.10 Voorbeeld van een scoreraportage zelfevaluatie

3.2.3 Vragenlijst over de Toets Situatiewustzijn

Om zicht te krijgen op de achtergrondkenmerken van de respondenten, de afnamecondities en op ongevalbetrokkenheid is een vragenlijst afgenomen, via een internetsurvey. De resultaten van deze evaluatielijsten zullen worden gebruikt om de afnameomstandigheden te onderzoeken: werden de condities als gunstig ervaren? Daarnaast dient aan de hand van de vragenlijsten duidelijk te worden of de vormgeving van de toets werkbaar is en of de toetsscores samenhangen met persoonsgebonden factoren. Gevraagd wordt naar geslacht, opleiding, slechtziendheid, kleurenblindheid en ervaring met (spel)computers. Daarnaast wordt gevraagd hoe lang men in het bezit is van een rijbewijs en hoeveel kilometer per jaar gereden wordt. Tevens wordt gevraagd naar ongevalbetrokkenheid. Ten slotte worden enkele stellingen voorgelegd met betrekking tot motivatie. De vragenlijst is opgenomen in bijlage C.

3.3 Design en dataverzameling

De Toets Situatiewustzijn is doorlopen door 81 deelnemers, zoals beschreven in paragraaf 3.1. De deelnemers zijn uitgenodigd op basis van hun eerdere deelname aan de afname van de Zelfevaluatie Rijgedrag. De deelnemers aan de toetsafnames zijn uitgenodigd per email. In verband met de capaciteit van de testserver zijn dagelijks maximaal 20 deelnemers uitgenodigd om aan de toets mee te doen. In de email werd een toelichting over het project CitoDrive gegeven en werd uitgelegd wat het belang van de toets was in het kader van dit project. Tevens werd een link opgenomen naar de vragenlijst over de toets. De deelnemers kregen binnen een week per email de uitslag toegestuurd. De deelnemers die de toets nog niet gemaakt hadden kregen nogmaals een uitnodiging.

3.4 Analyseplan

In tabel 3.2 wordt het analyseplan weergegeven. Per onderzoeksvraag wordt de methode van dataverzameling weergegeven. De data wordt verzameld aan de hand van de vragenlijst, de Toets Situatiewustzijn en de Zelfevaluatie Rijgedrag. Daarnaast worden de variabelen genoemd die gebruikt zullen worden voor de analyse, zoals scores van deelnemers en de stellingen die in de vragenlijsten worden bevraagd. Ten slotte wordt de analysemethode genoemd.

Tabel 3.2
Analyseplan

Onderzoeksvraag	Methode dataverzameling	Variabelen	Analysemethode
1. In hoeverre worden bij de Toets Situatiewustzijn passende afnamecondities geschapen, die de kandidaten gelegenheid geven hun kunnen te tonen? (afnamefase Crooks & Kane, 1996)	Vragenlijst	Stellingen: - Het doel van de toets was duidelijk. - De geschreven instructie voorafgaan aan de toets was helder. - Bij elke vraag was het duidelijk wat ik moest doen. over de toets kunt u ze in - De filmpjes waren duidelijk. onderstaand testvak noteren.' - De indeling van het beeldscherm was overzichtelijk. - Ik had voldoende tijd om elke vraag in te vullen. Overige opmerkingen.	Beschrijvende statistieken en inhoudsanalyse van de antwoorden op de open vraag 'Als u nog opmerkingen heeft
2. Hoe groot is de interne consistentie van de Toets Situatiewustzijn? (generalisatiefase Crooks & Kane, 1996)	Toets Situatiewustzijn	Scores deelnemers op de afzonderlijke items.	Toets- en itemanalyses (berekening p-waarden, item restcorrelaties, alpha-coëfficiënten en standaard meetfouten.
3. In hoeverre is het gekozen antwoordmodel bij de itemtypes passend bij het beoogde taakgedrag in de opgaven? (scoringfase Crooks & Kane, 1996)	Toets Situatiewustzijn	Itemscores op de vier itemtypen.	Analyses van de attractiewaarden van de itemtypen 'recollection' en 'action selection' en analyses van de responsietijden bij itemtype 'timed response'.

4. In hoeverre hangen scores op de Toets Situatiebewustzijn samen met maten voor rijvaardigheid in het algemeen en relevante externe criteria, zoals ongevalbetrokkenheid? (extrapolatiefase Crooks & Kane, 1996)	Toets Situatiebewustzijn Vragenlijst Zelfevaluatie Rijgedrag	Score Toets Situatiebewustzijn. Score Zelfevaluatie Rijgedrag. Externe Criteria (vragenlijst): - Bent u, in de afgelopen vier jaar, Betrokken geweest bij een auto-ongeluk of schadeclaim?	Beschrijvende statistieken en pearson correlatie tussen scores en externe criteria.
5. In hoeverre hangt de mate van situatiebewustzijn samen met persoonlijke achtergrondvariabelen, die uit de literatuur als relevant naar voren komen? (extrapolatiefase Crooks & Kane, 1996)	Toets Situatiebewustzijn Vragenlijst	Score Toets Situatiebewustzijn Persoonlijke achtergrondvariabelen (vragenlijst): - Wat is uw leeftijd? - Geslacht - Wat is uw hoogst afgerond opleiding? - Ik zie slecht. - Ik ben kleurenblind. - Speelt u spelletjes op de spelcomputer? - Wanneer heeft u uw rijbewijs B behaald? - Hoeveel kilometer rijdt u per jaar? - Ik vond het leuk om de toets te maken. - Ik vond de toets leerzaam. - Ik kon me goed concentreren tijdens de toets.	Beschrijvende statistieken en pearson correlatie tussen scores en persoonlijke achtergrondvariabelen.

4. Resultaten

In dit hoofdstuk worden de resultaten van het onderzoek gerapporteerd. In elke paragraaf worden per onderzoeksvraag de resultaten besproken.

4.1 Betrouwbaarheid van de Toets Situatiebewustzijn en haar subschalen

In deze paragraaf worden de resultaten van toets- en itemanalyses van de Toets Situatiebewustzijn besproken. Uit toets- en itemanalyses van TIA-plus en SPSS blijkt de toets aanvankelijk een betrouwbaarheid te hebben van .502 (Cronbach's Alpha). Dit is een lage betrouwbaarheid. Gekeken naar de afzonderlijke opgaventypen is de itemset van het type 'recollection' het meest betrouwbaar, met een alpha van .620. De betrouwbaarheidscoëfficiënten van de subschalen worden weergegeven in Tabel 4.1.

Tabel 4.1

Betrouwbaarheid Subschalen

Subschaal	Cronbach's Alpha
Recollection	.620
Action Selection	.300
Hazard Perception	.441
Timed Response	-.001

Na het verwijderen van een tiental types met een lage item restcorrelatie loopt de betrouwbaarheidscoëfficiënt van de Toets Situatiebewustzijn op tot .65 (zie bijlage D voor een overzicht van de resterende 30 opgaven).

In Tabel 4.2 worden per opgave de P-waarde, Rir-waarde en Cronbach's Alpha van de toets na verwijdering van het betreffende item weergegeven.

Tabel 4.2

Statistieken Per Opgave

Itemnummer	P-waarde	Rir-waarde	α wanneer dit item verwijderd is
1	.88	.23	.49
2	.97	.25	.49
3	.89	.13	.50
4	.47	.03	.51
5	.60	.24	.48
6	.88	.11	.50
7	.64	.23	.50
8	.87	.36	.47
9	.96	.33	.49
10	.92	.28	.48
11	.56	.30	.47
12	.33	-.06	.52
13	.73	.03	.51
14	.72	-.05	.52
15	.49	-.05	.52
16	.97	.27	.49
17	.89	.18	.49
18	.37	-.003	.51
19	.45	.16	.49
20	.84	.22	.48
21	.89	.10	.50
22	.89	.003	.51
23	.55	.10	.50
24	.84	.15	.50
25	.85	.34	.50
26	.95	.20	.49
27	.88	.21	.49
28	.88	.13	.50
29	.77	.41	.46
30	.99	.14	.50
31	.64	.05	.51
32	.63	-.08	.52
33	.60	.11	.50
34	.64	-.001	.51
35	.89	.20	.49
36	.31	.01	.51
37	.49	.15	.49
38	.40	-.11	.53
39	.52	.06	.50
40	.52	.18	.49

Het grootste deel van de items in de categorie 'recollection' heeft een voldoende hoge item restcorrelatie. Tevens bevatten zeven van de tien opgaven hoge p-waarden en de subschaal heeft een Cronbach's Alpha van .620. Opvallendheden worden in paragraaf 4.4.1 besproken.

De subschaal 'action selection' heeft een Cronbach's Alpha van .300. De opgaven met lage item restcorrelaties (lager dan 0.1) worden besproken in paragraaf 4.4.2.

De subschaal 'hazard perception' heeft een Cronbach's Alpha van .441 en kent één opgave met een item restcorrelatie lager dan 0.1. Deze opgave wordt besproken in paragraaf 4.4.3. Uit Tabel 4.2 is af te leiden dat de meeste opgaven van het type 'hazard perception' door een ruime meerderheid juist worden beantwoord. Ten slotte is in Tabel 4.2 te zien dat de item restcorrelaties van de opgaven uit de subschaal 'timed response' veelal laag of negatief zijn. De subschaal heeft een Cronbach's Alpha van -.001. In de huidige vorm bezitten de opgaven onvoldoende psychometrische kwaliteit. Het opgaventype 'timed response' wordt uitgebreid besproken in paragraaf 4.4.4.

De correlatiecoëfficiënt van de score op 'recollection' en de totaalscore op de toets hangen significant positief samen ($r = .76$, $p < .01$). Deze subschaal draagt substantieel bij aan de totaalscore. De samenhang tussen de vier subschalen wordt weergegeven in Tabel 4.3.

Tabel 4.3

Samenhang Totaalscores Per Subschaal en Algemene Totaalscore

Subschaal	1	2	3	4	Totaalscore
1. Recollection	1	.52(**) $p < .01$	.38(**) $p < .01$	.23(*) $p < .05$	.76(**) $p < .01$
	82	82	75	75	75
2. Action Selection	.52(**) $p < .01$	1	.25(*) $p < .05$	.18 $p < .05$	.65(**) $p < .01$
	82	82	75	75	75
3. Hazard Perception	.38(**) $p < .01$	.25(*) $p < .05$	1	.23(*) $p < .05$	.71(**) $p < .01$
	75	75	75	75	75
4. Timed Response	.23(*) $p < .05$	.18 $p < .05$	.23(*) $p < .05$	1	.59(**) $p < .01$
	75	75	75	75	75
Totaalscore	.76(**) $p < .01$	.65(**) $p < .01$	.71(**) $p < .01$	.59(**) $p < .01$	1
	75	75	75	75	75

Noot

* Correlatie is significant $\alpha = 0.05$ (tweezijdig).

** Correlatie is significant bij $\alpha = 0.01$ (tweezijdig).

In Tabel 4.3 is te zien dat vrijwel alle subschalen significante samenhang vertonen met de overige subschalen. Daarop is één uitzondering: de subschalen 'action selection' en 'timed response' hangen niet significant samen. Alle subschalen hebben een significante samenhang met de totaalscore op de Toets Situatiebewustzijn. Verwacht werd dat de subschalen samenhangen, omdat ze samen beogen één construct te meten.

4.2 Realisatie van de afnamecondities

Met behulp van de antwoorden op de digitale vragenlijst werd gekeken in welke mate de waardering van de afnamecondities samenhangt met de toetsscores. De resultaten zijn samengevat in Tabel 4.4.

Tabel 4.4

Resultaten Waardering Afnamecondities

Stelling*	Gem.	SD	Samenhang met RC score	Samenhang met totaalscore
Het doel van de toets was duidelijk.	2.4	0.9	.05	-.09
De geschreven instructie voorafgaand aan de toets was helder.	2.5	0.8	.17	.17
Bij elke vraag was duidelijk wat ik moest doen.	2.2	1.0	.24 (p < .05)	.17
De filmpjes waren duidelijk.	1.5	0.9	.14	-.06
De indeling van het beeldscherm was overzichtelijk.	1.8	0.9	.16	.23
Ik had voldoende tijd om elke vraag in te vullen.	2.5	0.9	.25 (p < .05)	.22
Speelt u spelletjes op de spelcomputer**	0.5	0.7	.11	.13
Ik vond het leuk om de toets te maken.	2.5	0.9	.10	.09
Ik vond de toets leerzaam.	1.7	0.4	-.17	-.11
Ik kon me goed concentreren tijdens de toets.	1.9	0.9	.28(p < .05)	.23
Ik was onzeker om goed te presteren.	2.2	0.9	-.14	-.02

Noot

* 0 = helemaal oneens, 3 = helemaal eens

** 0 = vrijwel nooit, 1 = ja, maar niet vaak, 2 = ja, ik ben inmiddels een ervaren speler

Doel van de toets

Uit de antwoorden blijkt dat 85,5 procent van de deelnemers het doel van de toets duidelijk vond. De gemiddelde waardering is 2.4 op een schaal van 0 tot 3. Daarnaast blijkt er geen significante samenhang te zijn tussen het oordeel over het doel en de totale toetsscore ($r = -.09$, $p = .48$).

Instructie toets

Uit de antwoorden op de vragenlijst blijkt dat 88,1 procent van de deelnemers de instructie vooraf duidelijk vond (gemiddelde waardering van 2.5). Er is geen significante samenhang tussen het oordeel over de instructie en de totale toetsscore ($r = .17$, $p = .18$).

Een aantal deelnemers vond de instructie minder helder. Acht deelnemers hebben opmerkingen gemaakt over de instructie. Gesteld werd dat er te veel instructietekst was, de instructie onduidelijk was, de tekst te klein was, er meer oefenopgaven hadden moeten zijn en dat er feedback had moeten worden gegeven na beantwoording van de oefenopgaven, zodat duidelijk zou worden of de opgave juist of onjuist werd beantwoord en waarom dit het geval was.

Duidelijkheid vragen

Het merendeel van de deelnemers vond de opgaven in de toets duidelijk, namelijk 76,4 procent (gemiddelde waardering was 2.2). Tussen het oordeel over de duidelijkheid van de opgaven en de totale toetsscore is geen significante samenhang ($r = .17$, $p = .18$), maar tussen het oordeel en de subschaal 'recollection' wel ($r = .24$, $p < .05$).

Hoewel een groot aantal deelnemers de opgaven duidelijk vond, werd een aantal kanttekeningen gemaakt door de deelnemers die niet elke opgave duidelijk vonden. De vraagstelling werd soms als onduidelijk ervaren. Dat was bijvoorbeeld het geval bij de volgende vraag van het itemtype 'recollection': 'Hoe hard mag u nu rijden?' Volgens een aantal deelnemers was het onduidelijk of gevraagd werd naar de toegestane snelheid of naar de snelheid die men zelf zou hanteren in de situatie van het videofragment. Daarnaast vonden deelnemers het onderscheid niet duidelijk tussen twee antwoordmogelijkheden bij het opgaventype 'action selection', namelijk 'gas los' en 'remmen'. Tevens was niet altijd duidelijk welke doelen de bestuurder in het videofragment heeft, terwijl dit in de werkelijkheid wel het geval is. Daar hangt soms het antwoord vanaf in de ogen van de respondenten.

Bij de categorie 'timed response' werden een aantal specifieke opmerkingen gemaakt. Een aantal deelnemers vond de opgaven onduidelijk en moest wennen aan de opgavenvorm. Aangegeven werd dat het beter ging na een aantal opgaven gemaakt te hebben. Daarnaast misten deelnemers een tijdsweergave. Hierdoor wisten deelnemers niet hoe lang het filmpje nog zou duren, waardoor ook niet duidelijk was wanneer geklikt zou kunnen worden. Tevens werd gesteld dat iemand niet aan kan geven wat het beste moment is om te remmen als niet het gehele filmpje gezien is. Na het klikken op het filmpje kwamen deelnemers erachter dat het filmpje meteen stopt. Volgens een deelnemer is het wellicht handig om dit te vermelden bij de instructie van het opgavenblok. De set opgaven van het type 'timed response' werd als aanmerkelijk moeilijker ervaren dan de andere opgaventypes.

Ten slotte vonden enkele deelnemers dat het opgaventype 'hazard perception' geen realistische meting van situatiebewustzijn is, omdat het beeld stilstaat en de deelnemer onbepakt de tijd kan nemen het (potentieel) meest gevaarlijke object te kiezen. Deze tijd ontbreekt in werkelijkheid tijdens het besturen van een auto.

Duidelijkheid filmpjes

De duidelijkheid van de videofragmenten werd door de deelnemers als matig tot slecht ervaren (de gemiddelde waardering was 1.5). Van een significante samenhang tussen het oordeel over de filmpjes en de totale toetsscore is echter geen sprake ($r = -.03$, $p = .83$). De filmpjes werden als onduidelijk ervaren, soms omdat het formaat te klein was, soms omdat de filmpjes te donker waren. Daarnaast duurden volgens een aantal deelnemers de filmpjes te kort, waardoor geen goed beeld kon worden gevormd van de verkeerssituatie. De achteruitkijkspiegel trekt volgens sommige deelnemers te veel aandacht en staat in geen verhouding tot de grootte van de voorruit. Tevens werden door sommige deelnemers de zijspiegels gemist en werden de opgaven minder realistisch gevonden, omdat deelnemers niet naast zich konden kijken. Vooral bij het itemtype 'hazard perception' werd dit als belemmerend ervaren. Ten slotte dacht één deelnemer aan het begin van de toets dat de opgaven gemaakt moesten worden vanuit het perspectief van de auto vóór de camera-auto reed.

Indeling beeldscherm

Een meerderheid van 69,8 procent van de deelnemers vond de indeling van het beeldscherm overzichtelijk (gemiddelde waardering: 1.8). Het blijkt dat het oordeel over deze stelling niet significant samenhangt met de totaalscore op de toets ($r = .23$, $p = .07$).

Een tweetal opmerkingen over de indeling van het beeldscherm werd gemaakt. Het werd als verwarrend ervaren dat bij de categorieën 'timed response' en 'hazard perception' de vragen ineens boven het videofragment getoond werden in plaats van eronder, zoals het geval is bij de opgaventypen 'recollection' en 'action selection'. Daarnaast vonden deelnemers het verwarrend dat er verschillende knoppen moesten worden aangeklikt om door te gaan naar de volgende vraag, namelijk een enkele pijl rechtsonder in beeld, een dubbele pijl rechtsonder in beeld en om af te sluiten een kruisje rechtsboven in beeld. De laatste knop zou te verwarren zijn met het browserkruisje.

Voldoende tijd

88,2 procent van de deelnemers had voldoende tijd om elke vraag in te vullen (gemiddelde waardering: 2.5). Het merendeel (89,6 procent) had nergens onvoldoende tijd of bij één tot vijf opgaven onvoldoende tijd. Er is geen significante correlatie tussen het oordeel over voldoende tijd en de toetsscore ($r = .22$, $p = .08$). Wel is er een significante samenhang tussen het oordeel en de score op de subschaal 'recollection' ($r = .25$, $p < .05$).

De deelnemers geven aan dat het prettig zou zijn wanneer de duur van het videofragment in beeld zou staan. Dit is in de toets niet het geval en dit werd gemist. Vooral bij het opgaventype 'timed response' was het voor een aantal deelnemers een belemmerende factor dat de deelnemer niet weet hoe lang het filmpje nog duurt.

Slechtziend en kleurenblindheid

De factoren slechtziendheid en kleurenblindheid zouden, in verband met de digitale afname, invloed kunnen hebben op de toetsscores. Voor beide categorieën zijn geen significante afwijkingen in toetsscores. Echter, het aantal deelnemers dat slechtziend ($n = 7$) of kleurenblind ($n = 3$) is, is te klein

om een betrouwbare uitspraak te kunnen doen over de invloed van kleurenblindheid en slechthoortheid op de toetsscores.

Spelcomputer

Een volgende factor die invloed zou kunnen hebben op de toetsscores is de ervaring met spelcomputers. Aangezien de Toets Sitatiebewustzijn digitaal wordt afgenomen bestaat de kans dat ervaren spelers beter op de toets scoren dan onervaren spelers. De samenhang tussen ervaring met de spelcomputer en toetsscores blijkt niet significant ($r = .11$, $p = .37$ voor 'recollection' en $r = .13$, $p = .31$) voor de totaalscore.

Motivatie en concentratie

Motivatie en concentratie kunnen invloed hebben op de toetsscores: vonden de deelnemers de toets leuk en/of leerzaam? Kon men zich voldoende concentreren en was er sprake van onzekerheid? Daarnaast werd gekeken naar de samenhang tussen concentratie en de invloed van onzekerheid en de prestatie op de toets.

De meeste deelnemers – 86,9 procent – vonden het leuk om de toets te maken (gemiddelde waardering: 2.5). Er is echter geen significante samenhang tussen dit oordeel en de totale toetsscore ($r = .093$, $p = .464$). Dit is positief: een lage motivatie heeft dus geen invloed op toetsscores. Tevens vond een meerderheid, namelijk 60,5 procent, de toets leerzaam (gemiddelde waardering: 1.7), maar ook hier is geen sprake van een significante samenhang tussen het oordeel over leerzaamheid en de totale toetsscore ($r = -.11$, $p = .38$). Een andere factor is concentratie. Een meerderheid van 69,8 procent (gemiddelde waardering: 1.9) kon zich goed concentreren. Tussen de mate van concentratie en de totale toetsscore is net geen significante correlatie ($r = .23$, $p = .07$). Tussen de mate van concentratie en de score op 'recollection' is er een positieve significante samenhang ($r = .28$, $p < .05$).

De personen die zich niet goed konden concentreren werden gestoord door binnenlopende collega's, telefoontjes en binnenkomende emailberichten. Daarnaast werd een gebrek aan concentratie gewijd aan het kleine formaat van de filmpjes.

Tot slot was het grootste deel van de deelnemers niet onzeker over hun prestatie op de toets (gemiddelde waardering: 2.2). Variaties in de mate van gevoelde onzekerheid hangen niet samen met de totale toetsscore ($r = -.02$, $p = .89$).

4.3 Analyse van antwoordgedrag en antwoordmodel bij de vier opgaventypen

Om een uitspraak te kunnen doen over de kwaliteit van de antwoordsleutel en de afleiders zijn de afzonderlijke vragen onder de loep genomen. Gekeken wordt naar de aantrekkelijkheid en plausibiliteit (attractiewaarden) van de afleiders en naar de gepastheid van het 'juiste' antwoord. Bespreekgevallen zijn items met alternatieven met hogere attractiewaarden dan p-waarden, items met een lage item restcorrelatie en items waarbij een afleider niet éénmaal gekozen is door de respondenten.

4.3.1 Opgaventype 'recollection'

Bij dit opgaventype 'recollection' zijn twee items gevonden met respectievelijk een lage item restcorrelatie en/of een geringe aantrekkelijkheid van alternatieven.

Recollection Opgave 4

In Tabel 4.5 is te zien dat het goede antwoord 'Nee' is, maar dat de meerderheid 'Ja' kiest. De item restcorrelatie van deze opgave is .033. Dit is meteen de laagste waarde van de vragen in de categorie 'recollection'. In het videofragment is in de achteruitkijkspiegel te zien dat er een auto met aanhanger op de rijbaan naast de deelnemer rijdt. Op het eindscherm is deze niet meer te zien (zie Figuur 4.1). De deelnemers lijken hun aandacht onvoldoende op de achteruitkijkspiegel te hebben gericht. Ook kan het zijn dat een deel veronderstelt dat de auto op de linkerbaan rijdt en er dus ingevoegd kan worden, hetgeen acceptabel is in de praktijk.

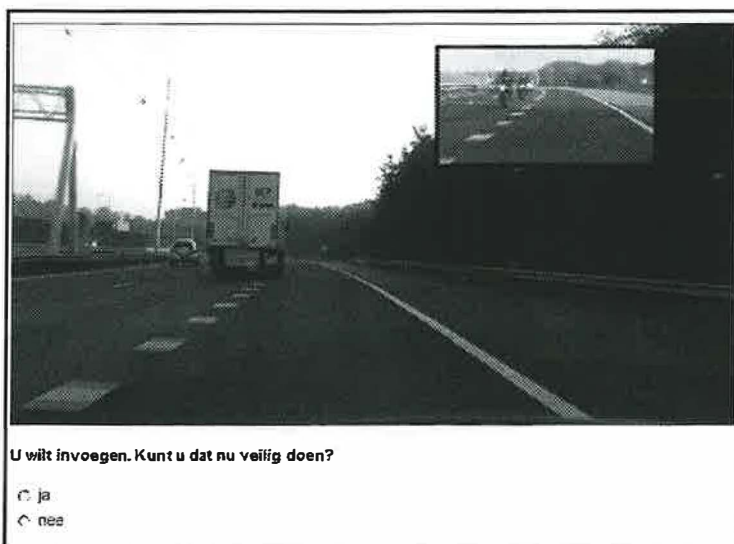
Tabel 4.5

Recollection Opgave 4

	Frequentie	Percentage	Cumulatief geldig percentage
Ja	47	56,0	57,3
Nee*	35	41,7	100,0
Totaal	82	97,6	--
Missing System	2	2,4	--
Totaal	84	100,0	--

Noot

* Juist volgens antwoordsleutel



Figuur 4.1 Recollection opgave 4

Recollection Opgave 8

In Tabel 4.6 is te zien dat een grote meerderheid het juiste antwoord kiest. Twee van de overige antwoordmogelijkheden zijn niet gekozen. Het videofragment laat duidelijk zien dat men binnen de bebouwde kom rijdt, de maximumsnelheid is daar 50 km/uur. Verondersteld kan worden dat alle deelnemers hiervan op de hoogte zijn en om die reden alleen voor '30' en '50' kiezen. Deze opgave had een item restcorrelatie van .356 en kan dus gezien worden als zeer illustratief voor dit opgaventype. De eindsituatie wordt weergegeven in Figuur 4.2.

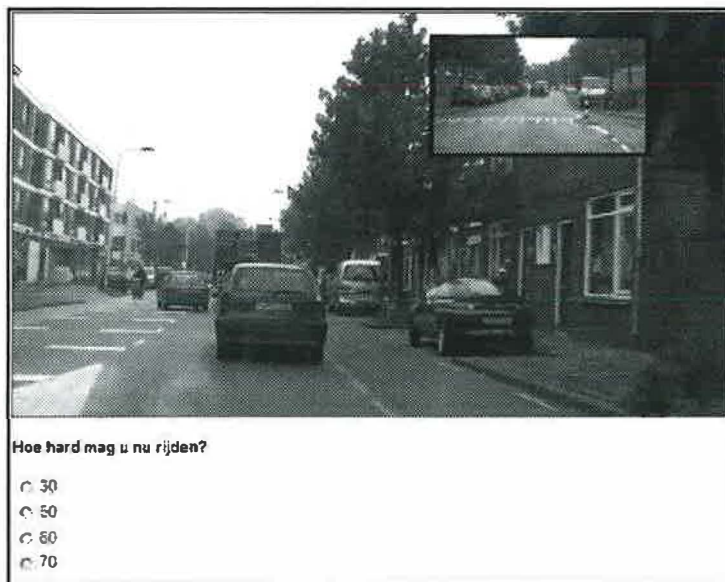
Tabel 4.6

Recollection Opgave 8

	Frequentie	Percentage	Cumulatief geldig percentage
30*	69	84,1	87,3
50	10	12,2	100,0
60	0	0	--
70	0	0	--
Totaal	79	96,3	--
Missing system	3	3,7	--
Totaal	82	100,0	--

Noot

* Juist volgens antwoordsleutel



Figuur 4.2 Recollection opgave 8

Desondanks heeft 12,8 procent het bord niet gezien (zie Figuur 4.3).



Figuur 4.3 Recollection opgave 8

Recollection Opgave 10

Opgave 10 is opvallend, omdat wederom twee antwoordalternatieven niet gekozen worden (zie Tabel 4.7). Ook in deze opgave wordt naar snelheid gevraagd. Hier worden langzaam twee borden gepasseerd (zie Figuur 4.4). Verondersteld wordt dat iedereen daarvan op de hoogte is en daarom twee antwoordalternatieven niet worden gekozen.

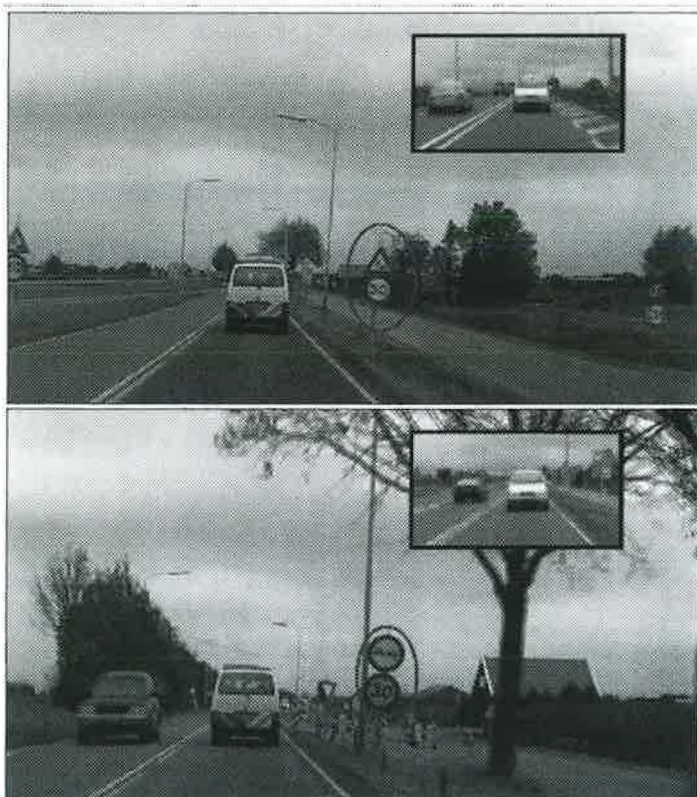
Tabel 4.7

Recollection Opgave 10

	Frequentie	Percentage	Cumulatief geldig percentage
30*	73	89,0	92,4
50	6	7,3	100,0
60	0	0	--
70	0	0	--
Totaal	79	96,3	--
Missing system	3	3,7	--
Totaal	82	100,0	--

Noot

* Juist volgens antwoordsleutel



Figuur 4.4 Recollection opgave 10

4.3.2 Opgaventype 'action selection'

Bij alle items van deze opgavencategorie werden alle antwoordmogelijkheden gekozen. De frequenties van afleiders verschilden sterk bij de items die hieronder besproken worden. Bij deze items werd ook een onjuist antwoord het meest gekozen.

Action Selection Opgave 12

In Tabel 4.8 is te zien dat meer deelnemers het onjuiste antwoordalternatief ('niets, zo blijven rijden') hebben gekozen dan het goede ('gas los'). De eindsituatie is weergegeven in Figuur 4.5. Te zien is dat de remlichten en richtingaanwijzer niet goed te zien zijn op het eindshot. Mogelijk verklaart dit de massale keuze om 'niets te doen'. Tevens wordt het antwoordalternatief 'uitwijken naar links' door niemand gekozen.

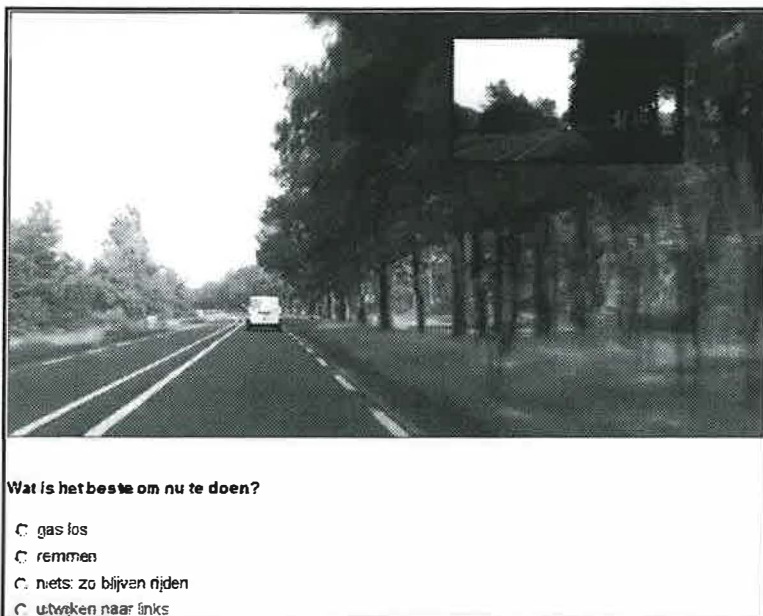
Tabel 4.8

Action Selection Opgave 10

	Frequentie	Percentage	Cumulatief geldig percentage
Missing system	2	2,4	2,4
Gas los*	28	34,1	36,6
Niets: zo blijven rijden	40	48,8	85,4
Remmen	12	14,6	100,0
Uitwijken naar links	0	0	100,0
Totaal	82	100,0	--

Noot

* Juist volgens antwoordsleutel

**Figuur 4.5 Action Selection Opgave 12***Action Selection Opgave 13*

In Tabel 4.9 is wederom te zien dat meer deelnemers een onjuist antwoordalternatief 'achter de fietsers blijven' kiezen dan het goede 'de fietsers passeren'. De eindsituatie is weergegeven in Figuur 4.6. De bestuurder rijdt in het eindshot op een verhoging en zal een drempel omlaag gaan passeren. Wanneer de bestuurder in zou gaan halen, zou harder gereden moeten worden, met het risico dat de voorbumper tegen het wegdek komt. Na opnieuw bekijken van het videofragment is in samenspraak met een verkeersdeskundige gekozen voor het aanpassen van de antwoordsleutel. Bij twijfel in het verkeer wordt geacht de meest veilige beslissing te maken en dat is 'achter de fietsers blijven'.

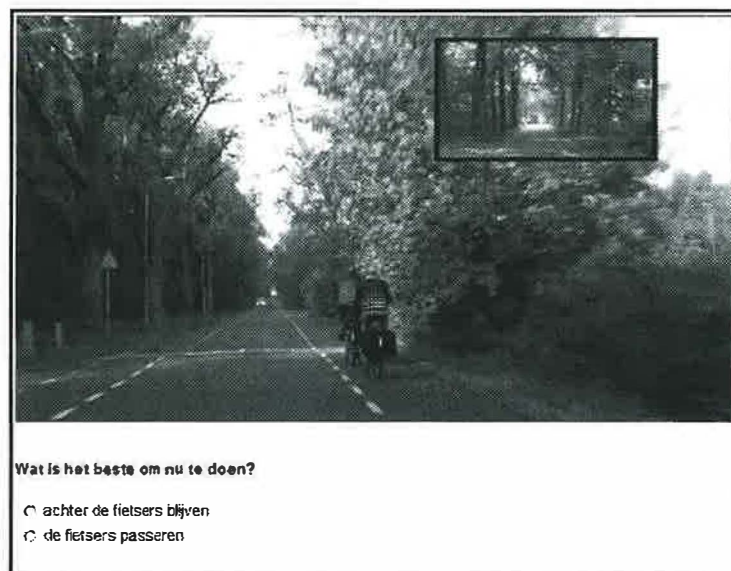
Tabel 4.9

Action Selection Opgave 13

	Frequentie	Percentage	Cumulatief geldig percentage
Missing system	3	3,7	3,7
Achter de fietsers blijven	59	72,0	75,6
De fietsers passeren*	20	24,4	100,0
Totaal	82	100,0	--

Noot

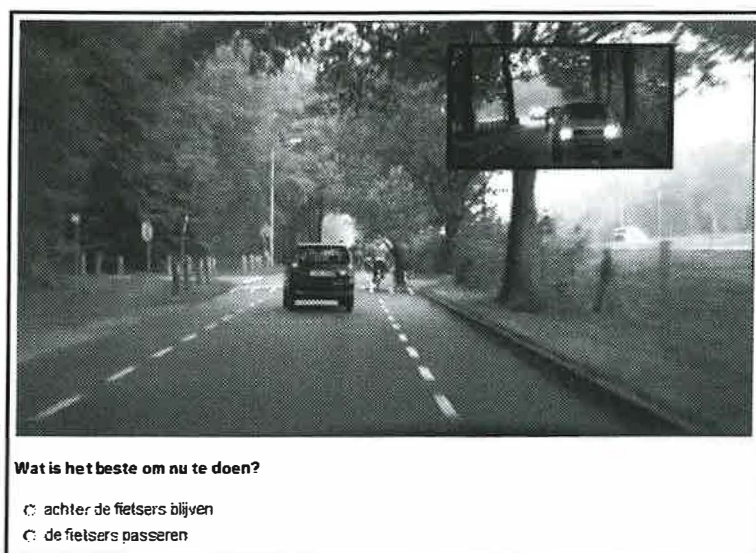
* Juist volgens antwoordsleutel



Figuur 4.6 Action Selection opgave 13

Action Selection Opgave 14

In deze opgave heeft 71 procent van de deelnemers het juiste antwoord 'achter de fietsers blijven' gekozen. De opgave heeft een item restcorrelatie van $-.054$. Net als in opgave 13 kan bij deze opgave gekozen worden uit twee antwoordalternatieven. De opgave is niet illustratief voor de toets. De opgave is weergegeven in Figuur 4.7. De voorgaande auto gaat de fietsers inhalen. Voor de bestuurder is het niet zichtbaar of hij ook kan inhalen. Het is aannemelijk dat de bestuurder achter de fietsers blijft.



Figuur 4.7 Action Selection opgave 14

Action Selection Opgave 15

In Tabel 4.10 is te zien dat ook bij deze opgave meer mensen een onjuist antwoordalternatief kiezen, namelijk 'gas los'. De eindsituatie is weergegeven in Figuur 4.8. Rechts naast de weg ligt een verharde berm. De vraag is of de bestuurder daar al op rijdt of nog niet. Na het bekijken van de oorspronkelijke situatie, waarin ook het vervolg van de manoeuvre te zien is, blijkt dat de bestuurder inderdaad naar rechts uitwijkt. De antwoordsleutel is aangepast, het goede antwoord is 'uitwijken naar rechts'. De betrouwbaarheidsanalyses zijn herhaald op basis van de nieuwe antwoordsleutel.

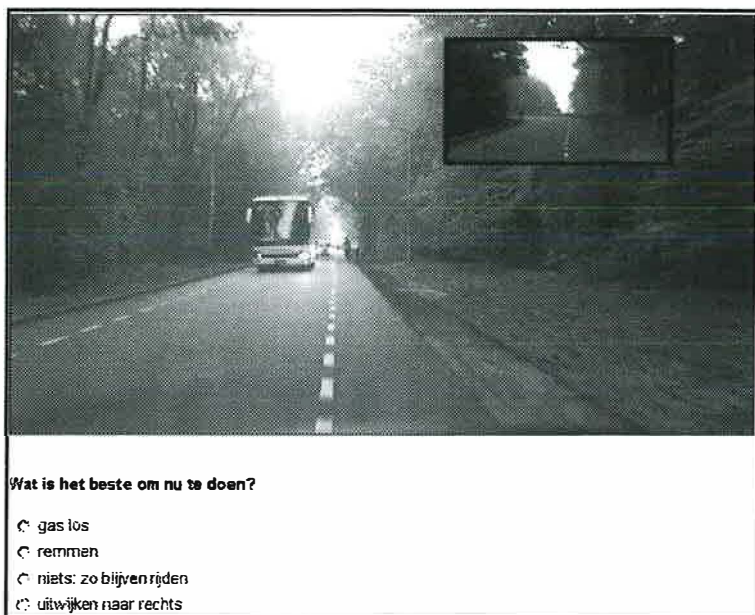
Tabel 4.10

Action Selection Opgave 15

	Frequentie	Percentage	Cumulatief geldig percentage
Missing system	3	3,7	3,7
Gas los*	20	24,4	28,0
Niets: zo blijven rijden	9	11,0	39,0
Remmen	10	12,2	51,2
Uitwijken naar rechts	40	48,8	100,0
Totaal	82	100,0	--

Noot

* Juist volgens antwoordsleutel



Figuur 4.8 Action Selection Opgave 15

Action Selection Opgave 18

In Tabel 4.11 is te zien dat ook bij deze opgave een groter aantal deelnemers een onjuist antwoordalternatief kiest, namelijk 'niets, zo blijven rijden', dan het goede. De eindsituatie is weergegeven in Figuur 4.9. Voor dit eindshot waren er boven de weg knipperende matrixborden geactiveerd die een maximumsnelheid van 50 aangaven, wat aanleiding zou moeten zijn tot het verminderen van de snelheid. Een mogelijke verklaring zou kunnen zijn dat de deelnemer verwacht dat de bestuurder daardoor al gas terugneemt of al gas terug heeft genomen, waardoor het niet direct noodzakelijk is om snelheid te minderen.

Tabel 4.11

Action Selection Opgave 18

	Frequentie	Percentage	Cumulatief geldig percentage
Missing system	2	2,4	2,4
Inhalen	8	9,8	12,2
Niets: zo blijven rijden	40	48,8	61,0
Snelheid minderen*	32	39,0	100,0
Totaal	82	100,0	--

Noot

* Juist volgens antwoordsleutel



Figuur 4.9 Action Selection Opgave 18

4.3.3 Opgaventype 'hazard perception'

Een nadere analyse van het antwoordgedrag op opgaven van het type 'hazard perception' is niet mogelijk. Het bleek niet mogelijk met de gebruikte toetsafname de coördinaten te bewaren van de plaatsen waarop kandidaten hadden geklikt. De enige gegevens die uit de toets gekomen zijn is of de deelnemers binnen of buiten het juiste gebied hebben geklikt. Opgave 22 heeft de laagste item restcorrelatie van dit opgaventype, namelijk .003. Dit item is niet illustratief voor de toets. 89 procent van de deelnemers heeft de vraag goed beantwoord. De opgave is weergegeven in Figuur 4.10.



Figuur 4.10 Hazard Perception Opgave 22

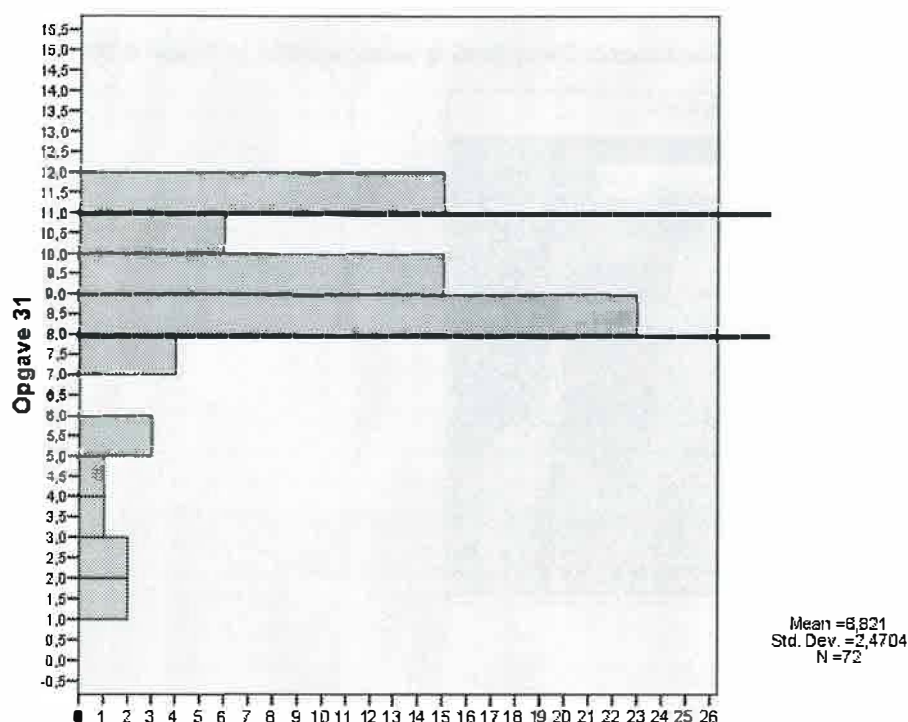
4.3.4 Opgaventype 'timed response'

Het juiste antwoord bij de categorie 'timed response' werd gegeven door op het juiste moment aan te geven om een manoeuvre in te zetten. Dit kan met behulp van een klik op het filmpje. Wanneer de deelnemers binnen het juiste kritieke tijdvenster klikten, was de opgave juist beantwoord, wanneer er buiten het kritieke tijdvenster werd geklikt was de opgave onjuist beantwoord. Wanneer nader

gekeken wordt naar het antwoordgedrag van de deelnemers kan een trend worden gesignaleerd. Het merendeel van de eerste vragen in de reeks wordt vaak te vroeg beantwoord, dus vóór de kritieke tijdsrange. Bij het merendeel van de laatste vragen wordt vaak ná de kritieke tijdsrange geklikt. Dit kan betekenen dat deelnemers in de loop van de toets wennen aan de vraagvorm, waardoor het antwoordgedrag verandert. Alle opgaven zullen kort besproken worden, met uitzondering van opgave 33, 35 en 37 die een bevredigende item restcorrelatie laten zien. Deze opgaven kennen voldoende spreiding over mogelijke reactietijden (zie bijlage F voor de grafieken met de reactietijden bij elke opgave uit de categorie 'timed response'). Opgave 40 kent ook een wat hogere item restcorrelatie ($r_{ir} = .183$), maar het antwoordgedrag bij die opgave is opvallend, waardoor ook opgave 40 besproken wordt.

Opgave 31

De meeste deelnemers klikken binnen het kritieke tijdsvenster, met een piek tussen de 8 en 9 seconden (zie Figuur 4.11). Het hoogste aantal deelnemers klikt op het moment dat de fietser de weg opstuurt om de geparkeerde auto te ontwijken (zie Figuur 4.12). De antwoordsleutel ging uit van een correcte reactie als de fietser op de weg rijdt en het te smal wordt om als autobestuurder voorbij te rijden. De reactie tussen 8 en 9 seconden is echter aannemelijk. Na nadere bestudering van het videofragment is daarom besloten de antwoordsleutel aan te passen. Tevens is een piek te zien net na de 11 seconden. Deze piek ontstaat door de deelnemers die niet hebben geklikt, waardoor de eindtijd van het filmpje wordt geregistreerd.



Figuur 4.11 Grafiek met reactietijden



Figuur 4.12 Opgave 31 'Wat is het beste moment om te remmen? Klik op dit moment op het filmpje.'

Opgave 32

De meeste deelnemers klikken vóór het kritieke tijdvenster, met een piek tussen vier en vijf seconden (zie bijlage F). Op dat moment wordt achter de geparkeerde vrachtauto een tweede vrachtauto zichtbaar en gaat het portier open van de geparkeerde vrachtauto, zie Figuur 4.13. Deelnemers remmen op dat moment al af. Volgens de antwoordsleutel is het beste moment om te remmen wanneer de bestuurder van de geparkeerde vrachtauto uitstapt en de tweede vrachtauto daarnaast in beeld komt.



Figuur 4.13 Opgave 32 'Wat is het beste moment om te remmen? Klik op dit moment op het filmpje.'

Opgave 34

De meeste deelnemers kiezen vóór het kritieke tijdvenster het moment van inhalen, met een piek tussen de 12 en 15 seconden (zie bijlage F). In het filmfragment wordt de autobestuurder ingehaald door drie auto's, wat te zien is in de achteruitkijkspiegel. De piek ontstaat wanneer de tweede auto in het beeld van de voorruit komt. De derde auto zit nog in de dode hoek en is voor de deelnemers niet zichtbaar, zie Figuur 4.14. Wanneer goed opgelet is, in de achteruitkijkspiegel, is het aannemelijk dat de deelnemer weet dat er nog een auto moet passeren, voordat ingehaald kan gaan worden. Kennelijk heeft 69 procent van de deelnemers dit niet opgemerkt.



Figuur 4.14 Opgave 34 'Wat is het beste moment om naar links te gaan? Klik op dit moment op het filmpje.'

Opgave 36

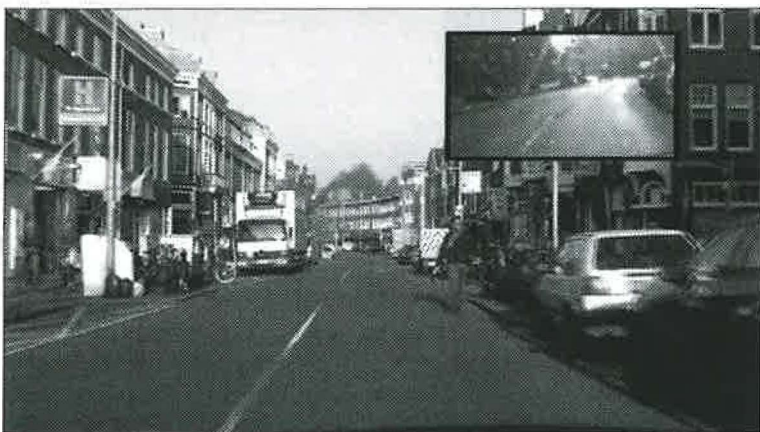
De opgave luidt: 'Wat is het beste moment om gas terug te nemen. Klik op dit moment op het filmpje.' Een kleine meerderheid van de deelnemers kiest vóór het kritieke tijdsvenster het moment om gas terug te nemen, met pieken tussen de 3 en 6 seconden (zie bijlage F). Het juiste moment van reageren lag tussen de 11 en 13 seconden. Op dat moment wordt de fietser zichtbaar die de weg oversteekt. Het kan zijn dat de deelnemers eerder klikten, omdat de auto een kruispunt over rijdt met een verkeerslicht dat op groen staat (zie Figuur 4.15), en bij zes seconden staat er op de weg 'zone' aangegeven (zie Figuur 4.16). Dit kan leiden tot afremgedrag. Een verklaring zou kunnen zijn dat de deelnemer niet weet hoe hard de auto rijdt en daarom wat vaart wil minderen in deze situaties. Een andere piek zit tussen 14 en 15 seconden, ná het kritieke tijdvenster dus. Op dat moment komt heel duidelijk een voetganger in beeld die op de weghelft van de bestuurder loopt (zie Figuur 4.17).



Figuur 4.15 Druk kruispunt met verkeerslicht



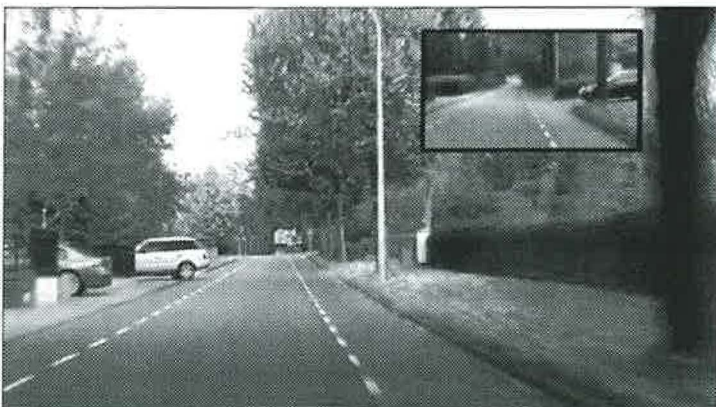
Figuur 4.16 Zone



Figuur 4.17 Voetganger

Opgave 38

In deze opgave wordt gevraagd naar het beste moment om gas terug te nemen. Het kritieke tijdvenster van deze opgave is 9 tot 11 seconden. Een grote meerderheid van de deelnemers kiest ná het kritieke tijdvenster het moment om gas terug te nemen, met een piek tussen de 11 en 12 seconden (zie bijlage F). Vlak daarvóór komt er mogelijk onverwachts een auto uit een uitrit (zie achteruitkijkspiegel in Figuur 4.18), wat kan afleiden en waardoor iets later geklikt wordt. Tussen de 11 en 12 seconden staat de auto waar het om gaat echter alweer stil (zie ter illustratie Figuur 4.18).



Figuur 4.18 Opgave 38 'Wat is het beste moment om gas terug te nemen. Klik op dit moment op het filmpje.'

Opgave 39

In deze opgave wordt gevraagd naar het beste moment om te remmen. De meeste deelnemers hebben de opgave goed beantwoord, zij reageerden binnen het kritieke tijdsvenster tussen de 6 en 10 seconden. Een aantal deelnemers heeft te vroeg geklikt. Dit zou te maken kunnen hebben met de remmende auto's die ver voor hen rijden. De witte auto, te zien in Figuur 4.19, komt de rijbaan van de bestuurder op, maar dit gebeurt soepel, waardoor er niet per se gas teruggenomen hoeft te worden. De piek na het kritieke tijdsvenster geeft het aantal deelnemers weer dat niet gereageerd heeft en de opgave dus fout heeft, zie bijlage F.



Figuur 4.19 Opgave 39 'Wat is het beste moment om te remmen? Klik op dit moment op het filmpje.'

Opgave 40

De bestuurder rijdt op een rechte weg op een ladder af. De ladder staat op straat tegen de gevel van een stadspand. Op de ladder staat een man. In bijlage F is te zien dat iets meer dan de helft van de deelnemers binnen het kritieke tijdsvenster zouden afremmen. Echter, vrijwel evenveel deelnemers zouden dit later doen. Dit kan op een verschil in interpretatie duiden van het begrip 'afremmen'. Zie Figuur 4.20 voor een weergave van de situatie.



Figuur 4.20 Opgave 40 'Wat is het beste moment om te remmen? Klik op dit moment op het filmpje.'

4.4 Samenhang tussen de toetsscores en externe criteria

In deze paragraaf wordt in lijn met onderzoeksvraag 4 gekeken naar de samenhang tussen toetsscores en externe criteria.

4.4.1 Samenhang met Zelfevaluatie Rijgedrag

Gekeken is naar de samenhang tussen de scores op de Toets Situatiebewustzijn en een algemene maat voor rijvaardigheid: de Zelfevaluatie Rijgedrag. In Tabel 4.12 wordt de correlatiematrix weergegeven. Een positieve significante samenhang ($\alpha = .05$) blijkt te bestaan tussen de score op de subschaal 'recollection' van de Toets Situatiebewustzijn en de subschaal 'milieuvriendelijk' van de zelfevaluatie ($r = .28, p < .05$). Daarnaast is er een significant positieve samenhang tussen de subschaal 'recollection' van de Toets Situatiebewustzijn en de subschaal 'controle over het voertuig' ($r = .27, p < .05$). Tevens is er een significante samenhang tussen de set opgaven van de subschaal 'recollection' van de Toets Situatiebewustzijn en de totaalscore op de Zelfevaluatie. Gezien de subschaal 'recollection' het meest betrouwbaar is, werd verwacht dat deze subschaal samen zou hangen met onderdelen van de zelfevaluatie. De samenhang tussen de overige opgaventypen van de Toets Situatiebewustzijn met de categorieën van de zelfevaluatie zijn niet significant, de correlaties zijn grotendeels zeer laag.

Tabel 4.12

Correlatiematrix Toets Situatiebewustzijn en Zelfevaluatie Rijgedrag

	<u>Subschalen Toets Situatiebewustzijn</u>				
	Recollection	Action Selection	Hazard Perception	Timed Response	Totaalscore
Veilig Rijgedrag	.11	.01	-.02	.004	-.04
Sociaal Rijgedrag	.15	-.01	-.04	.07	-.10
Doorstromen	.20	.17	-.02	-.07	-.05
Milieuvriendelijk	.28(*)	-.04	.02	.22	.11
Controle	.27(*)	.16	.08	-.08	.04
Totaalscore Zelfevaluatie	.28(*)	.07	.02	.06	.01

Noot

* Correlatie is significant bij $\alpha = 0.05$ (tweezijdig)

4.4.2 Samenhang met ongevalbetrokkenheid

Een tweede extern criterium dat in dit onderzoek betrokken is, is ongevalbetrokkenheid. Het zou kunnen dat er een verschil bestaat in toetsscores tussen deelnemers die betrokken zijn geweest bij een ongeval en deelnemers die niet betrokken zijn geweest bij een ongeval. Er blijkt echter geen significante samenhang te bestaan tussen toetsscores en ongevalbetrokkenheid (respectievelijk $r = -.02$, $p = .90$ voor de subschaal 'recollection' en $r = -.05$, $p = .72$). De verwachting was dat ongevalbetrokkenheid zou samenhangen met de toetsscores. Daarbij dient vermeld te worden dat slechts 16 procent van de deelnemers betrokken is geweest bij een ongeval.

4.5 Relatie tussen toetsresultaten en achtergrondkenmerken

In deze paragraaf worden de toetsscores in verband gebracht met persoonlijke relevante achtergrondvariabelen van de deelnemers.

Leeftijd

Gekeken is naar de samenhang tussen leeftijd en de scores op de toets (zie Tabel 4.13).

Tabel 4.13

Correlatie Leeftijd en Toetsscores

	Leeftijd	Score op Recollection	Totaalscore
Leeftijd	1	-.33(*)	-.41(**)

Noot

* Correlatie is significant bij $\alpha = 0.05$ (tweezijdig).

** Correlatie is significant bij $\alpha = 0.01$ (tweezijdig).

Uit Tabel 4.13 blijkt dat de samenhang tussen leeftijd en score op 'recollection' negatief significant is (bij $p < .05$). Dat wil zeggen hoe hoger de leeftijd, hoe minder goed gescoord wordt op de herinneringsopgaven. Hetzelfde geldt in sterkere mate voor de correlatiecoëfficiënt van leeftijd en totale score ($p < .01$). Verwacht werd dat leeftijd positief samen zou hangen met de toetsscores, in verband met rijervaring.

Geslacht

De tweede persoonlijke factor die bekeken werd is geslacht. Gekeken is of de scores van mannen significant afwijken van de scores van vrouwen. De gemiddelde scores van mannen en vrouwen op wijken zowel bij het opgaventype 'recollection' als bij de totaalscore niet significant af (respectievelijk $t(67) = -.63$, $p = .53$ en $t(64) = -.97$, $p = .39$).

Opleiding

De derde persoonlijke factor die de scores zou kunnen beïnvloeden is het niveau van opleiding. In Tabel 4.14 wordt een overzicht gegeven van verschil in opleiding en gemiddelde scores. De resultaten laten zien er geen grote verschillen zijn tussen hoogte van scores en de hoogst afgeronde opleiding. De verschillen zijn niet significant ($p = .88$ bij 'recollection' en $p = .81$ bij de totaalscores).

Tabel 4.14

Beschrijvende Statistieken Opleidingniveau

Opleiding	Gemiddelde Recollection	N	%	SD	Gemiddelde Totaalscore	N	%	SD
Havo	7,50	6	8,7	0,837	22,67	6	9,1	3,077
VWO	7,00	5	7,2	1,225	23,25	4	6,1	2,363
MBO	7,64	11	16,2	1,027	23,09	11	16,7	3,590
HBO	7,76	17	24,6	2,278	22,63	16	24,2	4,303
WO	7,30	27	39,1	1,068	22,42	26	39,4	3,443
Totaal	7,48	69	100,0	1,612	22,92	66	100,0	3,404

Rijbewijsbezit

Een vierde factor die invloed zou kunnen hebben op de toetsscores is het aantal jaren dat de deelnemer in bezit is van een rijbewijs. De samenhang tussen het aantal jaren rijbewijsbezit en de score op 'recollection' is $r = .21$ ($p = .08$) en tussen het aantal jaren rijbewijsbezit en de totaalscore is $r = .30$ ($p < .05$). Naarmate deelnemers langer hun rijbewijs bezitten scoren ze hoger op situatiebewustzijn. Verwacht werd dat langer rijbewijsbezit positief samenhangt met de toetsscores.

Kilometrage

Een vijfde externe factor is het aantal kilometers dat per jaar gereden wordt. Hoe hoger het aantal kilometers dat jaarlijks gereden wordt, hoe meer een bestuurder doorgaans aan gevaren wordt blootgesteld. Dit zou juist ook kunnen leiden tot meer ervaring met complexe verkeerssituaties, wat tot uiting zou kunnen komen in een groter situatiebewustzijn. Echter, de hoeveelheid kilometers dat per jaar gereden wordt blijkt niet significant samen te hangen met de toetsscores (respectievelijk $r = -.08$, $p = .52$ voor de score op de subschaal 'recollection' en $r = .02$, $p = .89$ voor de totaalscore). Verwacht werd dat kilometrage positief samen zou hangen met de toetsscores.

5. Conclusie en discussie

In dit laatste hoofdstuk zal antwoord gegeven worden op de onderzoeksvragen en worden de resultaten in perspectief van de literatuur geplaatst. Aan de hand van de resultaten van het onderzoek zullen aanbevelingen gedaan worden ter verbetering van de Toets Situatiebewustzijn. Mede aan de hand van de eerder besproken literatuur en resultaten zullen beperkingen van de toets in kaart gebracht worden en een antwoord worden gegeven op de hoofdvraag. Ten slotte zullen de beperkingen van het onderzoek besproken worden alsmede de implicaties voor vervolgonderzoek.

5.1 Onderzoeksvragen

In dit onderzoek stond de volgende hoofdvraag centraal met betrekking tot de Toets Situatiebewustzijn:

In hoeverre meet de ontwikkelde toets op een betrouwbare en valide wijze het construct situatiebewustzijn bij automobilisten?

Om te onderzoeken in welke mate de Toets Situatiebewustzijn valide en betrouwbaar het construct situatiebewustzijn meet, werden de volgende deelvraagstellingen onderzocht:

1. In hoeverre worden bij de Toets Situatiebewustzijn passende afnamecondities geschapen, die de kandidaten gelegenheid geven hun kunnen te tonen?
(afnamefase model Crooks & Kane, 1996)
2. Hoe groot is de interne consistentie van de Toets Situatiebewustzijn?
(generalisatiefase model Crooks & Kane, 1996)
3. In hoeverre is het gekozen antwoordmodel bij de itemtypes passend bij het beoogde taakgedrag in de opgaven?
(scoringfase model Crooks & Kane, 1996)
4. In hoeverre hangen scores op de Toets Situatiebewustzijn samen met maten voor rijvaardigheid in het algemeen en relevante externe criteria, zoals ongevalbetrokkenheid?
(extrapolatiefase model Crooks & Kane, 1996)
5. In hoeverre hangt de mate van situatiebewustzijn samen met persoonlijke achtergrondvariabelen, die uit de literatuur als relevant naar voren komen?
(extrapolatiefase model Crooks & Kane, 1996)

5.2 Conclusies

5.2.1 Taakdomein

Uit de resultaten blijkt dat de Toets Situatiebewustzijn een matig tot zwak valide en betrouwbare toets is. De toets kent vier verschillende opgaventypen. Met de vier verschillende typen opgaven wordt beoogd het construct situatiebewustzijn zo volledig mogelijk te meten. Uit de literatuur blijkt dat herinneringsitems situatiebewustzijn meten, maar dat alleen een goede herinnering niet als voldoende wordt gezien voor een goed situatiebewustzijn (Gugerty, 1997). Ook (deel)prestatiemetingen zijn

belangrijk volgens Gugerty. Hier is rekening mee gehouden door de opgaventypen 'action selection', 'hazard perception' en 'timed response' in de toets te gebruiken. De deelnemer moet in het verkeer keuzes maken, maar ook op tijd reageren. De vier opgaventypen passen tevens in het model van Endsley (1995b). Waarnemen, begrip van de situatie, projectie (voorspellen van de situatie), beslissen en handelen vormen de kern van het model. De set opgaven van het type 'recollection' toetst of de situatie goed is waargenomen. De opgaventypen 'action selection', 'hazard perception' en 'timed response' toetsen of de waargenomen situatie begrepen wordt, of men inziet hoe de situatie zal ontwikkelen en of op het juiste moment de juiste beslissingen genomen worden en de juiste handelingen worden uitgevoerd. Daarnaast bevestigen McKenna, Horswill & Alexander (2006) dat gevaarherkenning erg belangrijk is in het verkeer. Het opgaventype 'hazard perception' dekt dit onderdeel van situatiebewustzijn in het verkeer.

Volgens de literatuur zouden deze vier opgaventypen gezamenlijk het construct situatiebewustzijn moeten kunnen meten. De mate waarin dit betrouwbaar en valide is gerealiseerd komt in de volgende paragrafen aan de orde.

5.2.2 Interne consistentie in relatie tot afnamecondities en antwoordmodel

Onderzoeksvraag 1, 2 en 3 hadden betrekking op de afnamecondities, interne consistentie en het antwoordmodel van de Toets Situatiebewustzijn. De interne consistentie van de Toets Situatiebewustzijn was aanvankelijk laag. De subschaal gevormd door de set opgaven van het type 'recollection' is het meest betrouwbaar. Het is, zoals eerder in de literatuur vermeld (Gugerty, 1997) belangrijk voor het situatiebewustzijn dat relevante zaken herinnerd worden. Het verkeer achter en naast de bestuurder, verkeersregels en snelheden zijn zeer belangrijk in het verkeer. Kanttekening bij dit opgaventype is dat een aantal items betrekking had op maximumsnelheden. Uit de resultaten op de toets en de opmerkingen van deelnemers blijkt dat er gewenning optreedt en dat er na verloop van tijd automatisch al extra op snelheidsborden gelet wordt.

Het opgaventype 'action selection' is weinig betrouwbaar, zo blijkt uit de resultaten. Een verklaring zou kunnen zijn dat niet de gehele situatie te overzien is en niet duidelijk is in welk soort auto gereden wordt. Dit heeft verschil in keuzes tot gevolg. Tevens roept de formulering van antwoordmogelijkheden vragen op bij de kandidaten. Zo blijkt dat voor een aantal deelnemers het verschil tussen 'gas los' en 'remmen' niet duidelijk is. Deze begrippen zijn op verschillende manieren te interpreteren, omdat ze niet vooraf in de toetsinstructie zijn gedefinieerd.

Op het opgaventype 'hazard perception' wordt door vrijwel iedereen goed gescoord. De set opgaven is weinig betrouwbaar, zo blijkt. Een mogelijke verklaring hiervoor is dat wanneer het videofragment stopt, het eindshot in beeld blijft totdat de deelnemer naar de volgende opgave klikt. Hierdoor kan de deelnemer de gehele situatie uitvoerig bekijken en vervolgens een weloverwogen keuze maken. In de realiteit vindt gevaarherkenning plaats in een dynamische, snel veranderende omgeving.

Het laatste opgaventype is 'timed response'. Uit de resultaten blijkt dat de subschaal met opgaven van dit type vooralsnog onbetrouwbaar is. Het merendeel van de opgaven heeft een zwak negatieve item restcorrelatie. Een mogelijke verklaring hiervoor kan zijn dat er in het antwoordmodel geen rekening gehouden wordt met reactiesnelheid. Een andere mogelijke verklaring kan zijn dat

mensen verschillende keuzes maken afhankelijk van het soort auto waarin gereden wordt. De meest plausibele verklaring lijkt te zijn dat niet te zien is hoe lang het fragment duurt en deelnemers niet weten wanneer het juiste moment daar is, wellicht komt er nog een beter moment.

Uit de resultaten blijkt dat over het algemeen het doel, de instructie, de vraagstelling en de indeling van het beeldscherm duidelijk gevonden werd en de oordelen niet samenhangen met toetsscores. De videofragmenten vond een meerderheid onduidelijk, maar dit oordeel vertoonde ook geen samenhang met de toetsscores. Ervaring met de (spel)computer blijkt niet samen te hangen met toetsscores, wat een positief resultaat is.

Over het algemeen vond men de toets leuk en leerzaam. Ook deze oordelen hangen niet samen met toetsscores. Dit is mogelijk een indicatie dat motivatie geen invloed heeft gehad op toetsscores. Dat is een voordeel wanneer de deelnemers weinig gemotiveerd zijn: de toetsscores zullen daardoor niet beïnvloed worden. De factor 'concentratie' blijkt samen te hangen met de scores op het opgaventype 'recollection'. Dit is een plausibel resultaat. Wanneer een deelnemer door afleiding bepaalde zaken uit het videofragment mist, kan hij die zaken niet herinneren en is de kans groot dat hij de opgave onjuist beantwoordt. Dit komt overeen met de aanname dat verstoring van aandacht ten koste gaat van situatiebewustzijn (Endsley, 2000; Durso & Gronlund, 1999). De feitelijke afnamecondities, afname onder werktijd in het kantoor, zijn mogelijk niet optimaal geweest.

5.2.3 Samenhang Toets Situatiebewustzijn met externe criteria

De vierde onderzoeksvraag had betrekking op de relatie tussen situatiebewustzijn, rijvaardigheid en externe criteria. Verwacht mag worden dat situatiebewustzijn samenhangt met rijvaardigheid in het algemeen, zoals gemeten met de Zelfevaluatie Rijgedrag. De hoogst betrouwbare subschaal van de Toets Situatiebewustzijn is 'recollection'. De score op de subschaal 'recollection' hangt significant samen met milieuvriendelijk rijgedrag, gecontroleerd rijgedrag en de totaalscore op de zelfevaluatie. We vinden hier dus enig empirisch bewijs voor extrapolatie van de resultaten naar externe taakomstandigheden, zoals nu gemeten met behulp van de zelfevaluatie. Twee opmerkingen dienen gemaakt te worden. Ten eerste is het criterium gemeten met een zelfevaluatie, waar feitelijke performance assessment op zijn plaats zou zijn geweest. Ten tweede wordt de correlatie onderschat doordat de toets niet perfect betrouwbaar is. De werkelijke samenhang tussen situatiebewustzijn en rijvaardigheid zou hoger kunnen zijn. Dit biedt potentie voor samenhang tussen de zelfevaluatie en verbeterde versies van de toets met een hogere betrouwbaarheid.

Vooralsnog zijn er geen samenhangen gevonden tussen de mate van situatiebewustzijn en de betrokkenheid bij ongevallen. Mogelijk speelt de beperkte en enigszins selectieve steekproef een rol in dit resultaat. Bij een meer uitgebreide en gevarieerde steekproef zou wel degelijk een samenhang gevonden kunnen worden tussen ongevalbetrokkenheid en toetsscores.

5.2.4 Samenhang Toets Situatiebewustzijn met persoonlijke achtergrondvariabelen

De vijfde onderzoeksvraag had betrekking op de relatie tussen situatiebewustzijn en persoonlijke achtergrondvariabelen, zoals kilometrage en het aantal jaren rijbewijsbezit. Deze achtergrondvariabelen blijken invloed te hebben op de toetsscores. Het aantal jaren rijbewijsbezit hangt substantieel samen met de toetsuitslagen. Dit is in overeenstemming met de door Durso &

Gronlund (1999) benoemde verschillen tussen beginners en experts. Naarmate bestuurders langer hun rijbewijs bezitten, scores ze hoger op situatiebewustzijn. Opvallend is dat leeftijd, in tegenstelling tot uitspraken in eerder bestudeerde literatuur, negatief samenhangt met de toetsscores. Wellicht speelt de beperkte en enigszins selectieve steekproef ook in dit resultaat een rol. Ook is er bij de set opgaven van de subschaal 'recollection' geen positieve samenhang tussen kilometrage en het herinneren van de situatie, zoals wél uit eerder onderzoek van Randel & Pugh (1996) bleek.

Overige persoonlijke achtergrondvariabelen, zoals geslacht en autotype vertonen geen significante samenhang met de toetsscores. Ook hier geldt dat wanneer de steekproef groter zou zijn de mate van samenhang met persoonlijke achtergrondvariabelen beter onderzocht kan worden.

5.3 Beperkingen onderzoek en toets

Het onderzoek kende een aantal algemene beperkingen. De constructie van de toets verliep moeizaam, waardoor dataverzameling veel later van start kon gaan dan gepland. Dit had tot gevolg dat de Toets Situatiebewustzijn in dit onderzoek niet afgenomen kon worden bij leerlingbestuurders van een rijsschool en ervaren bestuurders die rijvaardigheidstrainingen volgden, zoals oorspronkelijk beoogd. Dit maakt de steekproef minder representatief. Tevens kon geen gebruik gemaakt worden van een krachtiger extern criterium, namelijk een praktijkoordeel over de rijvaardigheid van de kandidaten. Het was beoogd de rijvaardigheid in de praktijk te meten met behulp van het in de inleiding genoemde Performance Assessment Rijvaardigheid (PAR). Dit kon in dit onderzoek niet en daarom kan vooralsnog geen krachtige uitspraak gedaan worden over de extrapoleerbaarheid van de toets.

De eerste onderzoeksvraag had betrekking op de afnameomstandigheden ofwel de afnamefase uit het model van Crooks en Kane (1996). Een beperking van het onderzoek naar de toets was dat het werd afgenomen bij een selecte steekproef. Alle medewerkers van Cito die eerder de Zelfevaluatie Rijgedrag gemaakt hadden, werden uitgenodigd voor de Toets Situatiebewustzijn. Het personeel van Cito is op een aantal variabelen niet representatief voor de populatie automobilisten, zeker wat betreft opleiding. Daarnaast waren de afnameomstandigheden niet gecontroleerd. Cito-medewerkers konden elk moment van de dag, wanneer het hen uitkwam, de toets maken. Hierdoor waren de toetscondities wellicht niet optimaal. Ook bleek een aantal respondenten, om verschillende redenen, niet in voldoende mate geconcentreerd te zijn. Het belang van voldoende aandacht werd eerder genoemd door Endsley (2000) en Durso & Gronlund (1999).

De Toets Situatiebewustzijn bevat alle instructies om de toets te maken, dus de toets kon zonder hulp gemaakt worden. Uit de resultaten van de vragenlijst bleek echter dat de instructie door enkele respondenten niet begrepen werd en dat sommige deelnemers vonden dat er onvoldoende oefenvragen waren. Tevens werd geen feedback ontvangen bij de oefenopgaven, waardoor het leereffect minder groot was en deelnemers niet altijd begrepen wat de bedoeling was. Ook dit is een beperking van de afnamefase.

De tweede onderzoeksvraag had betrekking op het antwoordmodel, oftewel de scoring van de toets, fase 2 van het model van Crooks en Kane (1996). De toets bestaat uit 40 items, voor elk opgaventype een set van 10 vragen. De score werd in percentages gegeven: per subschaal van een

opgaventype kon 100 procent gescoord worden. De eindscore werd in het onderzoek het resultaat van alle opgaven. Alle onderdelen hadden eenzelfde gewicht. De antwoordsleutel is vooraf opgesteld door verkeerskundigen van een verkeerskundig adviesbureau en de juiste antwoorden werden digitaal geregistreerd. De manier van scoren is bij alle items dichotoom. Voor 'recollection' en 'action selection' is dat niet nadelig, maar voor de opgaventypen 'hazard perception' en 'timed response' kan dit nadelig zijn in verband met de precisie van antwoorden. Door reactievermogen of onwennigheid bij deze laatst genoemde categorie kan de score negatief beïnvloed worden, terwijl de deelnemers in de werkelijkheid op het juiste moment zouden handelen.

De derde onderzoeksvraag had betrekking op de interne consistentie van de toets. Dat komt aan bod in de generalisatiefase van het model van Crooks en Kane (1996). Een beperking in deze fase was dat de steekproef een beperkte omvang had en niet volledig representatief was voor alle automobilisten. Mogelijk zijn de geschatte itemparameters hierdoor niet voldoende stabiel en enigszins vertekend.

De vierde en vijfde onderzoeksvraag hadden betrekking op samenhang van de toetsscores met externe criteria en persoonlijke achtergrondvariabelen. Een beperking bij deze onderzoeksvragen is, zoals eerder vermeld, het ontbreken van een krachtiger extern criterium, namelijk een praktijkoordeel over de rijvaardigheid van de kandidaten. Een deugdelijk oordeel kan gegeven worden met behulp van het Performance Assessment Rijvaardigheid, zoals dat is ontwikkeling in het samenwerkingsverband CitoDrive.

5.4 Verbetervoorstellen Toets Situatiebewustzijn

Naar aanleiding van de in de vorige paragrafen besproken onderwerpen zijn enkele verbetervoorstellen geformuleerd. Eerst zullen enkele algemene verbetervoorstellen worden besproken, hierna zullen verbetervoorstellen per opgaventype worden besproken.

5.4.1 Algemene verbetervoorstellen Toets Situatiebewustzijn

Instructieteksten

De instructieteksten in de toets moeten aangepast worden. In de huidige toets is volgens enkele deelnemers te veel instructietekst aanwezig, waardoor deelnemers met een minder goede leesvaardigheid nadeel kunnen ondervinden. Een suggestie is om de instructieteksten te laten inspreken, waardoor dit probleem ondervangen wordt.

Opgaven

De Toets Situatiebewustzijn bevat vier opgaventypen die situatiebewustzijn beogen te meten. Zoals bleek uit de literatuur (Gugerty, 1997; McKenna, Horswill & Alexander, 2006; Endsley, 2000) is dit een goede opzet om situatiebewustzijn te meten. Het is aan te raden per opgaventype meer opgaven dan het huidige opgavenaantal in de toets op te nemen. Aan de hand van de subschaal kan vervolgens een meer betrouwbare uitspraak gedaan kan worden over het situatiebewustzijn in het verkeer.

Voorbeeldopgaven

Vooral voor het opgaventype 'timed response' is het belangrijk dat deelnemers kunnen oefenen en kunnen wennen aan het vernieuwende opgaventype. Bij de set opgaven van het type 'recollection' komen twee soorten vraagstellingen voor. Om die reden is het bij dat opgaventype ook aan te raden om van elke soort vraagstelling een voorbeeldopgave te behouden in de toets. Bij de overige opgaventypen volstaat één voorbeeldfragment. De verkeerssituaties in de voorbeeldfragmenten kunnen duidelijker. In de huidige toets is de kwaliteit van de voorbeeldfragmenten lager dan de overige fragmenten in de toets. Daarbij is het van belang dat de deelnemer feedback krijg op het antwoord dat gekozen is. Wanneer het antwoord onjuist is, dient het correcte antwoord gepresenteerd te worden met daarbij de uitleg waarom dat het beste antwoord is. Dit geldt alleen voor de voorbeeldopgaven.

In de huidige toets krijgen de deelnemers eerst vijf voorbeeldopgaven. Daarna wordt voorafgaand aan elk blok de instructietekst voor dat blok nogmaals getoond. Hierdoor wordt deels voorkomen dat deelnemers bij het vierde opgavenblok vergeten zijn wat er moest gebeuren. Het is aan te raden om bij elk blok de instructietekst én een voorbeeldopgave/oefenopgave te tonen in plaats van alle voorbeeldopgaven voorafgaand aan de toets te laten maken.

Ten slotte geldt ook voor de instructie bij voorbeeldopgaven dat de teksten ingesproken moeten worden, zodat voor deelnemers met een minder goede leesvaardigheid de instructie ook duidelijk is.

Videofragmenten

De videofragmenten in de beproefde toets zijn niet voldoende duidelijk, vanwege de kleine afmetingen en de te lage resolutie. In werkelijkheid is de voorruit van de auto veel groter. Ook is naar verhouding de achteruitkijkspiegel te groot ten opzichte van de voorruit. Op dit moment worden de videofragmenten omgezet naar een groter formaat. De fragmenten zullen hierdoor een stuk duidelijker worden. In een eventuele nieuwe versie kan overwogen worden om opnieuw videofragmenten te filmen. Ook is het van belang dat in de toekomst meer verkeerssituaties in beeld gebracht worden, zodat de toets een grotere variatie aan verkeerssituaties bevat. Dit geldt vooral voor het opgaventype van de subschaal 'recollection'. Tevens kan het uitbreiden van videofragmenten op lange termijn resulteren in een itemdatabank, zodat de toets verschillende keren gemaakt kan worden en daarbij niet dezelfde opgaven worden gebruikt.

Zijspiegels worden door een aantal deelnemers gemist in de videofragmenten. In de toekomst is het wellicht verstandig om zijspiegels toe te voegen. De deelnemer krijgt daardoor een zo compleet mogelijk beeld van de situatie, waardoor de deelnemer de situatie mogelijk beter kan beoordelen. Een mogelijk nadeel daarvan is dat de aanwezigheid van zijspiegels zou kunnen zorgen voor afleiding. In een authentieke situatie wordt alleen in de zijspiegel gekeken wanneer dat nodig is, maar in de toets zouden de zijspiegels permanent de aandacht kunnen trekken. Vervolgonderzoek met opnames vanuit zijspiegels zou hierover uitsluitel kunnen geven.

Verbetervoorstellen constructiefase

De huidige antwoordsleutel beschrijft alleen de juiste antwoorden. Hier worden geen argumenten bij gegeven. Dit wordt ook niet gedaan bij de onjuiste antwoordalternatieven. Het is aan te raden om bij de constructie van nieuwe opgaven, waar nuttig, alle antwoordalternatieven te voorzien van argumentaties, zodat verantwoord wordt waarom een vraag juist of onjuist is. Wellicht kan dit gerealiseerd worden met behulp van een expertbijeenkomst voor instructeurs, examinatoren en verkeersvaardigheidstrainers.

5.4.2 Verbetervoorstellen opgaventype 'recollection'

De subschaal met 'recollection'-opgaven was gezien het kleine aantal opgaven voldoende betrouwbaar, maar deelnemers gaven aan dat na een paar dezelfde opgaven extra gelet werd op verkeersborden met maximumsnelheid. Het is daarom van belang dat er een grotere variatie aan verkeerssituaties verwerkt wordt in de videofragmenten die bij dit type opgaven getoond worden. Daarnaast is het belangrijk dat een videofragment niet te duidelijk toont wat het goede antwoord moet zijn. Tevens is het van belang dat alle antwoordmogelijkheden plausibel zijn. In de onderzochte toets zijn er twee opgaven waarbij twee van de vier antwoordalternatieven niet gekozen zijn. In beide gevallen zijn wegwerkzaamheden en moet met een aangepaste snelheid gereden worden, wat duidelijk aangegeven staat. Ook hier kan gedacht worden aan meer variatie in videofragmenten.

5.4.3 Verbetervoorstellen opgaventype 'action selection'

De subschaal met 'action selection'-opgaven was gezien het kleine aantal opgaven matig betrouwbaar. De verkeerssituaties zijn voldoende gevarieerd in vergelijking met het opgaventype 'recollection'. Het is echter aan te raden om situaties uit enkele onduidelijke videofragmenten in een nieuw, duidelijker videofragment te verwerken. Daarnaast zullen antwoordalternatieven duidelijker geformuleerd moeten worden. 'Remmen' en 'Gas los' zijn begrippen die verschillend te interpreteren zijn. Tevens komen enkele videofragmenten voor waarbij veel twijfel over de meest correcte oplossing is. Dit hoeft geen probleem te zijn, omdat in het werkelijke verkeer ook veel twijfelmomenten zijn: de veiligste optie moet dan gekozen worden. Het is bij 'twijfelopgaven' wel belangrijk dat een duidelijke argumentatie beschreven is bij het correcte antwoord in de antwoordsleutel.

5.4.4 Verbetervoorstellen opgaventype 'hazard perception'

De analyse van de subschaal met 'hazard perception'-opgaven was beperkt. Aan te raden is om de coördinaten van de deelnemers te registreren, zodat bekeken kan worden waar de deelnemers geklikt hebben. Dit is vooral belangrijk wanneer buiten het 'hotspotgebied' geklikt is.

Daarnaast is gebleken dat opgaven van dit type veelal correcte antwoorden oplevert. De verkeerssituaties kunnen daarom wellicht gevarieerder en complexer. Tevens krijgen de deelnemers onbeperkt het laatste beeld van het videofragment te zien, zodat lang nagedacht kan worden wat als het meest potentiële gevaar wordt beschouwd. In werkelijkheid is de verkeerssituatie dynamisch, in dit geval niet. Het is daarom van belang dat het laatste beeld niet eindeloos te zien blijft, maar na enkele seconden verdwijnt. Daarbij kan de hoeveelheid tijd die de deelnemer nog heeft om de opgave te

beantwoorden in beeld gebracht worden. De deelnemer weet op deze manier dat snel een keuze moet worden gemaakt, zoals in werkelijkheid ook het geval is.

5.4.5 Verbetervoorstellen opgaventype 'timed response'

De betrouwbaarheid van dit opgaventype is erg laag. De tijdsfactor in de opgaven is een zeer lastige factor. Om ervoor te zorgen dat de betrouwbaarheid toeneemt worden hier enkele verbeterpunten genoemd. Allereerst vergt dit specifieke type opgaven meer oefening. Het is daarom aan te raden meer oefenopgaven in de toets op te nemen, zodat voor de deelnemers geheel duidelijk is hoe dit type opgaven beantwoord dient te worden. Het is dan aannemelijk dat opgaven minder snel foutief beantwoord worden door onwennigheid met het opgaventype.

Daarnaast is het van belang dat de tijdsduur van het videofragment weergegeven wordt. Deelnemers kunnen dan zien hoe lang het fragment nog duurt en wellicht wordt hiermee voorkomen dat deelnemers zo snel mogelijk klikken. Onderzocht zou kunnen worden of het nogmaals afspelen van het videofragment aan te raden is: in een werkelijke verkeerssituatie moet direct gereageerd worden en bestaat er geen tweede kans. Om deelnemers een beter beeld van de situatie te laten krijgen kan ook gekozen worden voor langer durende videofragmenten. Eventueel kunnen in het langere videofragment meerdere kritische situaties aan bod komen, zoals bij de Hazard Perception Test, die in Engeland deel uitmaakt van het theorie-examen. Dit is mogelijk wanneer het filmpje bij een eerste klik tegelijkertijd op pauze springt en weer verder gaat bij een tweede klik, waarna zich een nieuwe kritische situatie aandient.

Tevens geldt voor dit opgaventype, net als bij het opgaventype 'action selection', dat de begrippen 'remmen' en 'gas los' duidelijk moeten zijn voor de deelnemers.

5.5 Implicaties voor vervolgonderzoek

In voorgaande paragraaf zijn verbetervoorstellen gedaan voor de Toets Sitatiebewustzijn. Dit heeft implicaties voor vervolgonderzoek. Natuurlijk zullen, indien haalbaar, de eerder genoemde verbetervoorstellen doorgevoerd moeten worden om een meer betrouwbare en valide score toe te kunnen kennen aan de deelnemers. Een aantal punten dat van belang is voor het vervolgonderzoek zal hieronder worden besproken.

In het vervolgonderzoek zullen leerlingbestuurders en ervaren leaserijders deelnemen aan de Toets Sitatiebewustzijn. Dit maakt de steekproef meer representatief voor de populatie automobilisten. Bij de nieuwe afnamen van de toets, zullen de afnameomstandigheden gestandaardiseerd moeten worden. Bij voorkeur dient de afname van de toets gecontroleerd plaats te vinden, zodat zo correct mogelijk de vaardigheden van de deelnemer in beeld worden gebracht en de deelnemer zo min mogelijk wordt afgeleid.

Tevens is het van groot belang dat de toetsscores vergeleken worden met het praktisch rijgedrag van de deelnemers, om een krachtige uitspraak te kunnen doen over de extrapoleerbaarheid van de toets. Dit kan met behulp van het Performance Assessment Rijvaardigheid (PAR).

In het onderzoek, waarvan hier verslag wordt gedaan, is aandacht besteed aan verschillende fasen uit het validiteitsmodel van Crooks en Kane (1996). De afnamefase, scoringfase,

generalisatiefase en extrapolatiefase. Een aantal fasen uit het validiteitsmodel zijn niet onderzocht. In vervolgonderzoek zal aandacht besteed moeten worden aan met name de beslissingfase en aan de impact.

Daarnaast zal onderzocht moeten worden in hoeverre de prestatie van de deelnemer juist geïnterpreteerd wordt. In de evaluatiefase van het validiteitsmodel wordt namelijk gekeken naar de interpretatie van toetsresultaten. Daarnaast kan in die fase nader onderzocht worden of de toets daadwerkelijk toetst hoe iemand in de praktijk functioneert. Dit kan onderzocht worden met een Performance Assessment Rijvaardigheid. Pas wanneer deze zaken duidelijk zijn kan een beslissing aan de toets worden gekoppeld.

Ook dient in vervolgonderzoek aandacht te worden besteed aan de impact die de toets heeft. De impact van een toets kan positief, maar ook negatief zijn. Het is mogelijk dat een deelnemer erachter komt dat zijn rijgedrag niet zo goed is als hij dacht. Leidt dit tot langdurige onzekerheid op de weg, dan heeft de toets een negatieve impact op de deelnemer. Een voorbeeld van een positieve impact is dat het rijgedrag op basis van de resultaten verbeterd zal worden door een rijvaardigheidstraining en onzekerheid afneemt. Interessant is om te onderzoeken hoe groot de impact van de toets kan zijn, zowel positief als negatief.

Ten slotte is het interessant om te kijken naar de verschillen tussen resultaten op de opgavenvorm in deze toets (dynamisch videomateriaal) en de resultaten op de opgavenvorm in het theorie-examen van het CBR (statisch fotomateriaal).

Wanneer in vervolgonderzoek aandacht zal worden besteed aan bovengenoemde punten, wordt verwacht dat de Toets Situatiewaarschuwing een in hogere mate betrouwbaar en valide meetinstrument voor het construct situatiewaarschuwing in het verkeer zal zijn.

Op basis van de toets en de overige meetinstrumenten van het samenwerkingsverband CitoDrive, kunnen bestuurders met een verhoogd risico op verkeersongevallen, zoals jongeren en leaserijders, geselecteerd worden voor rijvaardigheidstrainingen. Wanneer de instrumenten bijdragen aan een verhoogde verkeersveiligheid, zouden de toets en het Performance Assessment Rijvaardigheid ingezet kunnen worden bij rijopleidingen. Bijvoorbeeld om te beoordelen of de kandidaat examenrijp is of vaardig genoeg is voor een tussentijdse toets. Uit de literatuur blijkt immers dat de hogere orde vaardigheden zoals gevaarherkenning, inzicht in verkeerssituaties en tijdig anticiperen op gevaar een belangrijke rol spelen bij besturing van een voertuig (Hattaka, Keskinen, Gregersen, Glad & Hernetkoski, 2002). Wanneer de resultaten op de instrumenten aantonen dat het gebruik van de instrumenten de verkeersveiligheid ten goede komt, zou men ook kunnen overwegen het belang van langdurende verkeerseducatie na behalen van het rijbewijs (Roelofs et al., 2008) nader te onderzoeken. De verkeersveiligheid kan dan blijvend worden verhoogd.

6. Literatuur

- CBR (2007). *Het rijexamen verandert per 1 januari 2008*. Binnengehaald op 15 januari 2008 van <http://rijbewijs.cbr.nl/pdf/LealetOVR.pdf>.
- Crooks, T. J. & Kane, M. T. (1996). Threats to the valid use of assessments. *Assessment in Education: Principles, Policy & Practice*, 3 (3), 265-285.
- Durso, F. T. & Gronlund, D. (1999). Situation Awareness. In F. T. Durso, R. Nickerson, R. Schvaneveldt, S. Dumais, M. Chi & S. Lindsay (Eds.), *The Handbook of Applied Cognition* (pp. 2-30), Chichester: Wiley.
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Human Factors Society 32nd Annual Meeting* (pp. 97-101), Santa Monica, CA.
- Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37 (1), 32-64.
- Endsley, M. R. (2000). Theoretical Underpinnings of Situation Awareness: A Critical Review. In M. R. Endsley & D. J. Garland (Eds.), *Situation Awareness Analysis and Measurement* (pp.3-32). Mahwah, New York: Lawrence Erlbaum Associates.
- Gugerty, L. J. (1997). Situation Awareness During Driving: Explicit and Implicit Knowledge in Dynamic Spatial Memory. *Journal of Experimental Psychology*, 3 (1), 42-66.
- Haertel, E.H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement 4th edition*. (pp. 65-110). Westport, CT: American Council on Education and Praeger Publishers.
- Hatakka, M., Keskinen, E., Gregersen, N. P., Glad, A. & Hernetkoski, K. (2002). From control of the vehicle to personal self-control; broadening the perspectives to driver education. *Transportation Research Part F*, 5 (3), 201-215.
- Horswill, M. S. & McKenna, F. P. (2004). Divers' Hazard Perception Ability: Situation Awareness on the Road. In S. Banbury & S. Tremblay (Eds.), *A Cognitive Approach to Situation awareness: Theory and Application* (pp. 155-175). Aldershot, United Kingdom: Ashgate Publishing Limited.
- Inspectie Verkeer en Waterstaat & Rijkswaterstaat (2008). *Veiligheidsbalans 2007*. Binnengehaald 15 januari 2008 van www.verkeerenwaterstaat.nl.
- Kane, M. T. (2006). Validation. In R. L. Brennan (ed.), *Educational Measurement 4th edition*. (pp. 17-64). Westport: American Council on Education and Praeger Publishers.
- McKenna, F. P., Horswill, M. S. & Alexander, J. L. (2006). Does Anticipation Training Affect Drivers' Risk Taking? *Journal of Experimental Psychology: Applied*, 12 (1), 1-10.

- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement 3rd edition*. (pp. 13–103). New York: American Council on Education and Macmillan.
- Moelands, H. A., & Sanders, P. F. (1996). Onderwijskundige meetinstrumenten. In: J. Scheerens (Ed.). *Kwaliteitszorg in het onderwijs*. (pp. 51-73). Alphen aan den Rijn: Samson H.D. Tjeenk Willink.
- O'Brien, K. S. & O'Hare D. (2007). Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics*, 50 (7), 1064-1091.
- Randel, J. M. & Pugh, H. L. (1996). Differences in expert and novice situation awareness in naturalistic decision making. *International Journal Human-Computer Studies* 45, 579-597.
- Roelofs, E., Van Onna, M., Nägele, R., Mesken, J., Kuiken, M. & Cozijnsen, E. (2008). Development of Multimedia Tests for Responsive Driving. In: Dom, L. (Ed.) *Driver Behaviour and Training Volume III*. (pp. 251-264). Hampshire: Ashgate Publishing Limited.
- Rothengatter, T. (1997). Psychological Aspects of Road User Behaviour. *Applied Psychology: An International Review*, 46 (3), 223-234.
- Stanners, M. & French, H. T. (2005). *An Empirical Study of the Relationship Between Situation Awareness and Decision Making*. Edinburgh South Australia: DSTO Systems Sciences Laboratory.
- Vissers, J. & Mesken, J. (2008). *Componenten van het vernieuwde praktijkexamen B: een pilotstudie*. Abstract ORD 2008.

Bijlage A Gegevens over subjecten

Tabel 1

Opleidingsniveau

Opleiding	Frequentie	Percentage
Basisschool	1	1,2
VMBO (mavo/vbo)	2	2,5
Havo	7	8,6
VWO	5	6,2
MBO	13	16,0
HBO	22	27,2
WO	30	37,0
Overig	1	1,2
Totaal	81	100,0

Tabel 2

Slechtziend

Slechtziend	Frequentie	Percentage
Ja	7	8,6
Nee	74	91,4
Totaal	81	100,0

Tabel 3

Kleurenblind

Slechtziend	Frequentie	Percentage
Ja	3	3,7
Nee	78	96,3
Totaal	81	100

Tabel 4

Ervaring met spelcomputer

Ervaring	Frequentie	Percentage
Nee, vrijwel nooit	46	56,8
Ja, maar niet vaak	26	32,1
Ja, ervaren speler	9	11,1
Totaal	81	100

Tabel 5

Autotype

Type auto	Frequentie	Percentage
Auto	62	76,5
Stationwagon	17	21,5
Bus	0	0
Totaal	79	97,5

Tabel 6

Ongevalbetrokkenheid

Betrokkenheid	Frequentie	Percentage
Ja	15	18,5
Nee	64	79
Totaal	79	97,5

Bijlage B Draaiboek filmen videofragmenten

Onderstaand draaiboek is overgenomen van de toetsconstructeurs en is gebruikt bij het filmen van de videofragmenten.

Materialen

- Minimaal twee digitale videocamera's + benodigdheden (extra accu's, tapejes, etc). Kies in ieder geval voor de vóórcamera een kwalitatief zeer goede consumentencamera, (bijvoorbeeld de Panasonic NV-GS 230). Te overwegen is om semi-professionele camera's te gebruiken, maar dan zal opnieuw moeten worden getest hoe dat in de auto werkt en hoe deze bevestigd moeten worden (kan niet met zuignapstatief).
- Statieven: bij voorkeur zuignapstatieven (Manfrotto autostatief, 2 beschikbaar bij DHV, ook te bestellen bij de betere foto/video handel). Er is weinig trilling en de camera kan precies in de juiste positie worden geplaatst. Een normaal statief kan ook, maar dat staat niet stabiel (in de kofferbak of voor de passagierstoel) en zal dus extra bevestigd moeten worden. Er is hiervoor een grotere auto nodig, bij voorkeur een automaat.
- Een auto met niet al te schuine voor – en achterrait
- iets om een geluidssignaal te produceren (voor de eerste sessie is een simpel *personal alarm* gebruikt dat een hard geluid geeft als er op een knop wordt gedrukt).

Vorbereiding

Omschrijf de wegtypen waar je wilt rijden, en de complexiteit van de verkeerssituatie. Dat levert een tabel op waar in de cellen de aantallen items kunnen worden weergegeven. Een meer gedetailleerde uitwerking naar scenario's is niet zinvol omdat het veel te onvoorspelbaar is wat je tegenkomt.

Voorbeeld:

	ETW 30	ETW50	ETW60	GOW80	SW100	SW120
Eenvoudig	3	3	3	3	3	3
Gemiddeld	3	3	3	3	3	3
Complex	3	3	3	3	3	3

ETW = Erftoegangsweg

GOW = Gebiedsontsluitingsweg

SW = Stroomweg

De cijfers geven de maximumsnelheid weer.

Aan de hand van een dergelijke soort tabel kan achteraf gecheckt worden of de videofragmenten gelijkmatig verdeeld zijn over de verschillende wegtypen.

Draaien

De beste dagen om te filmen zijn droge, bewolkte dagen.

Ga altijd met twee personen draaien zodat een persoon op de achterbank de achtercamera kan bedienen.

Zorg ervoor (in de instellingen van de camera) dat de teller van de band doorloopt na elke keer stoppen.

Zorg ervoor dat de beeld- en geluidsinstellingen op beide camera's gelijk zijn.

Zorg ervoor dat de positie van de voorcamera zo dicht mogelijk het gezichtspunt van de bestuurder benadert. Direct onder de binnenspiegel is een goede plaats. Plaats de achtercamera zo dicht mogelijk tegen de achterrait. Hierdoor heb je zo weinig mogelijk last van de achterrait verwarming (een auto zonder achterraitverwarming is uiteraard beter). Eventueel kun je de camera iets inzoomen.

Indien nodig kunnen beide camera's ondersteund worden met een extra stang die bij het statief hoort, om trillingen op te vangen. Check het beeld en zet dan de camera's goed vast. Klap tijdens het draaien de schermpjes dicht, dit scheelt energie.

Tijdens het rijden kun je zelf bepalen wanneer je de camera's wilt laten draaien of niet. Zorg ervoor dat je elk shot (of dat nu 30 seconden of 8 minuten duurt) beëindigt met een geluidssignaal. Dit is nodig om de beelden van de voor – en achtercamera achteraf synchroon te kunnen leggen. Het is niet verstandig om de camera's continu te laten draaien, omdat dan de accu snel op is en het bandje snel vol. Spreek met elkaar af wanneer je de camera start zodat de twee camera's zoveel mogelijk gelijk lopen.

De persoon op de achterbank dient bij te houden wat er gebeurt: zowel de begin- en eindtijd van elke take, als een karakterisering van de omgeving en een beschrijving van gebeurtenissen die mogelijk zinvol zijn. Daarvoor kan een formulier worden gemaakt, bijvoorbeeld

Datum

Take	Kloktijd		Inhoud
	start	eind	
1	9.11	+/-9.19	- fietskar inhalen (moeder+kind) - rotonde met 2 jongens die aanfietsen en op de rotonde rijden (beide einde film)
2	9.30		9.33 oranje stoplicht Waarschijnlijk prut
3	9.39	9.42	- Vrachtwagens + fietser (binnenstad A'foort) - (begin film)
4	9.45	9.52	(achter vrachtwagen, provinciale weg) - evt. auto rotonde (geeft voorrang aan fietser) - vrachtwagen verleent voorrang
5	9.58	+/-10.02	- +/- 10.00 busje uit uitrit
6	+/- 10.41	10.47	(bomenweg) - evt. einde film: auto geeft voorrang
7	10.49	10.51	- (begin take) evt. vrachtwagens - (eind take) rijdt achter vrachtwagen, wil deze inhalen, maar wordt zelf ingehaald (binnenspiegel)
8	10.55	+/-10.57	- invoegen - 10.56 inhalen caravan, vrachtwagens linkerstrook
9	11.06	11.11	Klaverbladje rijden
10	11.13	+/-11.19	(invoegen) - 11.16 uitvoegen - 11.17 invoegen - Wegenwacht
11	11.34	11.37	Drukke snelweg - 11.35 iemand voegt voor ons in (knipperlicht) - 11.35 hard remmen (remlicht)
12	11.47	+/-11.58	(bomenweg+ weg langs heide) - 11.51 voorligger remt ivm afslaan - 11.53 wegwerkzaamhedenbord - 11.54 vrachtwagen op weg geparkeerd - 11.55 fietsers (laan 1914)

Inladen

Bekijk al het materiaal door het op een pc in te laden (met usb kabel of firewirekabel) of door de camera rechtstreeks op een tv aan te sluiten. Zorg ervoor dat de teller van de band meeloopt in het

beeld. Selecteer de fragmenten die ingeladen moeten worden in het montageprogramma. Geef daarbij de begin- en eindtijd van de bandteller weer, van zowel de voor- als achtercamera. Dit kan bijvoorbeeld in het volgende format:

Take 1: kloktijd 9:11:34 – 9:17:58		
	Start	Eind
Camera voor	0:00:00:00	0:06:27:06
Camera achter	0:00:00:00	0:06:38:24
Events in take 1		
	Start	Eind
Brommer op rotonde	0:04:00:14	0:04:23:07
Fietskar	0:05:10:13	0:05:25:13

Take 2: kloktijd 9:30:13 - 09:35:31		
	Start	Eind
Camera voor	0:06:34:20	0:11:53:03
Camera achter	0:06:47:04	0:12:05:07
Events in take 2		
	Start	Eind
Verkeerslicht op oranje	0:09:37:11	0:09:47:20

Op basis van deze gegevens kan degene die de montage doet, makkelijk de fragmenten selecteren. Dat kan op basis van de tellernummers van de band. De kloktijd is een extra hulpmiddel (in dit voorbeeld is de kloktijd van de voorcamera genomen) maar is niet echt noodzakelijk. De fragmenten moeten worden ingenomen **tot en met** het geluidssignaal.

Monteren

Monteren kan in diverse montageprogramma's: Avid, Final Cut Pro (Apple), Pinnacle. Windows Moviemaker is niet geschikt, omdat daarin niet twee videobeelden over elkaar heen kunnen worden gelegd. Eerst moeten de takes worden ingeladen. Dan moet van de hele take het beeld van voor- en achtercamera synchroon worden gelegd op basis van het geluidssignaal. Als de beelden synchroon liggen, kan het beeld van de achtercamera als inzet (als een soort binnenspiegel) worden gemonteerd in het beeld van de voorcamera. Het is aan te raden om hier een dikker zwart kader omheen te zetten, zodat goed duidelijk is dat het om een ander beeld gaat. Let op: er wordt hier een spiegel gesimuleerd, dat betekent dat het beeld van de achtercamera gespiegeld dient te worden. Pas als de hele take op deze wijze gemonteerd is, kunnen de events eruit gesneden worden en apart als filmpjes worden opgeslagen. Dat gebeurt op basis van de teller van de voorcamera.

Bijlage C Vragenlijst

 zeker weten.

Definitief Vragenlijst Toets Situatiewustzijn (Cito, digitale versie)

Naar aanleiding van de 'Toets Situatiewustzijn' die u eerder gemaakt heeft, willen wij u een aantal vragen stellen. Het doel van deze vragen is dat wij inzicht krijgen in uw ervaringen met de gemaakte toets. Wij hechten veel waarde aan uw mening. Op basis van de uitkomsten kan bekeken worden op welke punten de toets eventueel verbeterd zou kunnen worden. Uw antwoorden op de vragen worden anoniem verwerkt, uw naam wordt alleen gebruikt om de vragenlijst te kunnen koppelen aan gemaakte keuzes.

De vragenlijst bestaat hoofdzakelijk uit vragen waarbij u het antwoord direct kunt aanvinken. U wordt verzocht alle vragen met een sterretje volledig te beantwoorden. Hierdoor krijgen wij het beste beeld van uw ervaringen met de toets.

Wij verzoeken u de vragenlijst in te vullen vóór 15 april 2008.
Bij voorbaat dank voor uw medewerking!

[Verder >>](#)

 zeker weten.

Definitief Vragenlijst Toets Situatiewustzijn (Cito, digitale versie)

*1. Vul hieronder uw voor- en achternaam in:

*2. Vul hieronder uw gebruikersnaam in die u aan het begin van de toets heeft ingevuld om in te loggen:

*3. Wat is uw leeftijd?

*4. Geslacht

☐ Man

☐ Vrouw

[<< Terug](#) [Verder >>](#)

Definitief Vragenlijst Toets Sitatiebewustzijn (Cito, digitale versie)

***5. Wat is uw hoogst afgeronde opleiding?**

- | | |
|--|---------------------------|
| ☐ Geen | ☐ VWO |
| ☐ Basisschool | ☐ MBO |
| ☐ VMBO (vbo/mavo) | ☐ HBO |
| ☐ HAVO | ☐ WO |
| ☐ Overig (vul antwoord in) | |

Anders, namelijk:

***6. Ik zie slecht.**

- ☐ Ja
☐ Nee

***7. Ik ben kleurenblind.**

- ☐ Ja
☐ Nee

***8. Speelt u spelletjes op de (spel)computer?**

- ☐ Ja, ik ben inmiddels een ervaren speler.
☐ Ja, maar niet vaak.
☐ Nee, vrijwel nooit.

<< Terug

Verder >>

Definitief Vragenlijst Toets Sitatiebewustzijn (Cito, digitale versie)

*9. Wanneer heeft u uw rijbewijs B behaald?

Vul de datum in
waarop u uw
rijbewijs heeft
behaald

DD MM YYYY

/ /

*10. Heeft u nog een ander rijbewijs in uw bezit?

- ☐ Rijbewijs A (motor)
- ☐ Rijbewijs BE (personenauto met aanhangwagen)
- ☐ Rijbewijs C (vrachtwagen)
- ☐ Rijbewijs CE (vrachtwagen met aanhangwagen)
- ☐ Rijbewijs D (bus)
- ☐ Rijbewijs DE (bus met aanhangwagen)
- ☐ Nee, ik heb alleen rijbewijs B (personenauto).

*11. Hoeveel kilometer rijdt u per jaar?

(vul het getal in, zonder punten, bijvoorbeeld 155000)

Aantal km

*12. In welk type voertuig rijdt u doorgaans?

- ☐ Auto
- ☐ Stationwagon (verlengde auto)
- ☐ Bus

<< Terug

Verder >>

Definitief Vragenlijst Toets Situatiewustzijn (Cito, digitale versie)

*13. Bent u, in de afgelopen vier jaar, betrokken geweest bij een auto-ongeluk of een schadeclaim?

- ☐ Ja
☐ Nee

14. Zo ja, geef dan aan hoe vaak en welk type ongeluk(ken)?

	1 keer	2 keer	3 keer	4 keer	5 keer of meer
Parkeerschade	☐	☐	☐	☐	☐
Door een andere auto aan voorzijde geraakt	☐	☐	☐	☐	☐
Zelf een andere auto aan voorzijde geraakt	☐	☐	☐	☐	☐
Door een andere auto aan achterzijde geraakt	☐	☐	☐	☐	☐
Zelf een andere auto aan achterzijde geraakt	☐	☐	☐	☐	☐
Controle verlies	☐	☐	☐	☐	☐
Diefstal / vandalisme	☐	☐	☐	☐	☐
Stilstaande objecten geraakt	☐	☐	☐	☐	☐
Anders namelijk:					

<< Terug

Verder >>

Definitief Vragenlijst Toets Situatiewustzijn (Cito, digitale versie)

*15. De toets

	Helemaal Oneens	Beetje Oneens	Beetje Eens	Helemaal Eens
Het doel van de toets was duidelijk	☐	☐	☐	☐
De geschreven instructie voorafgaand aan de toets was helder	☐	☐	☐	☐
Bij elke vraag was het duidelijk wat ik moest doen	☐	☐	☐	☐
De filmpjes waren duidelijk	☐	☐	☐	☐
De indeling van het beeldscherm was overzichtelijk	☐	☐	☐	☐
Ik had voldoende tijd om elke vraag in te vullen	☐	☐	☐	☐

*16. Geef een schatting van het aantal vragen in de toets waarbij u onvoldoende tijd had om te antwoorden.

<< Terug

Verder >>

Definitief Vragenlijst Toets Sitatiebewustzijn (Cito, digitale versie)

*17. De toets

	Helemaal Oneens	Beetje Oneens	Beetje Eens	Helemaal Eens
Ik vond het leuk om de toets te maken	☐	☐	☐	☐
Ik vond de toets leerzaam	☐	☐	☐	☐
Ik kon me goed concentreren tijdens de toets	☐	☐	☐	☐

18. Indien u zich niet goed kon concentreren kunt u hier de oorzaak daarvan in steekwoorden weergeven.

*19. De toets

	Helemaal oneens	Beetje oneens	Beetje eens	Helemaal eens
Ik was onreken om goed te presteren	☐	☐	☐	☐

<< Terug

Verder >>



maket weten.

Definitief Vragenlijst Toets Sitatiebewustzijn (Cito, digitale versie)

20. Als u nog opmerkingen heeft over de toets kunt u ze in onderstaand testvak noteren.

Dit is het einde van de vragenlijst.
Nogmaals bedankt voor uw medewerking!

<< Terug

Klaar >>

Bijlage D Itemkarakteristieken gereviseerde toets

Tabel 1

Itemkarakteristieken van de Gereviseerde Toets ($\alpha = .65$)

Itemnummer	Standaard deviatie	P-waarde	Std. error of the mean
RC1	,329	,88	,036
RC2	,217	,95	,024
RC3	,329	,88	,036
RC5	,496	,59	,055
RC6	,329	,88	,036
RC7	,485	,63	,054
RC8	,367	,84	,041
RC9	,241	,94	,027
RC10	,315	,89	,036
AS11	,502	,54	,055
AS13	,452	,72	,050
AS16	,189	,96	,021
AS17	,343	,87	,038
AS19	,502	,46	,055
AS20	,399	,80	,044
HP21	,311	,89	,036
HP23	,501	,55	,058
HP24	,369	,84	,043
HP25	,356	,85	,041
HP26	,226	,95	,026
HP27	,327	,88	,038
HP28	,327	,88	,038
HP29	,421	,77	,049
HP30	,115	,99	,013
TR33	,493	,60	,057
TR34	,421	,23	,049
TR35	,311	,89	,036
TR37	,503	,49	,058
TR39	,503	,52	,058
TR40	,503	,52	,058

Bijlage E Statistieken Toets Sitatiebewustzijn

E1 Samenhang tussen de waardering van afnamecondities en de toetsscores

Onderstaande tabellen geven beschrijvende statistieken en samenhang weer. Gerekend is met de gecorrigeerde toets, waardoor de maximumscore bij het opgaventype 'recollection' 9 is en de maximum totaalscore 30 is.

Tabel 1

Het doel van de toets was duidelijk

	N	%	Geldig %	Cumulatief %
Helemaal oneens	4	4,0	5,3	5,3
Beetje oneens	7	7,0	9,2	14,5
Beetje eens	19	19,0	25,0	39,5
Helemaal eens	46	46,0	60,5	100,0
Total	76	76,0	100,0	--
Missing System	24	24,0	--	--
Totaal	100	100,0	--	--
Gemiddelde	2,4	--	--	--
Standaarddev.	0,9	--	--	--

Tabel 2

Correlatie duidelijkheid doel toets met toetsscore

		Doel	Totaalscore
Doel	Pearson Correlatie	1	-,090
	Sig. (2-tailed)		,480
	N	67	64

Tabel 3

De geschreven instructie voorafgaand aan de toets was helder

Oordeel	N	%	Geldig %	Cumulatief %	Gemiddelde	SD
Helemaal oneens	4	4.0	5.3	5.3	--	--
Beetje oneens	5	5.0	6.6	11.8	--	--
Beetje eens	20	20.0	26.3	38.2	--	--
Helemaal eens	47	47.0	61.3	100.0	--	--
Totaal	76	76.0	100.0	--	--	--
Missing	24	24.0	--	--	--	--
Totaal	100	100.0	--	--	2.5	0.8

Tabel 4

Correlatie instructie toets met toetsscore

		Instructie	Totaalscore
Instructie	Pearson Correlatie	1	-,090
	Sig. (2-tailed)	--	,480
	N	67	64

Tabel 5

Bij elke vraag was het duidelijk wat is moest doen

Oordeel	N	%	Geldig %	Cumulatief %	Gemiddelde	SD
Helemaal oneens	7	7.0	9.2	9.2	--	--
Beetje oneens	11	11.0	14.5	23.7	--	--
Beetje eens	16	16.0	21.1	44.7	--	--
Helemaal eens	42	42.0	55.3	100.0	--	--
Totaal	76	76.0	100.0	--	--	--
Missing	24	24.0	--	--	--	--
Totaal	100	100.0	--	--	2.2	1.0

Tabel 6

Correlatie duidelijkheid van de vraag met toetsscore

		Duidelijkheid vraag	Totaalscore
Duidelijkheid vraag	Pearson Correlatie	1	,169
	Sig. (2-tailed)	--	,183
	N	67	64

Tabel 7

De filmpjes waren duidelijk

Oordeel	N	%	Geldig %	Cumulatief %	Gemiddelde	SD
Helemaal oneens	10	10.0	13.2	13.2	--	--
Beetje oneens	34	34.0	44.7	57.9	--	--
Beetje eens	20	20.0	26.3	84.2	--	--
Helemaal eens	12	12.0	15.8	100.0	--	--
Totaal	76	76.0	100.0	--	--	--
Missing	24	24.0	--	--	--	--
Totaal	100	100.0	--	--	1.45	0.9

Tabel 8

Correlatie duidelijkheid filmpjes met totaalscore

		Duidelijkheid filmpje	Totaalscore
Duidelijkheid filmpje	Pearson Correlatie	1	-.028
	Sig. (2-tailed)	--	.829
	N	76	64

Tabel 9

De indeling van het beeldscherm was overzichtelijk

Oordeel	N	%	Geldig %	Cumulatief %	Gemiddelde	SD
Helemaal oneens	9	9.0	11.8	11.8	--	--
Beetje oneens	14	14.0	18.4	30.3	--	--
Beetje eens	35	35.0	46.1	76.3	--	--
Helemaal eens	18	18.0	23.7	100.0	--	--
Totaal	76	76.0	100.0	--	--	--
Missing	24	24.0	--	--	--	--
Totaal	100	100.0	--	--	1.8	0.9

Tabel 10

Correlatie indeling beeldscherm met totaalscore

		Indeling	Totaalscore
Indeling	Pearson Correlatie	1	.226
	Sig. (2-tailed)	--	.073
	N	67	64

Tabel 11

Ik had voldoende tijd om elke vraag in te vullen

Oordeel	N	%	Geldig %	Cumulatief %	Gemiddelde	SD
Helemaal oneens	6	6.0	7.9	7.9	--	--
Beetje oneens	3	9.0	3.9	11.8	--	--
Beetje eens	17	17.0	22.4	34.2	--	--
Helemaal eens	50	50.0	65.8	100.0	--	--
Totaal	76	76.0	100.0	--	--	--
Missing	24	24.0	--	--	--	--
Totaal	100	100.0	--	--	2.5	0.9

Tabel 12

Correlatie voldoende tijd om de vraag te beantwoorden met totaalscore

		Voldoende Tijd	Totaalscore
Voldoende Tijd	Pearson Correlatie	1	.226
	Sig. (2-tailed)	--	.073
	N	67	64

Tabel 13

Beschrijvende statistieken slechtziendheid en kleurenblindheid

Score	Slechtziend	N	Gemiddelde	SD
Recollection	Ja	7	7.43	1.272
	Nee	62	7.50	1.457
Totaalscore	Ja	6	22.50	4.231
	Nee	60	22.72	3.508

Score	Kleurenblind	N	Gemiddelde	SD
Recollection	Ja	3	7.67	.577
	Nee	66	7.48	1.460
Totaalscore	Ja	3	21.67	2.517
	Nee	63	22.75	3.592

Tabel 14

Beschrijvende statistieken ervaring spelcomputer

Spel PC	Gem. RC	N RC	SD RC	Gem. Totaal	N Totaal	SD Totaal
Nee, vrijwel nooit	7.43	42	1.625	22.40	40	3.862
Ja, maar niet vaak	7.40	20	1.142	22.89	19	3.862
Ja, ik ben inmiddels een ervaren speler	8.14	7	.690	23.86	7	1.773
Totaal	7.49	69	1.431	22.70	66	3.543

Tabel 15

Correlatie spelcomputerervaring met toetsscores

		Spel PC	Recollection	Totaalscore
Spel PC	Pearson Correlatie	1	.110	.127
	Sig. (2-tailed)	..	.369	.310
	N	81	69	66

Tabel 16

Ik vond het leuk om de toets te maken

Oordeel	N	%	Geldig %	Cumulatief %	Gemiddelde	SD
Helemaal oneens	5	5.0	6.6	6.6	..	..
Beetje oneens	5	5.0	6.6	13.2	..	..
Beetje eens	16	16.0	21.1	34.2	..	..
Helemaal eens	50	50.0	65.8	100.0	..	..
Totaal	76	76.0	100.0	..	..	..
Missing	24	24.0	..	..	..	..
Totaal	100	100.0	..	..	2.5	0.9

Tabel 17

Ik vond de toets leerzaam

Oordeel	N	%	Geldig %	Cumulatief %	Gemiddelde	SD
Helemaal oneens	13	13.0	17.1	17.1	--	--
Beetje oneens	17	13.0	22.4	39.5	--	--
Beetje eens	27	27.0	35.5	75.0	--	--
Helemaal eens	19	19.0	25.0	100.0	--	--
Totaal	76	76.0	100.0	--	--	--
Missing	24	24.0	--	--	--	--
Totaal	100	100.0	--	--	1.7	0.4

Tabel 18

Ik kon me goed concentreren tijdens de toets

Oordeel	N	%	Geldig %	Cumulatief %	Gemiddelde	SD
Helemaal oneens	7	7.0	9.2	9.2	--	--
Beetje oneens	16	16.0	21.1	30.3	--	--
Beetje eens	30	30.0	39.5	69.7	--	--
Helemaal eens	23	23.0	30.3	100.0	--	--
Totaal	76	76.0	100.0	--	--	--
Missing	24	24.0	--	--	--	--
Totaal	100	100.0	--	--	1.9	0.9

Tabel 19

Ik was onzeker om goed te presteren

Oordeel	N	%	Geldig %	Cumulatief %	Gemiddelde	SD
Helemaal oneens	3	3.0	3.9	3.9	--	--
Beetje oneens	17	17.0	22.4	26.3	--	--
Beetje eens	16	16.0	21.1	47.4	--	--
Helemaal eens	40	40.0	52.6	100.0	--	--
Totaal	76	76.0	100.0	--	--	--
Missing	24	24.0	--	--	--	--
Totaal	100	100.0	--	--	2.2	0.9

Tabel 20

Correlatie concentratie met toetsscores

		Concentratie	Recollection	Totaalscore
Concentratie	Pearson Correlatie	1	.281(*)	.227
	Sig. (2-tailed)	--	.021	.071
	N	67	67	64

* Correlatie is significant bij $\alpha = 0.05$ (tweezijdig).

Tabel 21

Correlatie ongevalbetrokkenheid met toetsscores

		Ongevalbetrokkenheid	Recollection	Totaalscore
Ongevalbetrokkenheid	Pearson Correlatie	1	-.016	-.045
	Sig. (2-tailed)	--	.898	.721
	N	67	67	64

Tabel 22

Autotype

Score	Type	N	Gemiddelde	SD	Std.Fout
Recollection	Auto	53	7.47	1.514	.208
	Stationwagen	14	7.57	1.089	.291
Totaalscore	Auto	50	22.88	3.561	.504
	Stationwagen	14	22.36	3.296	.881

E2 Samenhang tussen achtergrondvariabelen en toetsscores

Onderstaande tabellen geven de beschrijvende statistieken en samenhang weer. Gerekend is met de gecorrigeerde toets, waardoor de maximumscore bij het opgaventype 'recollection' 9 is en de maximum totaalscore 30 is.

Tabel 1
Beschrijvende statistische maten geslacht en toetsscores

Score	M/V	N	Gemiddelde	SD	Std.Fout
Recollection	Man	38	7.39	1.603	.260
	Vrouw	31	7.61	1.202	.216
Totaalscore	Man	37	22.32	3.705	.609
	Vrouw	29	22.17	3.328	.619

Tabel 2

Correlatie aantal jaren rijbewijs met toetsscores

		Jaren rijbewijs	Recollection	Totaalscore
Jaren Rijbewijs	Pearson Correlatie	1	.214	.297(*)
	Sig. (2-tailed)	--	.084	.018
	N	66	66	63

* Correlatie is significant bij $\alpha = 0.05$ (tweezijdig).

Tabel 3

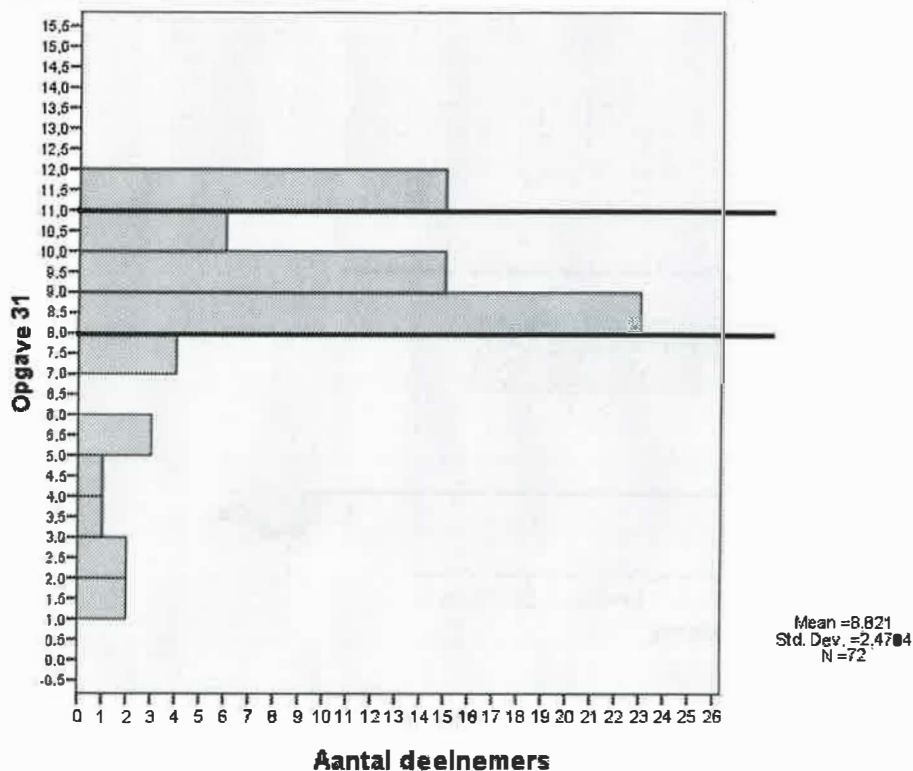
Correlatie kilometrage met toetsscores

		Km/Jaar	Recollection	Totaalscore
Km/Jaar	Pearson Correlatie	1	-.081	.017
	Sig. (2-tailed)	--	.516	.893
	N	67	67	64

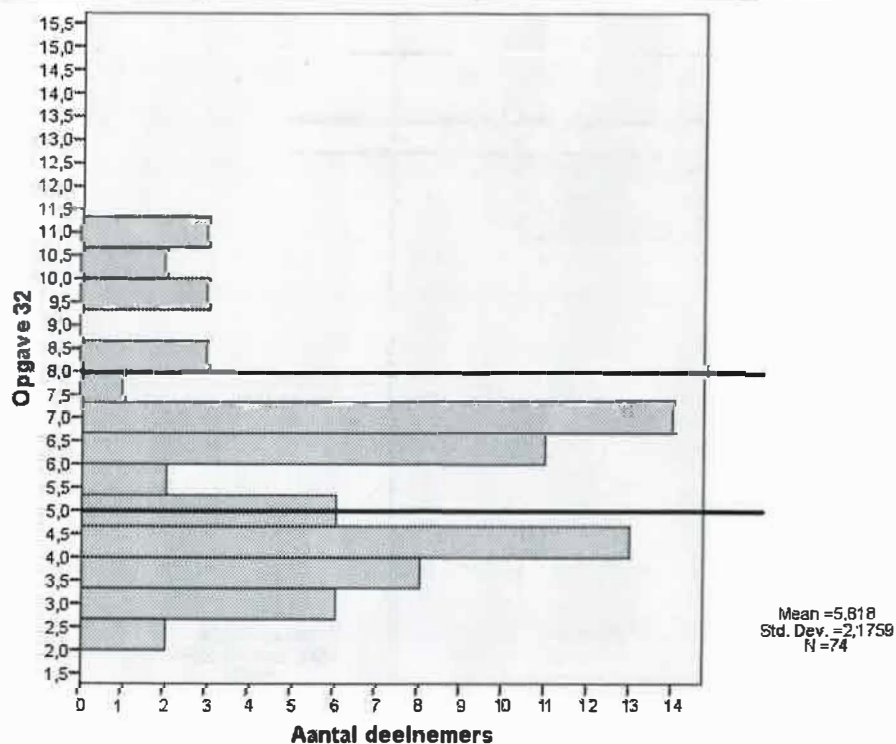
* Correlatie is significant bij $\alpha = 0.05$ (tweezijdig).

Bijlage F Grafieken Timed Response

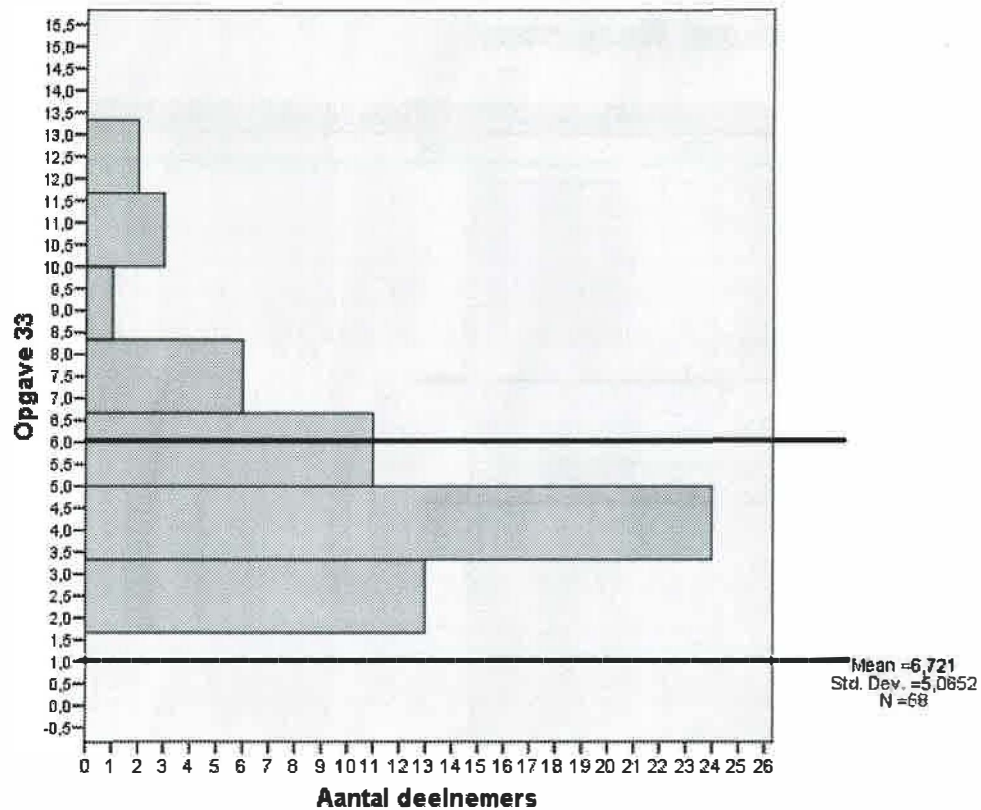
Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
31	.051	13	44	15



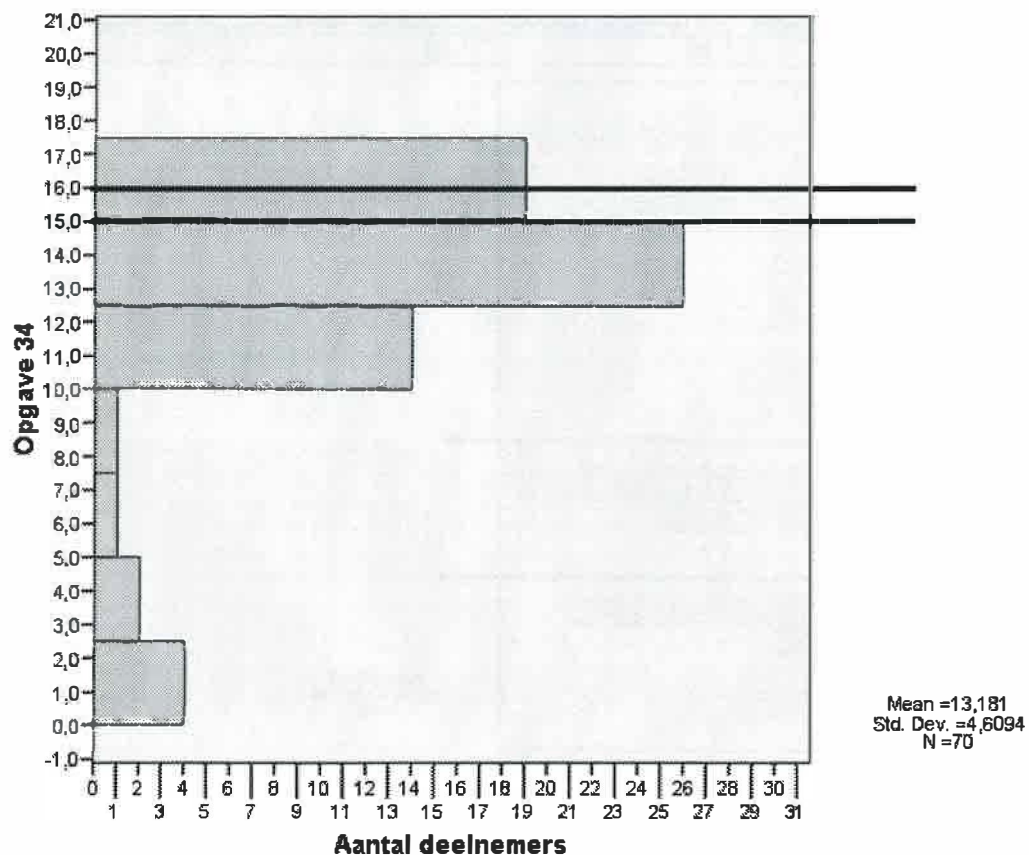
Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
32	-.083	35	28	11



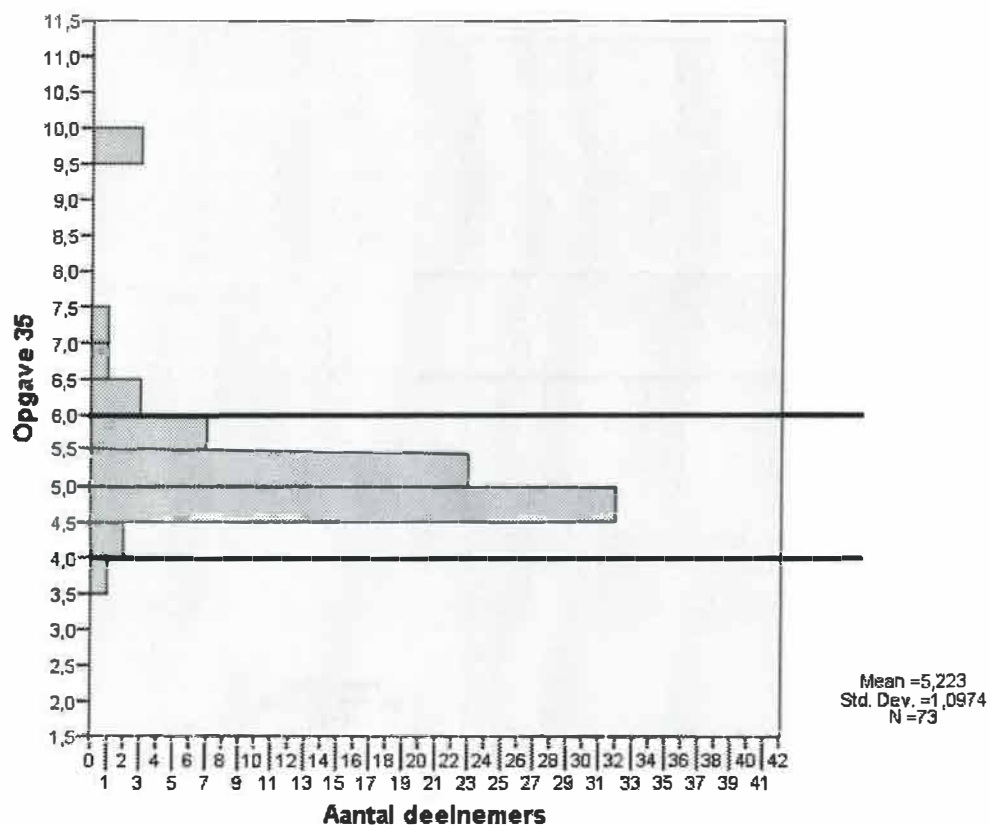
Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
33	.111	0	45	23



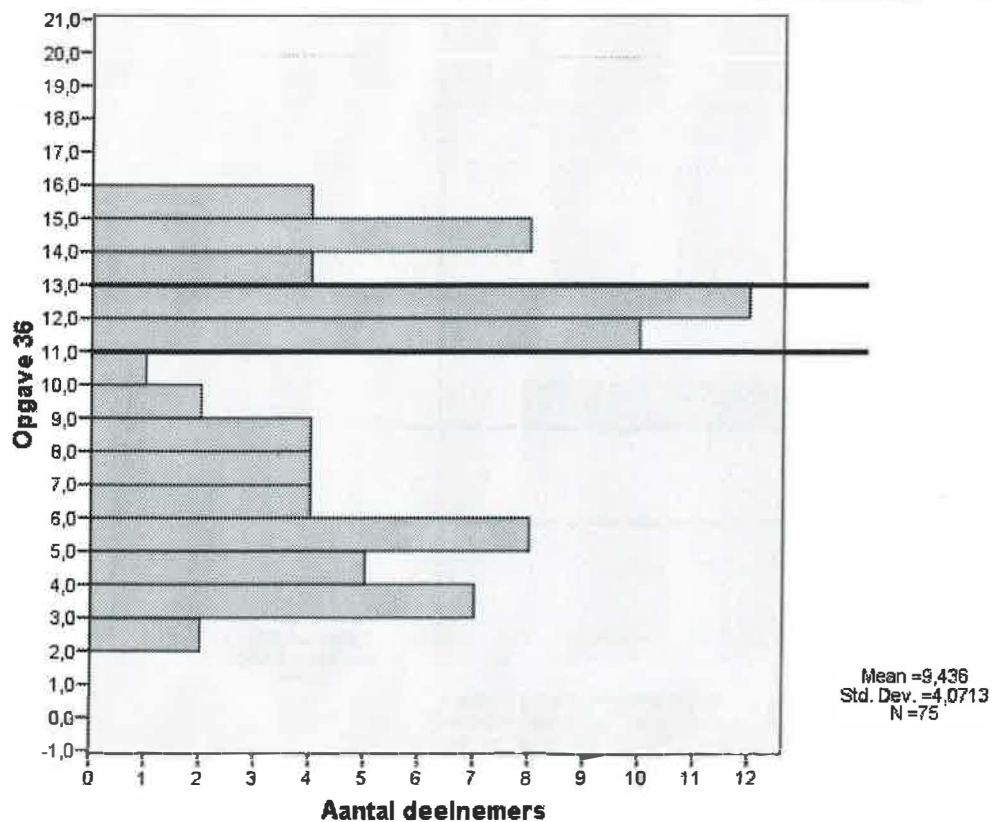
Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
34	-.001	48	9	13



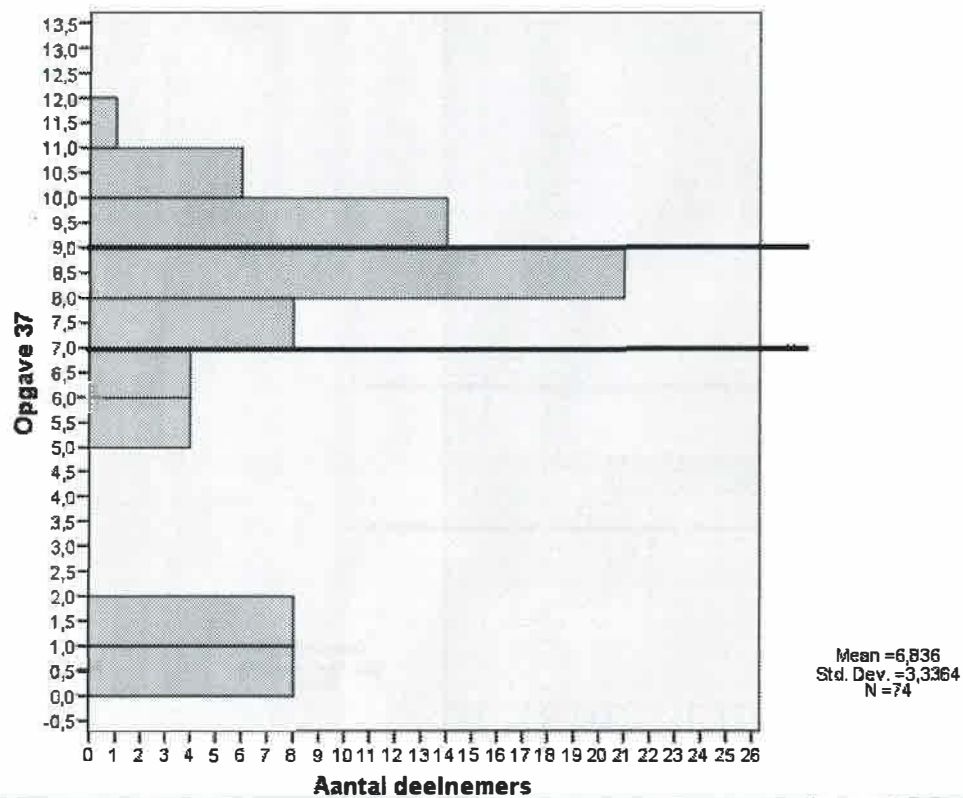
Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
35	.111	1	64	8



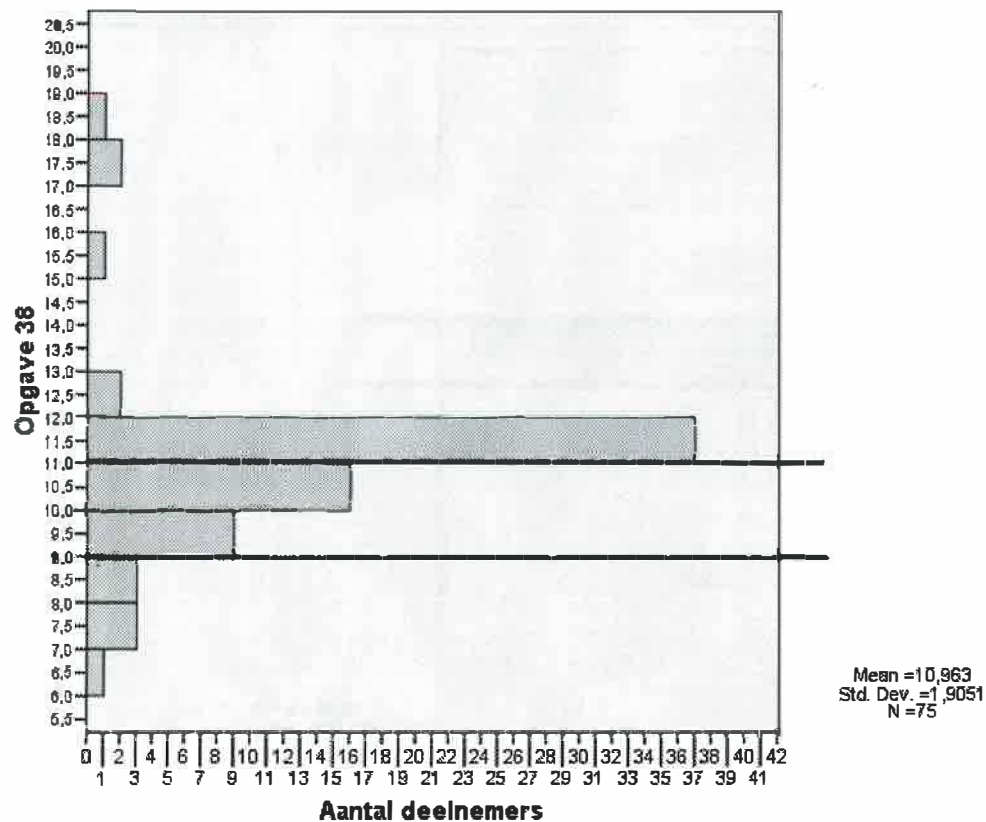
Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
36	.009	37	22	16



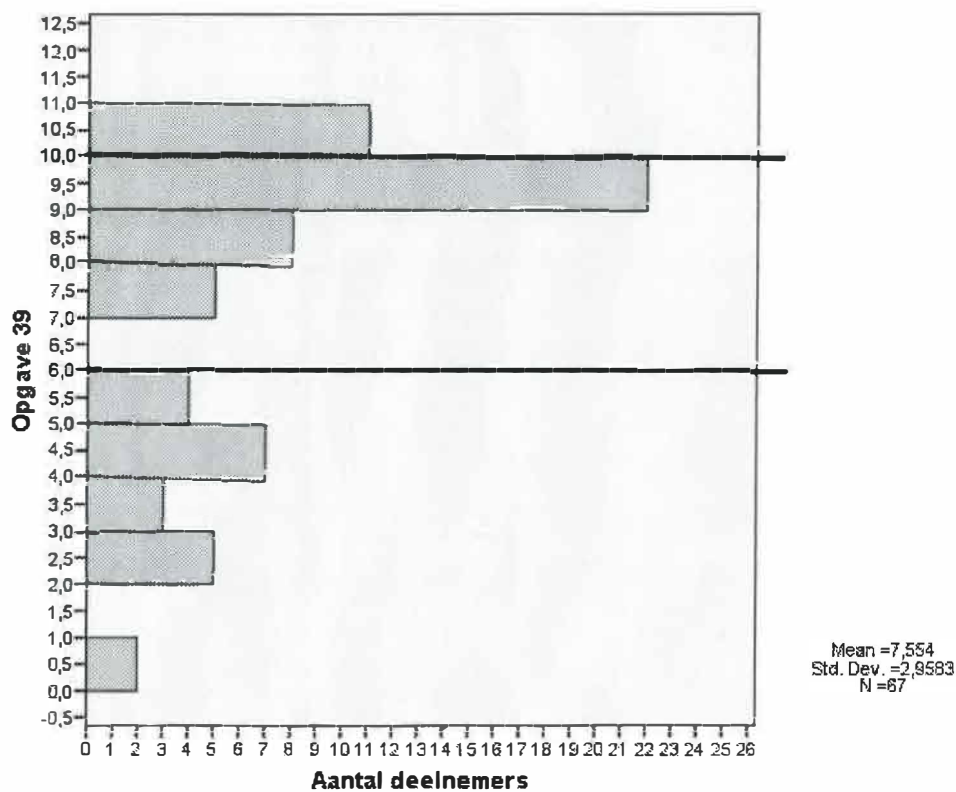
Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
37	.154	24	29	21



Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
38	-.108	7	25	43



Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
39	.057	21	35	11



Opgavenummer	Item Rest Correlatie	N te vroeg geklikt	N op tijd geklikt	N te laat geklikt
40	.183	3	36	35

