

Risk Assessment With the START: AV in Dutch Secure Youth Care

Citation for published version (APA):

De Beuf, T. (2022). *Risk Assessment With the START: AV in Dutch Secure Youth Care: From Implementation to Field Evaluation*. [Doctoral Thesis, Section Forensic Psychology]. Maastricht University. <https://doi.org/10.26481/dis.20220404td>

Document status and date:

Published: 04/04/2022

DOI:

[10.26481/dis.20220404td](https://doi.org/10.26481/dis.20220404td)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Unspecified

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

SAMENVATTING

Samenvatting [Dutch Summary]

In 2014 besloot de OG Heldringstichting, een Nederlandse Jeugdzorg^{plus}-instelling die vanaf 2021 bekend staat onder de naam Pactum, de Short-Term Assessment of Risk and Treatability: Adolescenten Versie (START:AV; Viljoen et al., 2014/2016) in te voeren. Het doel was om de risico's van de jeugdigen binnen hun zorg op een gestructureerde manier te schatten. Dit project werd uitgebreid met een promotieonderzoek en in deze context kwam het huidige proefschrift tot stand. Het onderzoek betreft de evaluatie van de implementatie en van de START:AV zelf in deze Jeugdzorg^{plus}-instelling en dient twee doelen. Ten eerste, het in kaart brengen van de uitdagingen en mogelijkheden van het START:AV-implementatieproces, alsook het evalueren van de effectiviteit van de implementatieprocedure aan de hand van verschillende implementatie-uitkomsten. Ten tweede, het onderzoeken van de betrouwbaarheid en de voorspellende waarde van de START:AV in deze specifieke setting, met behandelcoördinatoren als uitvoerders van de risicotaxatie.

Hoofdstukken in dit Proefschrift

Hoofdstuk 1 beschrijft de Nederlandse specialistische Jeugdzorg^{plus} behandelsetting en de kenmerken van de jongeren die (gedwongen) opgenomen worden in deze gesloten jeugdzorg. Kenmerkend voor de doelgroep is de veelheid en variëteit aan negatieve gebeurtenissen die ze hebben meegemaakt in de kindertijd en/of adolescentie (bijv. kindermishandeling, automutilatie, delinquent gedrag). Daarnaast hebben ze vele risicofactoren die hen kwetsbaar maken voor een herhaling van dergelijke negatieve gebeurtenissen in de toekomst. Daarom is het uitvoeren van een gestructureerde risicotaxatie met behulp van wetenschappelijk onderbouwde instrumenten relevant voor deze doelgroep. In dit hoofdstuk wordt de START:AV geïntroduceerd als risicotaxatie-instrument passend bij deze doelgroep. Het instrument wordt beschreven en wetenschappelijke studies naar de betrouwbaarheid en de voorspellende waarde van het instrument worden gepresenteerd. Het hoofdstuk eindigt met een beschrijving van het promotie-onderzoek alsook de setting waarbinnen het plaatsvond.

In **Hoofdstuk 2** onderzochten we welke uitdagingen en mogelijkheden medewerkers ondervonden tijdens de implementatie van de START:AV. Via focusgroepen zijn de ervaringen van zorgmedewerkers van verschillende professionele disciplines binnen de organisatie (bijv. pedagogisch medewerkers, behandelcoördinatoren) verzameld en geclassificeerd volgens de domeinen van het *consolidated framework of implementation research* (CFIR). Dit kader is ontwikkeld binnen de implementatiewetenschap en maakt inzichtelijk welke belemmeringen en mogelijkheden (kunnen) optreden tijdens een implementatieproces. In de focusgroepen en interviews werden in totaal 57 factoren genoemd die volgens de medewerkers de START:AV-implementatie beïnvloedden, zowel positief als negatief. Met name implementatiebelemmeringen werden benoemd, waaronder eigenschappen van: (1) het instrument zelf (bijv. overlap tussen items), (2) het personeel (bijv. weerstand tegen de

nieuwe procedure), en (3) het implementatieproces (bijv. een te lange voorbereidingsperiode). Daarnaast werden ook verschillende eigenschappen van de organisatie als belemmerend ervaren, zoals de hiërarchie binnen de organisatie, de organisatiecultuur, communicatielijnen, en de houding van de organisatie ten opzichte van verandering. Medewerkers identificeerden ook een aantal factoren die een positieve invloed hadden op de implementatie, zoals de voordelen van de START:AV ten opzichte van de vroegere werkwijze en de positieve houding van medewerkers ten opzichte van het instrument.

In **Hoofdstuk 3** evalueerden we het succes van de implementatie aan de hand van vijf uitkomstmaten uit het *implementation outcomes framework* van Enola Proctor en collega's (2011). Op verschillende momenten tijdens het implementatieproces evalueerden we de implementatie-uitkomsten "aanvaardbaarheid" (*acceptability*), "opname" (*adoption*), "bruikbaarheid" (*applicability*), "uitvoerbaarheid" (*feasibility*), en "penetratie" (*penetration*) met behulp van een vragenlijst die werd voorgelegd aan de behandelcoördinatoren (d.w.z., de hoofdgebruikers van de START:AV in deze instelling). Uit de vragenlijst bleek dat de kenmerkende onderdelen van het instrument, zoals het beoordelen van het risico voor meerdere schadelijke uitkomsten en het scoren van zowel krachten als kwetsbaarheden, door de meeste behandelcoördinatoren als zinvol beschouwd werden ("bruikbaarheid"). Toch bleek de tevredenheid over de START:AV te verminderen tijdens het implementatieproces ("aanvaardbaarheid"). Vermoedelijk kwam dit doordat behandelcoördinatoren vanaf de invoering van de START:AV een toegenomen werkdruk ervoeren ("uitvoerbaarheid"). Ondanks deze ontevredenheid was het invulpercentage voldoende: de meerderheid van de START:AVs die ingevuld moesten worden, werd daadwerkelijk ingevuld ("opname"). Er waren wel onderlinge verschillen, met een relatief laag invulpercentage bij sommige behandelcoördinatoren. Middels deze vijf implementatie-uitkomsten kregen we inzicht in verschillende facetten van de implementatie. Hierdoor konden verbeterpunten geïdentificeerd en aangepakt worden om de implementatie te verbeteren.

+

In **Hoofdstuk 4** werd een zesde implementatie-uitkomst geëvalueerd, namelijk "adherence", een term die zich moeilijk naar het Nederlands laat vertalen (bijv. "methodetrouw"). Adherence reflecteert de mate waarin START:AV-formulieren ingevuld worden volgens de instructies in de handleiding. Om deze uitkomst specifiek voor de START:AV te meten, hebben we een beoordelingsschaal ontwikkeld: de *START:AV adherence rating scale* (STARS; De Beuf et al., 2018). Met deze beoordelingsschaal toetsten we 306 START:AV-formulieren op het correct en compleet invullen van 11 centrale onderdelen (bijv. Zijn alle items ingevuld?). Uit de resultaten bleek dat de mate van "adherence" verschilde per START:AV-onderdeel. Zo werden de items en de schadelijke uitkomsten over het algemeen compleet ingevuld, maar START:AV-onderdelen waarbij een onderbouwing gevraagd werd (zoals de onderbouwing bij matig en hoog risico-

inschattingen), hadden meer ontbrekende informatie. Adherence aan de instructies kon wel verbeterd worden: na een opfrisworkshop over het correct invullen van de START:AV werden de risicotaxatie-formulieren vollediger ingevuld.

In **Hoofdstuk 5** onderzochten we de interbeoordelaarsbetrouwbaarheid van de START:AV bij gebruik in de Jeugdzorg^{Plus}. Voor deze studie werd de START:AV enerzijds door behandelcoördinatoren ingevuld, als onderdeel van het behandelproces, en anderzijds door onderzoekers als onafhankelijke tweede beoordeling. Zoals verwacht, op basis van eerder onderzoek, vonden we in onze praktijkstudie een lagere interbeoordelaarsbetrouwbaarheid dan de studies waarbij de START:AV alleen voor onderzoeksdoeleinden ingevuld werd. De interbeoordelaarsbetrouwbaarheid was onvoldoende voor de meeste items, zowel bij de krachten als bij de kwetsbaarheden. Hetzelfde werd gevonden voor de totaalscores: de betrouwbaarheid was onvoldoende voor de totaalscore van krachten en matig voor de totaalscore van kwetsbaarheden. Echter, de betrouwbaarheid van de eindoordelen (laag, matig, of hoog risico) was matig tot goed voor de meeste schadelijke uitkomsten. Alleen voor ongeoorloofde afwezigheden bleek dat de inschatting van het risico voor de nabije toekomst in deze steekproef niet betrouwbaar gemaakt kon worden. Dat de overige eindoordelen voldoende betrouwbaar zijn in de praktijk is bemoedigend aangezien ze cruciaal zijn voor beslissingen inzake behandeling en risicohantering (bijv. besluiten over begeleid versus onbegeleid verlof). Tot slot was de interbeoordelaarsbetrouwbaarheid van de voorgeschiedenis van de schadelijke uitkomsten matig tot goed voor de meeste uitkomsten. Omdat we in deze studie gebruik maakten van verschillende typen beoordelaars, namelijk behandelcoördinatoren en onderzoekers, konden we verschillen tussen beide groepen observeren. Zo bleek dat ze andere informatie gebruikten bij het coderen van de START:AV: de behandelcoördinatoren vertrouwden meer op hun herinneringen aan het dossier en de gebeurtenissen van de afgelopen periode en maakten, naar eigen zeggen, weinig gebruik van het dossier, terwijl de onderzoekers alleen gebruik maakten van dossierinformatie. Dit verschil in geraadpleegde informatie had invloed op de risicotaxatie: onderzoekers registreerden vaker dan behandelcoördinatoren incidenten in de voorgeschiedenis op basis hun dossierstudie, terwijl behandelcoördinatoren meer items konden invullen (d.w.z., minder items leeg lieten dan de onderzoekers) op basis van de informatie uit hun dagelijkse interacties met de jongere en het behandelteam. Beoordelaars met verschillende rollen in de behandeling van de jeugdige benutten dus andere informatiebronnen, en om hiervan te profiteren bij de risicotaxatie raden we de consensus-benadering aan. Dit houdt in dat verschillende zorgmedewerkers (bijv. mentor, behandelcoördinator, diagnosticus) onafhankelijk van elkaar de START:AV invullen en daarna een overleg organiseren waarbij de scores besproken worden en men tot een gezamenlijke (d.w.z., consensus) risicotaxatie komt. Naast deze aanbeveling beschrijven we andere mogelijkheden om in een behandelsetting de interbeoordelaarsbetrouwbaarheid van de START:AV op peil te houden.

In **Hoofdstuk 6** evalueerden we een tweede belangrijke psychometrische eigenschap van de START:AV: de predictieve validiteit. Evenals bij de interbeoordelaarsbetrouwbaarheid waren we vooral geïnteresseerd in de voorspellende waarde van de START:AV *in de praktijk*; in ons geval ingevuld door behandelcoördinatoren en met reële implicaties voor de (behandeling van de) jongere. We onderzochten of de totaalscores, de voorgeschiedenis, en de eindoordelen voorspellend waren voor incidenten binnen de instelling in de vier maanden volgend op de risicotaxatie. Deze incidenten werden geïdentificeerd in de dagelijkse rapportages over de jongeren geschreven door groepsleiding, en vervolgens gecodeerd met behulp van de START Outcome Scale: Adolescent Version Institutional (SOS-AVI; De Beuf et al., 2018). Via dit formulier werden negen typen incidenten gecodeerd: fysieke agressie naar anderen, niet-gewelddadig delict, ongeoorloofde afwezigheid (van school, leefgroep, behandeling), zelfbeschadiging, slachtofferschap, middelengebruik, suïcidaal gedrag, zelfverwaarlozing, en overtreding van de instellingsregels. Een onderzoeksassistent documenteerde wanneer en hoe vaak een bepaald type incident voorkwam in de vier maanden na de risicotaxatie, en dit werd de uitkomstmaat in deze studie. Uit de resultaten bleek dat de meeste eindoordelen inderdaad voorspellend waren voor incidenten van fysieke agressie, overtreding van de instellingsregels, middelenmisbruik, zelfbeschadiging, en slachtofferschap in de vier maanden na de risicotaxatie. Net als bij voorgaande START:AV-studies vonden we geen voorspellende waarde voor de eindoordelen van ongeoorloofde afwezigheid en zelfverwaarlozing. Verder was het eindoordeel van niet-gewelddadige delict niet voorspellend voor dergelijke incidenten (bijv. diefstal, vandalisme). Een voorgeschiedenis van bepaalde incidenten was voorspellend voor nieuwe incidenten van hetzelfde type in de nabije toekomst. Dit was het geval voor de meeste schadelijke uitkomsten behalve voor niet-gewelddadig delict, ongeoorloofde afwezigheid, en zelfverwaarlozing. Voor middelenmisbruik, zelfbeschadiging, en slachtofferschap onderzochten we tevens of de eindoordelen een significante meerwaarde hadden bovenop de voorspellende waarde van de voorgeschiedenis. Dit was niet het geval. Om onze studie te vergelijken met eerdere studies onderzochten we ook de voorspellende waarde van de totaalscores. Hierbij dient opgemerkt te worden dat dit voor de praktijk niet zo'n relevante analyse is, omdat er voor behandeldoeleinden geen START:AV-totaalscores berekend worden. In tegenstelling tot eerdere studies met de START:AV, hadden de totaalscores van zowel krachten als kwetsbaarheden weinig voorspellende waarde in onze steekproef. Beide totaalscores voorspelden wel institutionele overtredingen op korte termijn en de totaalscore voor krachten was eveneens voorspellend voor incidenten van fysieke agressie.

+

Het laatste hoofdstuk, **Hoofdstuk 7**, geeft een overzicht van de belangrijkste bevindingen uit dit proefschrift in de context van de beperkingen en de sterke aspecten van het onderzoek. Op basis van deze bevindingen worden aanbevelingen voor de praktijk geformuleerd, zowel voor het implementatieproces als het effectieve gebruik van de START:AV. Het hoofdstuk

eindigt met suggesties voor toekomstig onderzoek om meer inzicht te krijgen in de factoren die bijdragen aan een effectieve implementatie van de START:AV (en andere risicotaxatie-instrumenten) in de praktijk.

Algemene Conclusie

Dit proefschrift geeft inzicht in de belemmeringen en mogelijkheden die zich (kunnen) voordoen tijdens een START:AV-implementatie, alsook in de verschillende implementatie-uitkomsten die geëvalueerd kunnen worden om de effectiviteit van de implementatie in kaart te brengen. Daarnaast vonden we bemoedigende resultaten voor de betrouwbaarheid en voorspellende waarde van de START:AV-eindoordelen geformuleerd door behandelaars als onderdeel van het behandelproces in een Jeugdzorg^{Plus}-instelling. Met name de eindoordelen van fysieke agressie naar anderen, overtreding van de instellingeregels, middelenmisbruik, zelfbeschadiging, en slachtofferschap bleken significant voorspellend voor nieuwe incidenten van dezelfde aard tijdens de behandeling. Voor de voorspelling van andere incidenten, zoals niet-gewelddadig delicten, ongeoorloofde afwezigheid, en zelfverwaarlozing op basis van de START:AV vonden we onvoldoende evidentie. Verder vonden we in ons praktijkonderzoek onvoldoende interbeoordelaarsbetrouwbaarheid voor de meeste START:AV-items en de totaalscore voor krachten. Zowel de totaalscores voor krachten als kwetsbaarheden bleken tevens onvoldoende voorspellend voor de meeste schadelijke uitkomsten die we onderzochten in onze steekproef. Voor het gebruik van de START:AV in een behandelsetting heeft dit geen ernstige implicaties aangezien voor klinische doeleinden geen totaalscores berekend worden en de eindoordelen de centrale voorspellers zijn.

Tot slot kunnen we op basis van het huidige proefschrift twee algemene adviezen formuleren die relevant zijn voor het integreren en toepassen van risicotaxatie-instrumenten in de praktijk, namelijk:

1. Modellen en kaders uit de implementatiewetenschap zijn zinvol voor de implementatie van risicotaxatie-instrumenten. Een gestructureerde en evidence-based implementatie kan de kwaliteit van een instrument in de praktijk verbeteren waardoor potentieel de effectiviteit van het instrument voor het verminderen van herhaling van schadelijke uitkomsten toeneemt.
2. Steeds meer onderzoek wijst uit dat risicotaxatie-instrumenten minder goed presteren wanneer ze in de praktijk ingevuld worden door behandelaars en/of andere professionals in vergelijking met risicotaxaties die uitgevoerd worden door onderzoekers en/of onderzoeksassistenten onder meer gecontroleerde omstandigheden. De bevindingen in dit proefschrift sluiten hierbij aan. Dit bevestigt niet alleen het belang van (meer) praktijkonderzoek met risicotaxatie-instrumenten, maar ook de noodzaak om bij een nieuwe implementatie steeds de "lokale" psychometrische eigenschappen van een instrument te evalueren.