

# Samenvatting

## 1 Aanleiding, doel en onderzoeksvragen

Het instrument 'Recidive InschattingsSchalen' (RISc) is het diagnose-instrument van de Nederlandse reclasseringsorganisaties dat in 2002-2003 door Adviesbureau Van Montfoort is ontwikkeld in opdracht van de drie reclasseringsorganisaties — Reclassering Nederland (RN), Stichting Verslavingsreclassering GGZ Nederland (SVG) en Leger des Heils Jeugdzorg en Reclassering (LJ&R) — en het beleidsprogramma Terugdringen Recidive (TR). Voor de ontwikkeling van de RISc heeft de What Works benadering als uitgangspunt gediend (zie bijvoorbeeld McGuire, 1995). Deze benadering gaat ervan uit dat gedragsinterventies die erop gericht zijn om recidive terug te dringen, aangepast moeten zijn op het recidiverisico van een dader en dat zij gericht moeten zijn op het beïnvloeden van factoren die het delictgedrag bevorderen. In overeenstemming met deze uitgangspunten is het doel van de RISc om te komen tot een inschatting van de kans op recidive (gedefinieerd als een nieuwe veroordeling) en van de statische — niet te veranderen factoren zoals leeftijd, geslacht en justitiële voorgeschiedenis — en dynamische — in principe veranderbare en beïnvloedbare factoren — criminogene factoren die aan dit risico ten grondslag liggen. De RISc bestaat uit twaalf schalen die elk één van de criminogene factoren in kaart beogen te brengen: (1) Delictgeschiedenis; (2) Huidig delict en delictpatroon; (3) Huisvesting en wonen; (4) Opleiding, werk en leren; (5) Inkomen en omgaan met geld; (6) Relaties met partner, gezin en familie; (7) Relaties met vrienden en kennissen; (8) Druggebruik; (9) Alcoholgebruik; (10) Emotioneel welzijn; (11) Denkpatronen, gedrag en vaardigheden; en (12) Houding. Samen vormen deze schalen de totaalscore die het risico op recidive aangeeft.

De RISc is gebaseerd op het Britse Offender Assessment System (OASys; Howard, Clark & Garnham, 2003), het instrument dat door de reclassering en het gevangeniswezen in Engeland en Wales wordt gebruikt als een systeem van diagnostiek, indicatiestelling en sanctieplanning. OASys is aangepast aan de Nederlandse situatie en tussen april en september 2003 bij 465 delinquenten afgenomen. Op basis van de op deze manier verzamelde data is de eerste gebruikersversie van het instrument ontwikkeld (Adviesbureau Van Montfoort & Reclassering Nederland, 2004). Tijdens de ontwikkeling van de RISc heeft men zich geconcentreerd op de kwaliteit van de items en is aandacht besteed aan de interne consistentie van de schalen van het instrument. Verder onderzoek naar de betrouwbaarheid en de validiteit was echter, gezien het karakter van het instrument en het doel waarvoor het wordt ingezet, van het grootste belang. De RISc is immers een beoordelingsinstrument op basis waarvan belangrijke beslissingen over individuen worden genomen en het mag daarom voor de resultaten op de RISc niet uitmaken of het instrument door reclasseringswerker A of B wordt afgenomen. Anders geformuleerd moeten beoordelingen generaliseerbaar zijn over beoordelaars. Dit is de vraag naar de interbeoordelaarsbetrouwbaarheid van het instrument. Omdat op basis van de RISc verstrekkende beslissingen kunnen worden genomen over een delinquent, is het van het grootste belang dat het instrument daadwerkelijk meet wat het pretendeert te meten, namelijk recidiverisico en factoren die daaraan ten grondslag liggen. De vraag die hiermee wordt gesteld is de vraag naar de begripsvaliditeit van het in-

strument. Daarnaast is voor elk deugdelijk instrument een heldere, goed te interpreteren schaalstructuur essentieel. In de ontwikkelfase van de RISC is uitgegaan van een aantal a priori opgestelde schalen die tezamen een totaalscore vormen. Deze structuur diende in een grote onderzoeksgroep op basis van statistische analyses te worden getoetst. Dergelijke analyses lenen zich er tegelijkertijd voor om de interne consistentie van de RISC-totaalscore en de RISC-schalen in een grote groep te onderzoeken. Deze vorm van betrouwbaarheid geeft aan in hoeverre per schaal de score replicerbaar is over de items die deel uitmaken van die schaal

Het WODC is in 2005 gestart met een onderzoek naar de interbeoordelaarsbetrouwbaarheid, structuur, interne consistentie en de begripsvaliditeit van de RISC. Het doel van het onderzoek is om inzicht te krijgen in deze psychometrische kenmerken van de RISC, maar voornamelijk om aanbevelingen te doen om het instrument psychometrisch te optimaliseren. De volgende onderzoeksvragen stonden centraal:

- 1 Wat is de interbeoordelaarsbetrouwbaarheid van de RISC?
- 2 Hoe is de begripsvaliditeit van de RISC?  
De vraag naar de begripsvaliditeit van het instrument wordt in dit onderzoek toegespitst op de congruente validiteit ofwel soortgenootvaliditeit: de mate waarin een test correleert met een andere test waarvan wordt aangenomen dat die dezelfde eigenschap meet (Drenth & Sijtsma, 1990).
- 3 Welke aanbevelingen ter verbetering van het instrument zijn op grond van de antwoorden op vraag 1 en 2 te doen?
- 4 Hoe is de schaalstructuur van de RISC en wat is de interne consistentie van de RISC-schalen en de RISC-totaalscore en welke aanbevelingen ter verbetering zijn met betrekking tot de structuur van de RISC te doen?

Op basis van de ervaringen die in het eerste jaar met de RISC zijn opgedaan, is door de drie reclasseringsorganisaties geconstateerd dat bij reclasseringswerkers de indruk bestaat dat de RISC bij een aantal specifieke doelgroepen tot een lagere inschatting van het recidiverisico leidt dan zij in de praktijk aannemelijk achten. Als dit daadwerkelijk het geval zou zijn, betekent dit dat de validiteit van het instrument voor deze doelgroepen onvoldoende zou zijn. Naar aanleiding hiervan is het WODC gevraagd het onderzoek naar de congruente validiteit uit te breiden met de volgende onderzoeksvraag:

- 5 Wat is de relatie tussen de RISC-totaalscore en de score op het recidivevoorspellingsmodel StatRec voor de volgende subgroepen:
  - a Plegers van huiselijk geweld
  - b Zedendelinquenten
  - c Veelplegers
  - d Oplichters
  - e Oudere ouders
  - f Bestuurders onder invloed

Na het bespreken van de voor de beantwoording van de onderzoeksvragen gehanteerde onderzoekopzet en een weergave van de resultaten per deelonderzoek, wordt aan het eind van deze samenvatting een overzicht gegeven van de belangrijkste conclusies en aanbevelingen.

## 2      **Onderzoeksopzet**

Het onderzoek naar de interbeoordelaarsbetrouwbaarheid, schaalstructuur, interne consistentie en de congruente validiteit van de RISC is in drie deelonderzoeken uitgevoerd.

### *De interbeoordelaarsbetrouwbaarheid van de RISC*

Om de interbeoordelaarsbetrouwbaarheid van de RISC te onderzoeken is het instrument in de periode van november 2005 tot half mei 2006 bij een groep van 75 reclasseringscliënten twee maal afgenomen door twee verschillende reclasseringswerkers. Naast een reguliere (eerste) afname is de RISC nogmaals afgenomen door een tweede reclasseringswerker, die dit onafhankelijk van de eerste werker deed. De reclasseringswerkers werkten in vaste koppels, maar hadden onderling geen contact over de cliënt bij wie zij de RISC afnamen. De aan het onderzoek deelnemende reclasseringswerkers, die random waren geselecteerd door de onderzoekers, vroegen de cliënten bij wie zij een RISC afnamen of zij mee wilden werken aan het onderzoek. In totaal werkten negentien koppels van reclasseringswerkers mee aan het onderzoek: elf RN-koppels, vijf SVG-koppels en drie LJ&R-koppels. Wanneer een cliënt aangaf mee te willen werken aan het onderzoek, werd de collega-werker gewaarschuwd waarna die een afspraak maakte met de cliënt voor een tweede afname. Deze tweede afspraak diende ongeveer twee tot drie weken na de eerste afname te worden gepland. Cliënten kregen €25,- voor hun deelname aan het onderzoek.

### *De schaalstructuur en interne consistentie*

Het onderzoek naar de schaalstructuur en de interne consistentie van de RISC is uitgevoerd op basis van een van de Reclassering Nederland (RN) verkregen database waarin alle RISC's zijn opgenomen die tussen november 2004 en mei 2006 door de drie reclasseringsorganisaties zijn gestart. Na de essentiële datacleaning konden de analyses uitgevoerd worden op 11.666 RISC's.

### *De congruente validiteit van de RISC*

Aan het onderzoek naar de congruente validiteit van de RISC zijn twee onderdelen onderscheiden. Het belangrijkste doel van de RISC is om een inschatting van het risico op recidive te bepalen. De vraag of de RISC in voldoende mate correleert met een instrument dat eveneens het recidiverisico beoogt in te schatten, vormt het eerste deel van het onderzoek naar de congruente validiteit. Voor dit deel van het onderzoek is gebruik gemaakt van dezelfde database als bij de analyse van de schaalstructuur en de interne consistentie van de RISC. Voor 9.985 van de 11.666 RISC's uit deze database kon een score worden berekend op de StatRec-schaal. Dit gevalideerde instrument voorspelt op basis van een aantal statische achtergrondkenmerken de recidivekans van daders. Recidive is hierin gedefinieerd als het opnieuw in contact komen met justitie. Door de samenhang tussen beide instrumenten te onderzoeken, kon de congruente validiteit van de RISC-recidivevoorspelling worden vastgesteld. Voor het onderzoek naar de specifieke doelgroepen is volgens dezelfde onderzoeksopzet gewerkt.

Het tweede deel van het onderzoek naar de congruente validiteit van de RISC heeft betrekking op de inhoud van de RISC-schalen. Met behulp van deze schalen worden specifieke leefgebieden in kaart gebracht en de aan- of afwezigheid van criminogene problemen op deze gebieden wordt gebruikt om een indicatie te geven

voor het aanbod van de reclassering, bijvoorbeeld in de vorm van gedragsinterventies. De vraag of deze schalen daadwerkelijk de concepten meten die ze beogen te meten, is in het tweede deel van het validiteitsonderzoek bestudeerd. Het valideren van alle RISC-schalen bleek een te grote belasting te zijn voor de reclasseringsorganisaties en bovendien bestonden er inhoudelijke problemen met betrekking tot het valideren van de meer feitelijke RISC-schalen zoals huisvesting, opleiding / werk en financiën. Het lag namelijk het meest voor de hand om dergelijke schalen te valideren met behulp van dossierinformatie, maar omdat de reclasseringswerker bij het afnemen en invullen van de RISC gebruik maakt van dezelfde dossierinformatie, was dat geen mogelijkheid. Uiteindelijk is in overleg met de drie reclasseringsorganisaties besloten de drie minst feitelijke en meest subjectieve schalen uit het instrument te valideren. Het gaat daarbij om de schalen 10 *Emotioneel welzijn*, 11 *Denkpatronen, gedrag en vaardigheden* en 12 *Houding*. Dit zijn bovendien schalen die een belangrijke rol spelen in de beoordeling van de vraag of een cliënt een gedragsinterventie zou moeten volgen. Om de congruente validiteit van deze schalen te bestuderen, is op zoek gegaan naar betrouwbare en gevalideerde instrumenten die (delen van) dezelfde concepten meten. Drie vragenlijsten — de Nederlandse Persoonlijkheidsvragenlijst (NPV), Utrechtse Copinglijst (UCL) en de Buss-Durkee Hostility Inventory - Dutch (BDHI-D) — werden voor het onderzoek geselecteerd. Reclasseringswerkers vroegen hun cliënten na afloop van de afname van de RISC om medewerking aan het onderzoek naar de congruente validiteit. Iedere cliënt vulde slechts één van de drie geselecteerde vragenlijsten in. Tijdens het invullen van de vragenlijst was de reclasseringswerker aanwezig voor eventuele uitleg. Als de cliënt klaar was met invullen, controleerde de werker of alle vragen waren ingevuld en of geen bladzijden waren overgeslagen. Voor hun medewerking aan het onderzoek ontvingen cliënten €10,-. De vragenlijsten zijn tussen april en half september 2006 door 204 reclasseringscliënten ingevuld. In een aantal gevallen was de RISC van de cliënt onvindbaar, onbruikbaar of op de sluitingsdatum van de dataverzamelingsperiode nog niet afgesloten. In de totale validiteitsonderzoeksgroep resteren daarom 185 personen.

### **3 Resultaten van het onderzoek naar de interbeoordelaarsbetrouwbaarheid van de RISC**

Negentien koppels van reclasseringswerkers hebben de RISC onafhankelijk van elkaar twee maal afgenomen bij 75 reclasseringscliënten. De eerste RISC-afname werd gevolgd door de tweede RISC-afname met een tussenpoos van gemiddeld 29 dagen. Nadat is vastgesteld dat de onderzoeksgroep met betrekking tot hun achtergrondkenmerken en hun scores op de RISC in grote mate overeenkomt met de reguliere reclasseringspopulatie is de kwaliteit van de data in termen van missende waarden onderzocht. Deze bleek goed te zijn. De interbeoordelaarsbetrouwbaarheid van de RISC is zowel op itemniveau als op schaalniveau onderzocht. Ook de interbeoordelaarsbetrouwbaarheid van de RISC-totaalscore is bestudeerd. De gebruikte analysetechnieken om de interbeoordelaarsbetrouwbaarheid van de RISC-items en —schalen vast te stellen, zijn afgestemd op het meetniveau van items en schalen. Om de mate van overeenstemming tussen de reclasseringswerkers op *nominale* items te onderzoeken, is coëfficiënt  $\kappa$  berekend (Cohen, 1960). Voor *ordina-*le items, de schaalcores en de totaalscore is de volgende strategie gebruikt (cf. Born, 1995, p. 130-132).

- 1 Het berekenen van de proportie overeenstemming. Deze maat is de meest gebruikte en eenvoudigste maat, maar hiermee kon niet worden volstaan. Het ontbreekt namelijk zowel aan een controle voor kansovereenstemming als aan een formele toetsing van de mate van overeenstemming.
- 2 Het berekenen van de  $\chi^2$ -maat van Lawlis en Lu (1972). Deze maat geeft aan of de overeenstemming significant groter is dan op basis van toeval te verwachten is (kansovereenstemming).
- 3 T-maat van Tinsley en Weiss (1975) berekenen. Deze maat is een afgeleide van de  $\chi^2$ -maat van Lawlis en Lu en geeft een indicatie van de grootte van de overeenstemming (0=de overeenstemming is niet groter dan op basis van toeval is te verwachten, 1=perfecte overeenstemming).

In alle gevallen is op basis van de resultaten een uitspraak gedaan over de mate van overeenstemming tussen de reclasseringswerkers. Met betrekking tot de items van de RISc is de mate van overeenstemming tussen reclasseringswerkers over het algemeen redelijk tot sterk en ook op schaalniveau is de overeenstemming tussen de beoordelaars in alle gevallen redelijk tot sterk. Met betrekking tot de totaalscore van de RISc komen reclasseringswerkers in sterke mate tot hetzelfde oordeel. De interbeoordelaarsbetrouwbaarheid van de RISc is op basis van deze resultaten goed te noemen. Wel valt op dat de interbeoordelaarsbetrouwbaarheid van items afneemt naarmate minder naar feitelijkheden wordt gevraagd en dat een groot aantal items van de schalen 11 en 12 slechts een matige overeenstemming laat zien.

#### **4 Resultaten van het onderzoek naar de schaalstructuur en de interne consistentie van de RISc**

Om de schaalstructuur van het instrument te onderzoeken, is steeds voor elke schaal en voor de totaalscore een principale componentenanalyse (pca) uitgevoerd op de bijbehorende items respectievelijk schalen. Voor elke schaal en voor de totaalscore is nagegaan in hoeverre de pca één factor opleverde en zijn de factorloadingen bestudeerd. Op deze manier werd nagegaan of de items respectievelijk schalen eenzelfde achterliggend begrip in kaart brengen. Daarnaast is voor elke schaal en voor de totaalscore een betrouwbaarheidsanalyse uitgevoerd. De interne consistentie van elke schaal en van de totaalscore is berekend met behulp van Cronbach's coëfficiënt alfa.

De RISc-schalen vormen over het algemeen goede schalen: de meeste items van elke schaal meten ieder een deel van hetzelfde achterliggende concept dat de schaal in kaart beoogt te brengen en de scores zijn in de meeste schalen voldoende repliceerbaar over de items. Dat wil zeggen dat met betrekking tot de meeste schalen sprake is van een betrouwbare meting. Dit blijkt uit de interne consistentie van de schalen, die voor de meeste schalen voldoende tot goed is. De uitzonderingen worden gevormd door de schalen 6 *Relaties met partner, gezin en familie* en 9 *Alcoholgebruik*. Deze schalen bestaan uit items die in onvoldoende mate eenzelfde achterliggend begrip in kaart brengen en die onvoldoende betrouwbaar zijn in hun meting (onvoldoende interne consistentie).

De analyse van de RISc-totaalscore toont aan dat ook de RISc-schalen samen een goede schaal vormen: ze laden over het algemeen goed op de geëxtraheerde factor en meten dus elk een deel van hetzelfde achterliggende begrip. De enige schaal die

weinig bijdraagt aan de totaalscore is schaal 9 *Alcoholgebruik*. Aan de interne consistentie van de totaalscore worden strengere eisen gesteld dan aan de interne consistentie van de afzonderlijke RISC-schalen omdat op basis van deze score belangrijkere beslissingen over individuele daders worden genomen dan op basis van de afzonderlijke schaalesscores. De betrouwbaarheidsanalyse van de totaalscore toont aan dat aan deze strengere eisen wordt voldaan.

## 5 Resultaten van het onderzoek naar de congruente validiteit van de RISC

Ten behoeve van het onderzoek naar de congruente validiteit van de RISC in termen van recidivevoorspelling is de samenhang tussen RISC en de StatRec-schaal bestudeerd. Dit is op verschillende manieren gedaan. Als eerste stap is de onderzoeksgroep op basis van hun RISC-totaalscore in vier risicogroepen verdeeld en is nagegaan in hoeverre de StatRec-scores van de vier groepen significant van elkaar verschillen. Vervolgens zijn correlaties berekend tussen RISC en StatRec en tot slot zijn regressieanalyses uitgevoerd.

De resultaten van het onderzoek naar de congruente validiteit van de RISC in termen van de voorspelling van de recidivekans zijn gunstig. Er werd, conform de verwachting, een sterk verband gevonden tussen de RISC-totaalscore, die een inschatting beoogt te geven van de recidivekans, en de gevalideerde voorspelling van de recidive zoals deze met StatRec wordt berekend. Op basis van het uitgangspunt van de RISC — elke schaal houdt verband met (het risico op) delictgedrag — werd verwacht dat alle correlaties tussen de afzonderlijke RISC-schalen en de StatRec-recidivekans middelsterk positief zouden zijn. In de meeste gevallen werd deze verwachting in de totale onderzoeksgroep ondersteund. De uitzonderingen worden gevormd door de schalen 6 *Relaties met partner, gezin en familie*, 9 *Alcoholgebruik* en 10 *Emotioneel welzijn*. De samenhang tussen deze schalen en de StatRec-voorspelling van recidive is zwak. De totale onderzoeksgroep is op basis van haar achtergrondkenmerken — geslacht, leeftijd, herkomst en uitgangsdelict — verdeeld in een aantal subgroepen waarvoor de samenhang tussen RISC en StatRec eveneens is bestudeerd. Het beeld dat uit deze analyses naar voren komt, is vergelijkbaar met dat uit de analyse van de totale onderzoeksgroep.

Regressieanalyses zijn uitgevoerd om na te gaan in hoeverre de score op de StatRec te verklaren is op basis van de RISC-schalen. De resultaten laten zien dat de RISC-schalen samen 48% van de variantie van de StatRec-recidivekans verklaren. Dit is een zeer aanzienlijk deel en deze bevinding ondersteunt dan ook de congruente validiteit van de RISC. Ook in vrijwel alle subgroepen wordt een aanzienlijk deel van de StatRec-recidivekans verklaard door de RISC-schalen. De groep vrouwen vormt hierop een uitzondering. De mate waarin de StatRec-score van deze groep wordt verklaard door hun scores op de RISC is weliswaar groot, maar aanzienlijk minder groot dan bij mannen het geval is. Hoewel het, gezien de onderlinge samenhang tussen de afzonderlijke RISC-schalen, niet reëel was om te verwachten dat van elke afzonderlijke schaal de unieke bijdrage aan de verklaring van de StatRec-recidivekans kon worden bepaald, zijn enkele opmerkelijke resultaten gevonden. De schalen 4 *Opleiding, werk en leren* en 7 *Relaties met vrienden en kennissen* leveren, naast schaal 1&2 *Delictgegevens*, in de totale onderzoeksgroep een noemenswaardige, unieke bijdrage aan de verklaring van de StatRec-score. Dit betekent dat deze leefgebieden, ondanks de samenhang die ze vertonen met de andere leefgebieden, een eigen invloed hebben op de verklaring van de StatRec-score. De

invloed van opleiding en relaties met vrienden en kennissen op de verklaring van de op basis van StatRec berekende recidivekans is ook in veel van de onderzochte subgroepen gevonden. Ook schaal 8 *Druggebruik* blijkt in een aantal subgroepen een eigen bijdrage te leveren aan de verklaring van de StatRec-recidivekans. Daarnaast bestaan er verschillen tussen de bestudeerde subgroepen in de aard van de dynamische factoren die, ondanks de onderlinge samenhang tussen de RISC-schalen, een unieke bijdrage leveren aan de verklaring van de StatRec-voorspelling van de recidivekans. Ook is er onderscheid te maken tussen de diverse groepen in de sterkte van de effecten van RISC-schalen op de verklaring van de StatRec-score.

Ook in het kader van het onderzoek naar de congruente validiteit van de RISC is de samenhang bestudeerd tussen scores op de schalen 10 *Emotioneel welzijn*, 11 *Denkpatronen, gedrag en vaardigheden* en 12 *Houding* van de RISC en de scores op de Nederlandse Persoonlijkheidsvragenlijst (NPV), Utrechtse Copinglijst (UCL) en de Buss-Durkee Hostility Inventory-Dutch (BDHI-D). Voorafgaand aan de analyses is een aantal verwachtingen opgesteld met betrekking tot de samenhang tussen de RISC-schalen en bepaalde, geselecteerde schalen van de NPV, UCL en BDHI-D. De uitkomsten van het onderzoek naar de congruente validiteit van de schalen 10 *Emotioneel welzijn*, 11 *Denkpatronen, gedrag en vaardigheden* en 12 *Houding* laten een gunstig beeld zien van de congruente validiteit van de twee eerstgenoemde schalen. De verwachte, middelsterke samenhang wordt gevonden tussen Schaal 10 *Emotioneel welzijn* en alle schalen die een soortgelijk begrip meten. Ditzelfde geldt ook voor schaal 11 *Denkpatronen, gedrag en vaardigheden*. Dit vormt een onderbouwing voor de aanname dat deze schalen de begrippen in kaart brengen die ze beogen te meten. Schaal 12 *Houding* vertoont daarentegen met twee van de drie onderzochte schalen niet de verwachte, middelsterke samenhang. Ondanks de kanttekeningen die bij dit deelonderzoek zijn te plaatsten en ondanks het feit dat tussen schaal 12 en de derde van de onderzochte schalen wel de verwachte, middelsterke samenhang werd gevonden, kunnen de resultaten van dit onderzoek twijfels opwekken over de vraag of de RISC-schaal 12 er voldoende in slaagt de houding van de reclasseringscliënt ten opzichte van andere mensen, de maatschappij, het delict en criminaliteit in het algemeen in kaart te brengen.

## **6 Resultaten van het onderzoek naar de samenhang tussen RISC en StatRec in een aantal specifieke doelgroepen**

Omdat onder reclasseringswerkers de indruk bestond dat de RISC in het geval van een aantal specifieke doelgroepen een te laag recidiverisico inschat, is de samenhang tussen RISC en StatRec in deze groepen afzonderlijk bestudeerd. Zowel RISC als StatRec beoogt de kans in te schatten dat een dader opnieuw met justitie in aanraking komt. De verwachting is dat de RISC-totaalscore en de StatRec-voorspelling van de recidivekans ook in de bestudeerde specifieke doelgroepen een sterke samenhang vertonen. Dit blijkt voor alle groepen, behalve voor de groep veelplegers (daders met elf of meer eerdere strafzaken) te worden bevestigd. In deze groep is de correlatie tussen RISC en StatRec middelsterk. De afzonderlijke RISC-schalen die dynamische criminogene factoren in kaart beogen te brengen, worden verwacht een middelsterke samenhang met de StatRec-score te vertonen. Deze verwachting wordt over het algemeen bevestigd. Net als in de totale onderzoeksgroep vormen de schalen 6 *Relaties met partner, gezin en familie*, 9 *Alcoholge-*

*bruik* en 10 *Emotioneel welzijn* de belangrijkste uitzonderingen. Deze schalen vertonen in alle subgroepen zwakke correlaties met de door StatRec voorspelde recidivekans.

Regressieanalyses wijzen uit dat de StatRec-score in de groep bestuurders onder invloed en de groep oplichters het best door de RISC wordt verklaard. Ook in de groep plegers van huiselijk geweld en de groep zedendelinquenten verklaart de RISC een aanzienlijk deel van de variantie van de StatRec-voorspelling van het recidiverisico. De RISC-schalen verklaren de minste variantie van de StatRec-score bij de groep veelplegers. Hoewel dit overeenkomt met een sterke correlatie, slaagt RISC er beduidend minder goed dan in de andere groepen in om de StatRec-voorspelling van de recidivekans te verklaren.

Van de dynamische RISC-schalen levert, ondanks de samenhang met alle andere RISC-schalen, in de meeste groepen de schaal 4 *Opleiding, werk en leren* een noemenswaardig unieke bijdrage aan de verklaring van de StatRec-score. De schalen 5 *Inkomen en omgaan met geld* en 9 *Alcoholgebruik* leveren elk in drie van de zes bestudeerde groepen een noemenswaardige, unieke bijdrage aan de verklaring van de StatRec-voorspelling van de recidivekans. Tussen de specifieke subgroepen worden ook enkele verschillen gevonden in de aard van de dynamische factoren die een unieke bijdrage leveren aan de verklaring van de StatRec-voorspelling van de recidivekans.

Over het algemeen leveren de analyses geen aanwijzingen op dat de RISC voor de onderzochte doelgroepen een te laag recidiverisico inschatten. Dit geldt echter niet voor de groep veelplegers. Afgaande op het op basis van StatRec ingeschatte recidiverisico van deze groep lijkt het erop dat reclasseringswerkers gelijk hebben wanneer zij stellen dat RISC bij sommige veelplegers tot een te lage inschatting van de recidivekans komt. Veelplegers die volgens RISC een laag of een middelhoog recidiverisico hebben, hebben namelijk op basis van StatRec een recidivekans van 66% respectievelijk 76%.

## **7 Conclusies en aanbevelingen**

De resultaten van de drie uitgevoerde studies naar de interbeoordelaarsbetrouwbaarheid, de structuur en de congruente validiteit van de RISC in ogenschouw nemend, luidt de algemene conclusie dat de RISC in termen van de onderzochte kenmerken over gunstige psychometrische kwaliteiten beschikt. Om het instrument verder te verbeteren zijn op verschillende punten aanpassingen mogelijk en vervolgonderzoek, met name naar de predictieve validiteit van het instrument, is nodig. Voor het vergroten van de interbeoordelaarsbetrouwbaarheid van de RISC worden de volgende aanbevelingen gedaan:

- Formuleer de invulinstructies van de RISC nog eenduidiger en voorkom op die manier mogelijk bestaande ambivalentie en verschillen in interpretatie door reclasseringswerkers.
- Overwogen dient te worden de antwoordschalen aan te passen aan de inhoud van het item. De antwoordschaal van moeilijk in te schatten items dient ruimte te bieden voor nuance en antwoordschalen met slechts twee antwoordcategorieën dienen in dat soort gevallen te worden vermeden.
- Geadviseerd wordt om nog meer aandacht te besteden aan opleiding en deskundigheidsbevordering voor reclasseringswerkers.



- De matige interbeoordelaarsbetrouwbaarheid van de items van schalen 11 en 12 dienen aangepakt te worden. Op de eerste plaats door middel van deskundigheidsbevordering, maar daarnaast kan worden overwogen de scoringsinstructies te onderwerpen aan een onderzoek door een ervaren gedragsdeskundige.

De kwaliteit van de structuur van de RISC kan op een aantal manieren worden verbeterd. Door de items 2.10 *Verantwoordelijkheid nemen voor het delictgedrag*, 6.4 *Familie- of gezinslid heeft een justitieel dossier*, 8.5 *Motivatie om druggebruik aan te pakken* en 9.5 *Motivatie om alcoholgebruik aan te pakken* niet meer te betrekken in de berekening van de score op de schalen waar ze bij horen, kan de kwaliteit van deze schalen worden verbeterd. Hiermee wordt niet aangeraden deze items uit de RISC te verwijderen, maar te zoeken naar andere mogelijkheden om de informatie die met deze items wordt verzameld, nuttiger in te zetten. Aangeraden wordt om item 2.10 bij schaal 12 *Houding* in te delen. De items 8.5 en 9.5 zouden behouden kunnen worden als aparte indicatoren voor motivatie en zouden opgenomen kunnen worden in een aparte 'motivatieschaal'. Onderzoek naar de samenhang tussen item 6.4 en daadwerkelijke recidive zal uitsluitel moeten geven over de feitelijke waarde van dit item. Met betrekking tot de totaalscore worden geen aanbevelingen gedaan. Hoewel verbeteringen mogelijk zijn, wordt aangeraden eerst op schaalniveau aanpassingen te doen en vervolgens na te gaan welk effect dit op de psychometrische kenmerken van de totaalscore heeft.

De congruente validiteit van de RISC in termen van de voorspelling van recidive is goed. De RISC-totaalscore, die nieuwe contacten met justitie beoogt te voorspellen, hangt zowel in de totale onderzoeksgroep als in vrijwel alle bestudeerde subgroepen in sterke mate samen met de StatRec-voorspelling van de recidivekans. Uit de resultaten van de regressieanalyses blijkt dat de score op de StatRec in de totale onderzoeksgroep en in vrijwel alle onderzochte subgroepen goed te verklaren is op basis van de scores op de RISC-schalen tezamen. Er zijn voorts op basis van de resultaten in dit onderzoek, met uitzondering van de groep veelplegers, geen aanwijzingen dat de RISC in bepaalde groepen tot een te lage inschatting komt van de recidivekans. Deze resultaten vormen een overtuigende onderbouwing voor de congruente validiteit van de totaalscore van de RISC. Hier moet echter wel bij worden aangetekend dat de RISC is gevalideerd met behulp van een *voorspelling van recidive*. De samenhang tussen RISC en StatRec, bestudeerd met behulp van correlaties en regressieanalyses, geeft dan ook geen uitsluitel over de daadwerkelijke voorspellingskracht van de RISC en onderzoek naar de mate waarin de RISC samenhangt met daadwerkelijke recidive is daarom nodig. Speciale aandacht zal hierbij uit moeten gaan naar de voorspelling van recidive bij vrouwen en veelplegers. De congruente validiteit van de RISC is bij vrouwen niet in het geding, maar de RISC lijkt er toch wat minder goed in te slagen de recidivekans te verklaren. Voor veelplegers zijn er aanwijzingen dat de RISC ten onrechte een te laag recidive-risico inschat.

De vraag of de schalen 10, 11 en 12 meten wat ze beogen te meten, kan in het geval van de schalen 10 *Emotioneel welzijn* en 11 *Denkpatronen, gedrag en vaardigheden* vrij gunstig worden beantwoord. Vervolgonderzoek naar de congruente validiteit van schaal 12 *Houding* is echter aan te bevelen omdat de resultaten uit het onderhavige onderzoek niet eenduidig waren. Daarnaast zal de congruente validiteit

van de overige schalen van de RISC in een of meerdere vervolgstudies moeten worden onderzocht.

Wanneer de resultaten van de drie deelonderzoeken tezamen worden overzien, kan worden geconcludeerd dat er twee RISC-schalen zijn waarbij vraagtekens kunnen worden geplaatst met betrekking tot hun waarde voor het instrument: schaal 6 *Relaties met partner, gezin en familie* en schaal 10 *Emotioneel welzijn*. Van schaal 6 kon zowel de betrouwbaarheid (interne consistentie) als de congruente validiteit in dit onderzoek niet worden ondersteund. De congruente validiteit van schaal 10 in termen van *inhoud* (emotioneel welzijn) wordt weliswaar ondersteund, maar de schaal hangt zowel in de totale onderzoeksgroep als in alle onderzochte subgroepen slechts zwak samen met de voorspelling van recidive. De congruente validiteit van deze schaal in termen van de samenhang met de recidivekans kan in dit onderzoek dan ook niet worden onderbouwd.