

# Risico op afbreken van een opleiding in de eerste twee jaar van het mbo

J. Vugteveen & A.C. Timmermans

## Samenvatting

Ongeveer negentig procent van de leerlingen die een vmbo-opleiding hebben afgerond, vervolgt hun onderwijsloopbaan in het mbo. Daar verlaten ieder jaar duizenden jongeren het onderwijs zonder startkwalificatie. Om het afbreken van opleidingen te verklaren zijn gedurende de afgelopen vijftig jaar verschillende theoretische modellen voorgesteld (bijv., Tinto, 1975; Finn, 1989; Battin-Pearson et al., 2000). In de huidige studie is van 721 mbo-studenten onderzocht welke factoren uit de theoretische modellen, en enkele aanvullingen daarop, van voorspellende waarde zijn voor het afbreken van een opleiding gedurende de eerste twee leerjaren van het mbo. Daartoe is gebruik gemaakt van schooladministratiegegevens en is verspreid over een periode van twee jaar vier keer een vragenlijst aan studenten voorgelegd. Ruim 42 procent van de studenten ( $n = 304$ ) brak hun mbo-opleiding af. Uit de resultaten van een survivalanalyse bleek de prevalentie van het afbreken van de mbo-opleiding rond of vlak na de overgang van het eerste naar het tweede leerjaar het grootst. Een verlaagd risico op het afbreken van de eerst gekozen mbo-opleiding lijkt geassocieerd te zijn met een hoger opleidingsniveau van ouders, het hebben gevolgd van een vmbo-opleiding in de sector Economie (t.o.v. de sector Zorg & Welzijn), een lagere spijbelfrequentie, en het kwalificatieplichtig zijn. Er is onvoldoende bewijs gevonden voor een eventuele rol van andere over tijd variërende factoren, zoals motivatie en de mate waarin de student het idee heeft zelf sturing aan zijn/haar leertraject te kunnen geven. De afwezigheid van bewijs voor de invloed van deze factoren ligt mogelijk in het simultaan toetsen van verschillende variabelen of in de van de oorspronkelijke constructen afwijkende operationalisaties.

**Kernwoorden:** beroepsonderwijs (mbo), afbreken opleiding, risicofactoren

## 1 Inleiding

Sinds de invoering van de kwalificatieplicht in 2007 zijn jongeren bij wet verplicht tot hun achttiende levensjaar onderwijs te volgen, tenzij ze eerder een startkwalificatie (minimaal een havo-, vwo- of mbo-2 diploma) behalen. Volgens voorlopige cijfers van het Ministerie van Onderwijs, Cultuur en Wetenschap (2016) hebben in studiejaar 2014-2015 desondanks bijna 24.500 jongeren het onderwijs verlaten zonder startkwalificatie, waarvan ruim 19 duizend jongeren met een opleiding in het middelbaar beroepsonderwijs (mbo) stopten. Uit deze cijfers blijkt de bijzondere positie van het mbo: in geen enkele andere onderwijssector in het Nederlandse onderwijssysteem is de prevalentie van voortijdig schoolverlaten zo groot. Het voortijdig verlaten van het onderwijs is zorgelijk omdat jongeren zonder startkwalificatie vaker zijn aangewezen op het uitvoeren van ongeschoold werk en hun toekomstperspectief beperkt is (Ministerie van Onderwijs, Cultuur en Wetenschap, 2002). Dat jongeren toch hun opleiding afbreken is vaak het resultaat van een samenloop van omstandigheden (Tinto, 1975; Finn, 1989; Ensminger & Slusarcick, 1992; Alexander, Entwisle, & Kabbani, 2001).

In de onderhavige studie staat het afbreken van de gekozen mbo-opleiding gedurende de eerste twee leerjaren na de overgang vanuit het vmbo centraal. In de afgelopen vijftig jaar zijn reeds verschillende modellen voor het voorspellen van het voorzetten of verlaten van onderwijs en de daarbij horende processen voorgesteld (bijv., Rumberger & Lim, 2008). Talrijke onderzoeken naar schoolverlaten stoelen op (variëaties van) twee van deze theoretische modellen: het longitudinale model voor drop-out van Tinto (1975) en het participatie-identificatiemodel van Finn (1989). Een derde recenter

theoretisch model, waarin meer aandacht wordt besteed aan probleemgedrag van jongeren, is het samengestelde model van Battin-Pearson et al. (2000). De huidige studie beoogt een samengestelde model voor het afbreken van de mbo-opleiding te toetsen in de Nederlandse mbo-context (zie paragraaf 1.2). Dit model is samengesteld uit elementen van deze drie reeds bestaande modellen. We zullen derhalve eerste de bestaande theoretische modellen en het samengestelde model beschrijven, waarna de empirische toetsing van het samengestelde model besproken wordt.

### Bestaande theoretische Modellen

#### *Tinto's longitudinale model voor drop-out (1975)*

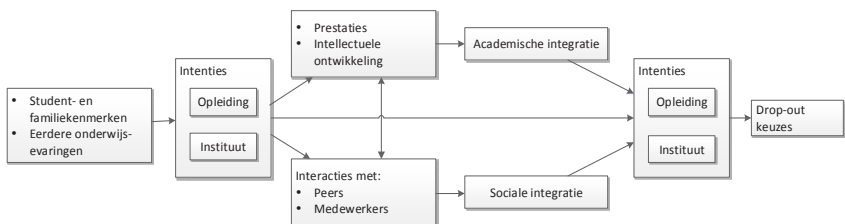
Het model van Tinto, weergegeven in Figuur 1, beschrijft het proces dat resulteert in het afbreken of voortzetten van een opleiding in het hoger onderwijs. De basis van het model wordt gevormd door eerdere onderwijservaringen van studenten, achtergronden en familiekenmerken. Deze kenmerken zijn van invloed op de twee in het model geëxpliciteerde dimensies van integratie: academische integratie en sociale integratie. Academische integratie is een product van de behaalde prestaties en de intellectuele ontwikkeling van de student, en is mede afhankelijk van de mate waarin een student zich bij aanvang voorneemt de gekozen opleiding af te maken. Sociale integratie is het resultaat van interactie van de student met medestudenten en docenten (of andere medewerkers binnen een onderwijsinstituut). Sociale integratie is mede afhankelijk van de mate waarin

de student bij aanvang van de opleiding van plan was die binnen het gekozen instituut af te maken. Gezamenlijk bepalen de mate van academische en sociale integratie of een student de opleiding voortzet of afbreekt.

Sinds de publicatie van het model is binnen verschillende onderwijscontexten getracht het model van Tinto (bijv., Terenzini & Pascarella, 1977; Pascarella & Chapman, 1983; Bers & Smith, 1991; Brunsten, Davies, Shevlin, & Bracken, 2000; Liu & Liu, 2000) of een deel daarvan (Thomas, 2000) te valideren. Zowel in de twee studies waar Pascarella bij betrokken was als in de studie van Bers en Smith werd gevonden dat studenten die het Amerikaanse college-onderwijs verlaten inderdaad minder academische en sociale integratie vertoonden dan studenten die hun opleiding voortzetten. Liu en Liu (2000) kwamen wat academische integratie betreft tot dezelfde conclusie, maar zij vonden onvoldoende bewijs dat sociale integratie samenhangt met de keuze om te blijven. Thomas (2000) richtte zich op de sociale dimensie van het model van Tinto en vond dat niet alleen het aantal leeftijdsgenoten waar studenten connecties mee hebben van belang bleek, maar ook dat diversiteit binnen het sociale netwerk, bijvoorbeeld qua opleiding, bevorderlijk is. Daarentegen trekken Brunsten, Davies, Shevlin en Bracken (2000) in twijfel of het model van Tinto het proces tot het wel of niet verlaten van het onderwijs voldoende beschrijft. Een groot deel van de studenten die het hoger onderwijs verlaat, doet dat volgens hen zo vroeg in de opleiding dat te betwijfelen valt of de twee vormen van integratie in die korte tijd tot stand kunnen komen

Figuur 1

*Bewerking van Tinto's oorspronkelijke longitudinale model voor drop-out.*



en een (grote) rol kunnen spelen in de beslissing tot het voortzetten of afbreken van een opleiding. Mogelijk geldt dat ook voor het afbreken van opleidingen in het Nederlandse beroepsonderwijs, aangezien het grootste deel van de uitvallers de mbo-opleiding in het eerste leerjaar verlaten (Vugteveen, Timmermans, Korpershoek, van Rooijen, & Opdenakker, 2016).

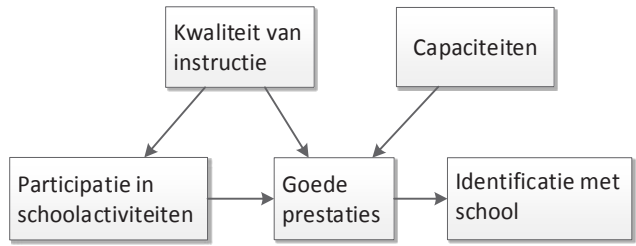
Net als in het model van Tinto, is in het merendeel van de studies naar schoolverlaten een rol weggelegd voor sociaal-demografische kenmerken van studenten. Het meest onderzocht zijn sekse, sociaaleconomische status en etniciteit van studenten. De consensus is dat jongens (bijv., Stage, 1989; Battin-Pearson et al., 2000; Elffers, 2011; Theunissen, de Man, Verdonk, Bosma, & Feron, 2015), jongeren met een lage sociaaleconomische status (bijv., Alexander, Entwisle, & Kabbani, 2001; Rumberger & Lim, 2008) en jongeren die tot een etnische minderheidsgroep behoren (bijv., Rumberger & Lim, 2008; Stevens, Clycq, Timmerman, & Van Houtte, 2011) hoger risico lopen om uit te vallen. Toch zijn de resultaten van voorgaand onderzoek niet volledig eenduidig en mogelijk deels afhankelijk van de manier waarop de rol van voorspellende variabelen wordt onderzocht (bijv. het al dan niet toevoegen van covariaten of interactie-effecten tussen voorspellende variabelen). Zo duiden bijvoorbeeld de resultaten van de studie van Luyten en collega's op afwezigheid van een hoofdeffect van etniciteit, maar komt etniciteit wel in beeld in een interactie-effect tussen etniciteit en het opleidingsniveau van de ouders (Luyten, Bosker, Dekker, & Derks, 2003).

In het model van Tinto zijn behaalde cijfers en eerder verworven academische en sociale bekwaamheden opgenomen. De relatie tussen enerzijds het afbreken van een opleiding en anderzijds capaciteiten, eerder opgedane onderwijservaringen en/of in het verleden behaalde prestaties komt veelvuldig in onderzoek aan de orde. Ethington (1990) vond in het beroepsonderwijs (in de Verenigde staten) een directe positieve relatie tussen in het voortgezet onderwijs behaalde prestaties en het voortzetten van opleidingen in het beroepsonderwijs. Elffers (2011) vond hetzelfde in het Nederlandse mbo. Een mogelijke verklaring voor dit verband is dat positieve onderwijservaringen uit het verleden het zelfvertrouwen van jongeren verhogen en het doorzettingsvermogen versterken (Dupin-Bryant, 2004).

*Finn's model voor het distantiëren van school (1989)*

Het Participatie-Identificatiemodel van Finn (zie Figuur 2) is toegespitst op studenten in high school in de Verenigde Staten (min of meer vergelijkbaar met de bovenbouw van het Nederlandse voortgezet onderwijs). Het model illustreert een cyclus waarin actieve participatie van studenten voorwaardelijk is voor het behalen van goede prestaties, en goede prestaties op hun beurt van belang zijn voor identificatie met school. De cyclus wordt voltooid doordat de mate waarin studenten zich identificeren met school indicatief is voor de mate van actieve participatie bij schoolactiviteiten. De capaciteiten van jongeren en de kwaliteit van de gegeven instructie beïnvloeden de beschreven cyclus: om onderwijsactiviteiten succesvol uit te kunnen voeren moet de instructie aan de stu-

Figuur 2  
Bewerking van Finn's oorspronkelijke participatie-identificatie model.



dent van voldoende niveau zijn. Tevens moet de student over de minimaal benodigde capaciteiten beschikken om op een zinvolle manier aan de activiteiten deel te kunnen nemen. Wanneer de cyclus op ernstige wijze verstoord raakt, bijvoorbeeld doordat het niveau van het onderwijs niet aansluit bij de capaciteiten van de student, zal het proces mogelijk in negatieve zin in gang gezet worden waarbij een student steeds minder participeert in schoolse activiteiten, steeds slechter presteert en zich steeds minder met school identificeert. Als de negatieve cyclus niet op tijd wordt doorbroken, zal het proces eindigen in het afbreken van een opleiding.

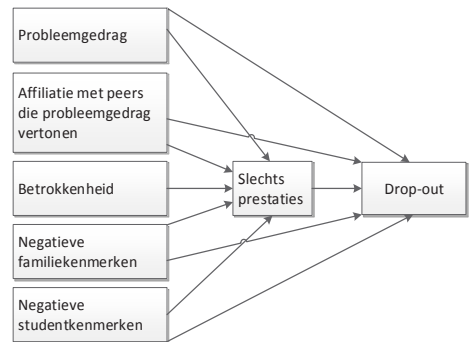
Ten opzichte van het model van Tinto is het model van Finn uniek door inclusie van de participatie van studenten en de kwaliteit van instructie, maar de modellen hebben met elkaar gemeen dat ze naast huidige prestaties ook capaciteiten van studenten meenemen. Het model van Finn veronderstelt dat capaciteiten van studenten de beschreven cyclus beïnvloeden middels samenhang met de huidige prestaties van studenten. In het model van Tinto spelen capaciteiten daarentegen een indirecte rol via de mate waarin studenten van plan zijn hun opleiding af te maken. Tinto hanteert daarnaast een smallere definitie van capaciteiten: prestaties en verworven vaardigheden voorafgaand aan de start van de nieuwe opleiding.

Participatie en identificatie zijn de twee centrale constructen in het model van Finn. In later werk (1993) beschrijft Finn deze twee constructen niet langer als participatie en identificatie, maar als gedragsmatige en emotionele betrokkenheid bij school. Het construct betrokkenheid heeft ten grondslag gelegen aan talrijke studies in verschillende onderwijscontexten. De consensus is dat betrokkenheid bij school een sleutelrol vervult in het al dan niet voortzetten van opleidingen (bijv., Fredricks, Blumenfeld, & Paris, 2004; Johnson, McGue, & Iacono, 2006; Janosz, Archambault, Morizot, & Pagani, 2008; Rumberger & Lim, 2008). Uit onderzoek van Hausmann, Schofield en Woods (2007) en van O’Keeffe (2013) bleek dat het gevoel van betrokkenheid bij school, een vorm van emotionele betrokkenheid, gerela-

teerd is aan intenties om een opleiding voort te zetten. Archambault, Janosz, Fallu en Pagani (2009) vonden daarentegen in een onderzoek naar vroege uitval in high school dat enkel gedragsmatige betrokkenheid voorspellend is in het proces tot het al dan niet verlaten van het onderwijs. De auteurs merken daar bij op dat ze vermoeden dat de eventuele rol van emotionele betrokkenheid mogelijk gemedieerd wordt door gedragsmatige betrokkenheid.

*Battin-Pearson’s model voor drop-out (2000)*

Het model van Battin-Pearson en collega’s, weergegeven in Figuur 3, richt zich specifiek op uitval in de eerste jaren van high school. Dit model betreft, in tegenstelling tot het model van Finn, een momentopname in plaats van een procesbeschrijving van het al dan niet verlaten van het onderwijs. Er wordt een mediërende rol van prestaties verondersteld in de relatie tussen enerzijds voortijdig schoolverlaten en anderzijds probleemgedrag, negatieve invloed van leeftijdsgenoten, familiekenmerken, studentkenmerken, en betrokkenheid bij school. Daarnaast bevat het model rechtstreekse relaties tussen schoolverlaten en de genoemde factoren, met uitzondering van betrokkenheid bij school.



Figuur 3  
Bewerking van het oorspronkelijke drop-out model van Battin-Pearson en collega’s.

Het model van Battin-Pearson is uniek door de toevoeging van probleemgedrag dat zowel direct, als indirect via prestaties,

samenhangt met het verlaten van het onderwijs. Met probleemgedrag worden misdragingen binnen school, overtredingen van de wet buiten school, onverantwoord seksueel gedrag en drug- en alcoholgebruik aangeduid (Rumberger, 2011). Er is consensus over dat het overtreden van de wet samenhangt met een hoger risico op drop-out (bijv., Sweeten, 2006; Rumberger & Lim, 2008; Hirschfield, 2009; Ter Bogt, Van Lieshout, Doornwaard, & Eijkemans, 2009). Ook onverantwoorde seksuele activiteit wordt algemeen erkend als risicofactor voor schoolverlaten (Anderson, 1993; Freudenberg & Ruglis, 2007). De relatie tussen binnenschools probleemgedrag en het afbreken van opleidingen wordt, zeker naarmate jongeren ouder zijn, veelal onderzocht aan de hand van spijbelgedrag. De meest voorkomende conclusie is dat spijbelgedrag indicatief is voor voortijdig schoolverlaten (bijv., Wehlage & Rutter, 1986; Garry, 1996; Henry, 2007).

Hoewel in de voorgaande alinea geschreven werd dat de inclusie van probleemgedrag uniek is ten opzichte van de modellen van Tinto en Finn, is het van belang daar enige nuance in aan te brengen. Het vertonen van probleemgedrag op school of daarbuiten kan ten dele worden gezien als het tegenovergestelde van Finn's gedragsmatige betrokkenheid (Archambault et al., 2009). Daarnaast bestaat er mogelijk een relatie tussen probleemgedrag en de door Tinto beschreven sociale integratie van studenten. Zowel Thomas (2000) als Battin-Pearson (2000) stellen dat de kans op probleemgedrag kleiner is wanneer studenten deel uitmaken van een sociaal netwerk waarbinnen weinig probleemgedrag voorkomt dan wanneer zij deel uitmaken van een netwerk waarin meer studenten zich er schuldig aan maken.

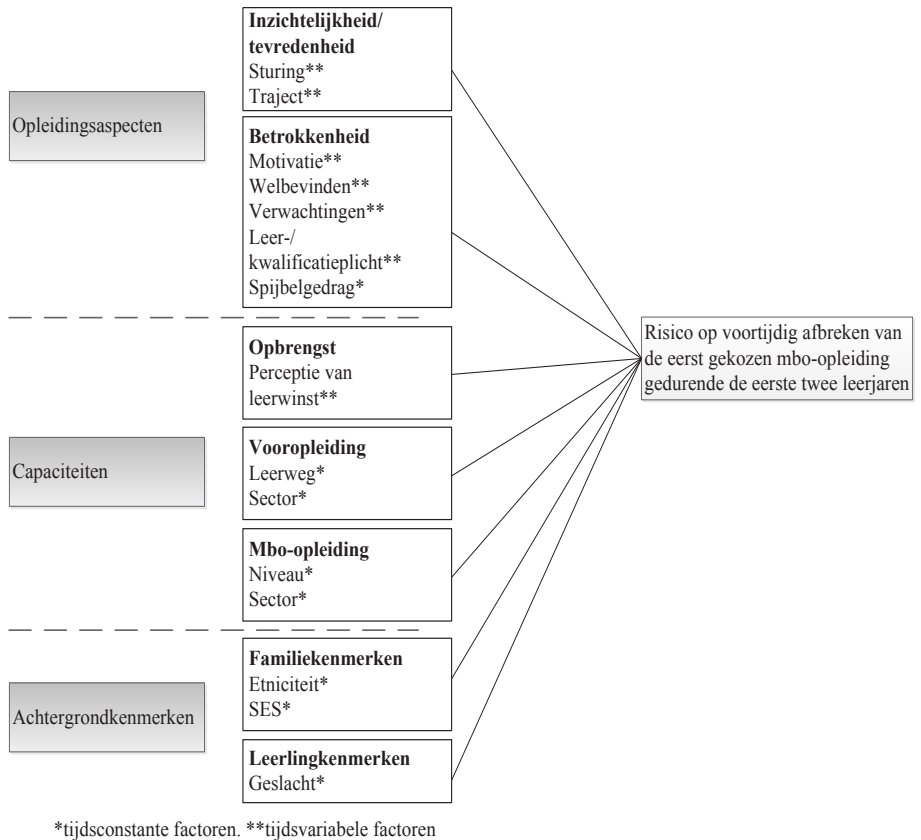
### **1.2 Samenvoeging van Modellen voor Voor- spellen van Uitval**

Het model dat ten grondslag ligt aan de onderhavige studie (zie Figuur 4), bestaat uit een combinatie van elementen uit de drie besproken modellen en enkele aanvullingen daarop. Voor zover wij hebben kunnen achterhalen, is nog niet eerder getracht de relatie tussen deze modellen te onderzoeken of de modellen samen te voegen tot een overkoepe-

lend model en dit model te valideren voor de Nederlandse mbo-context. In tegenstelling tot de reeds besproken modellen, die uitval in high school of in higher education (college) in de Verenigde Staten in kaart brengen, heeft het opgestelde model betrekking op de Nederlandse mbo-context. De onderwijscontexten wijken op bepaalde vlakken van elkaar af: de etnische minderheidsgroepen in de Verenigde Staten zijn anders dan de minderheidsgroepen in Nederland. Bovendien zijn Nederlandse mbo-studenten iets ouder dan Amerikaanse high school studenten, maar ze zijn niet per se ouder dan studenten in het Amerikaanse higher education. Ook betreft het Nederlandse mbo mogelijk een andere onderwijsniveaurange dan de Amerikaanse high school. Desondanks is het aannemelijk dat Nederlandse mbo-studenten in veel opzichten hetzelfde proces tot uitval of het voortzetten van hun opleiding doormaken. Veel factoren uit de drie besproken modellen betreffen namelijk psychologische factoren, waarvan het aannemelijk is dat die in de Nederlandse mbo-context en de Amerikaanse context een min of meer vergelijkbare rol spelen. Het model bestaat derhalve uit factoren die uit voorgaande modellen komen, aangevuld met voor Nederland specifieke kenmerken die mogelijk van belang zijn.

Aansluitend op het model van Finn, wordt de kern van het samengesteld model gevormd door de mogelijk over tijd heen variërende kans op het afbreken van een opleiding. De belangrijkste reden om te kiezen voor een longitudinaal model in plaats van een momentopname is dat het afbreken van een opleiding het gevolg is van een samenloop van omstandigheden die gedurende de tijd verandert en een variërende kans op het afbreken van een opleiding met zich meebrengt (Ensminger & Slusarcick, 1992; Alexander, Entwisle, & Kabbani, 2001). Het model komt daarmee tevens tegemoet aan de in Nederland voorkomende variatie in momenten van uitval uit mbo-opleidingen (Vugteveen et al., 2016).

De in het model opgenomen factoren die constant zijn gedurende de tijd zijn student-/familiekenmerken en tijdens een vooropleiding verworven kennis en vaardigheden.



Figuur 4  
Schematisch model van risico op het afbreken van een opleiding.

Evenals in de modellen van Tinto en Battin-Pearson het geval is, en vergelijkbaar met voorgaand onderzoek naar sociaal-demografische student- en familiekenmerken, zijn sekse, etniciteit en sociaaleconomische status in het samengestelde model opgenomen. Daarnaast zijn ook eerder verworven kennis en vaardigheden onderdeel van het model. In plaats van overwegend gebruik te maken van cijfers behaald tijdens een vooropleiding, zoals Tinto en Finn doen, is in het samengestelde model het niveau en de sector van de gevolgde vmbo-vooropleiding meegenomen. Uit eerder onderzoek is namelijk gebleken dat de mate van overeenkomst tussen vmbo-vooropleiding en de gekozen mbo-opleiding wat betreft het gekozen niveau en de gekozen sector sterk gerelateerd is aan uitval in het mbo (Vugteveen, et al., 2016).

Daarnaast zijn over tijd variërende factoren in het model opgenomen: prestaties, tevredenheid, intenties/verwachtingen, motivatie, welbevinden en spijbelgedrag. Zowel Tinto als Finn en Battin-Pearson operationaliseren huidige prestaties op een relatief objectieve manier: ze gebruiken cijfers op toetsen. Gegeven dat het huidige onderzoek heeft plaatsgevonden binnen meerdere opleidingen uit verschillende sectoren en van verschillende niveaus (met ieder eigen eindtermen, toetsvormen en -momenten etc.) was het niet mogelijk om gedurende de eerste twee jaar van de opleidingen herhaaldelijk prestaties te meten en die met elkaar te vergelijken. Daarom is gebruik gemaakt van de eigen inschatting van leerwinst van de student. De tevredenheid van studenten over de gekozen opleiding en het onderwijsinstituut is, evenals

de zojuist beschreven inschatting van leerwinst, in het samengesteld model opgenomen als maat voor Tinto's academische integratie. Verwachtingen die studenten hebben van het succesvol en naar tevredenheid afronden van hun opleiding zijn, in navolging van intenties van studenten in het model van Tinto, in het model opgenomen. De in het samengestelde model opgenomen motivatie voor school en voor schoolwerk is vergelijkbaar met emotionele betrokkenheid uit het model van Finn. Welbevinden is, net als Tinto's sociale integratie, een indicatie voor de mate waarin een student zich op sociaal vlak op zijn plek voelt binnen de opleiding en het opleidingsinstituut. Tot slot zijn gedragsmatige betrokkenheid uit het model van Finn en binnenschools probleemgedrag uit het model van Battin-Pearson in kaart gebracht aan de hand van spijbelgedrag.

Naast intrinsieke betrokkenheid met de opleiding, te meten aan de hand van tevredenheid, motivatie, welbevinden en verwachtingen van studenten, is er in de Nederlandse mbo-context ook sprake van extrinsiek opgelegde betrokkenheid. Kwalificatieplicht, zoals we die in Nederland kennen, komt in geen van de drie modellen voor, maar kan worden beschouwd als maat voor opgelegde of extrinsieke gedragsmatige betrokkenheid. De reden dat leerplicht weinig in onderzoek wordt meegenomen is mogelijk dat leerplicht in de Verenigde Staten niet landelijk maar per staat geregeld wordt (National Center for Education Statistics, z.j.), waardoor vergelijkingen op nationaal niveau lastiger zijn. In Nederland is leer-/kwalificatieplicht landelijk geregeld, maar ook hier wordt deze factor om onduidelijke reden, en voor zover bij ons bekend, niet frequent in onderzoek meegenomen. Het lijkt ons echter aannemelijk dat de verplichting tot het volgen van onderwijs wel een rol speelt in de keuzes die studenten maken met betrekking tot het al dan niet afbreken van de opleiding. Om recht te doen aan de Nederlandse onderwijscontext is kwalificatieplicht in het model opgenomen.

In deze studie zal het samengestelde model getoetst worden om antwoord te geven op de volgende onderzoeksvraag: Hoe ontwikkelt het risico op het afbreken van een opleiding

zich gedurende de eerste twee leerjaren van het mbo en welke factoren uit het samengestelde model zijn daarop van invloed?

## 2 Methode

### 2.1 Proefpersonen

In dit onderzoek is gebruik gemaakt van bestaande data die van 2007 tot en met 2010 in het kader van een onderzoek naar competentiegericht leren zijn verzameld binnen een ROC in het noorden van Nederland. Er zijn gegevens verzameld binnen acht opleidingen uit de beroepsopleidende leerweg (BOL) verdeeld over de sectoren Economie & Handel, Techniek en Zorg & Welzijn. Er zijn vier opleidingen op niveau 2 (Kok, Medewerker ICT, Helpende zorg/welzijn en Elektromonteur) en vier opleidingen op niveau 3/4 (Horeca ondernemer/manager, International Business Studies, Verzorgende/ verpleegkundige, Elektrotechniek/Middenkader engineering) geselecteerd. Per opleiding hebben twee achtereenvolgende studentcohorten aan het onderzoek meegewerkt: studenten die in augustus 2007 en studenten die in augustus 2008 aan één van de acht opleidingen zijn begonnen. Beide studentcohorten zijn in het eerste leerjaar in november en mei, en in het tweede leerjaar in november en april, bevraagd middels vragenlijsten. Deze vragenlijsten zijn ontworpen in het kader van het oorspronkelijke onderzoek naar competentiegericht leren en grotendeels gebaseerd op het COOL<sup>5-18</sup> cohortonderzoek met verschillende aanpassingen om het geschikt te maken voor het mbo. Daarnaast is de administratie van de onderwijsinstelling geraadpleegd om aanvullende achtergrondinformatie over studenten te verzamelen.

De analyses zijn uitgevoerd op basis van gegevens van 721 studenten die rechtstreeks vanuit het vmbo naar het mbo zijn doorgestroomd. Door ons te richten op rechtstreeks vanuit het vmbo instromende studenten, kunnen we het proces tot het voortzetten of verlaten van de eerst gekozen mbo-opleiding in kaart brengen en zijn de studenten binnen de onderzoeksgroep vergelijkbaar wat betreft onderwijsverleden.

## 2.2 Variabelen

Getalsmatige samenvattingen van de variabelen die betrekking hebben op student-/familiekenmerken, capaciteiten, tevredenheid/inzicht en betrokkenheid zijn weergegeven in Tabel 1 (tijdsconstante variabelen) en Tabel 2 (tijdsvariërende variabelen). De interne consistentie van de variabelen, die zijn samengesteld uit meerdere items, is weergegeven in Tabel 3.

### *Het afbreken van een opleiding*

De administratiedata van de onderwijsinstelling zijn geraadpleegd om de uitschrijfdatum van uitvallende studenten te achterhalen. Gezien enige onnauwkeurigheid van de uitschrijfdatum in de administratiegegevens is voor iedere student, wanneer van toepassing, vastgesteld tussen welke twee meetmomenten de opleiding afgebroken werd.

### *Student- en familiekenmerken*

Kenmerken van de studenten en hun familie zijn op twee manieren verzameld. Informatie voor de variabele 'seks' is uit de administratiegegevens van de onderwijsinstelling gehaald. Informatie voor de variabelen 'ethniciteit' en 'opleidingsniveau van de ouders' is verzameld met behulp van enkele vragen die bij het eerste meetmoment aan studenten zijn voorgelegd. De variabele 'ethniciteit' is gebaseerd op de geboortelanden van de ouders van de student. Dit is vergelijkbaar met operationalisaties van ethniciteit in voorgaande, grootschalige Nederlands cohortonderzoeken, zoals PRIMA (bijv., Driessen, Van Langen, & Vierke, 2006), VOCL'99 (bijv., Kuypers & van der Werf, 2003) en COOL<sup>5-18</sup> (bijv., Zijlstra, Keuning, Keizer-Mittelhaeuser, Naaijer, & Timmermans, 2017). Er wordt onderscheid gemaakt tussen studenten waarvan tenminste één ouder in Nederland geboren is en studenten waarvan beide ouders in een ander land dan Nederland geboren zijn (Karsen, Van der Veen, & Volman, 2015). Het opleidingsniveau van ouders wordt meegenomen als indicator van de sociaaleconomische status van het gezin. Ook deze operationalisatie is direct gebaseerd op de zojuist genoemde cohortonderzoeken. Studenten is gevraagd het opleidingsniveau van de hoogstopleide

ouder te noteren. De categorieën zijn: geen dan wel alleen basisonderwijs, middelbaar onderwijs, of hoger- dan wel universitair onderwijs.

### *Capaciteiten*

De capaciteiten van studenten zijn in kaart gebracht aan de hand van enkele (voor)opleidingskenmerken en de perceptie van leerwinst van de studenten gedurende de opleiding. De kenmerken van de vmbo-vooropleiding, opgevraagd uit de administratie van de onderwijsinstelling, zijn samengevat in de variabelen 'vmbo-leerweg' en 'vmbo-sector'. De variabele 'vmbo-leerweg' onderscheidt vier vmbo-leerwegen: de basisberoepsgerichte leerweg (bl), de kaderberoepsgerichte leerweg (kl), de gemengde leerweg (gl) en, tenslotte, de theoretische leerweg (tl). De variabele 'vmbo-sector' bestaat eveneens uit vier categorieën: Techniek, Landbouw, Economie & Handel en Zorg & Welzijn.

De kenmerken van de gekozen mbo-opleiding zijn ook in twee variabelen gevat. In de variabele 'mbo-sector' worden drie sectoren onderscheiden: Economie, Techniek en Zorg & Welzijn. Er zijn om praktische redenen geen data verzameld bij mbo-opleidingen in de sector Landbouw. Daarnaast worden de gekozen mbo-opleidingen in de variabele 'mbo-niveau' onderverdeeld in niveau 1/2 en niveau 3/4 opleidingen.

De prestaties van studenten zijn in de huidige studie in kaart gebracht aan de hand van de perceptie die studenten over de relevantie van de geleerde kennis en vaardigheden, als ook over de hoeveelheid daarvan, hebben. De reden daarvoor is dat het effect van factoren als prestaties sterk afhankelijk is van de mate waarin studenten bereid zijn ze mee te laten wegen in hun keuze tot het voortzetten of afbreken van een opleiding (McKeown, Macdonell, & Bowman, 1993). Studenten die graag goed presteren en daar ook hun best voor doen, zullen hun keuze bijvoorbeeld meer laten beïnvloeden door slechte prestaties dan studenten die minder waarde hechten aan goede prestaties en zich daar ook in beperkte mate voor inzetten. Deze perceptie van studenten is samengevat in de variabele 'leerwinst'. De variabele is uit zes vragen



Tabel 1  
*Samenstelling steekproef op basis van tijdsconstante variabelen*

Student-/familiekenmerken		N	%
Sekse	Man	391	54.2
	Vrouw	330	45.8
	Totaal	721	100.0
Opleidingsniveau ouders	Geen of alleen basisonderwijs	13	2.6
	Middelbaar onderwijs	168	33.1
	Hoger of universitair onderwijs	326	64.3
	Totaal	507	100.0
Etniciteit	Tenminste één in Nederland geboren ouder	465	91.0
	Twee niet in Nederland geboren ouders	46	9.0
	Totaal	511	100.0
Capaciteiten		N	%
Leerweg vooropleiding	vmbo-bl	133	18.4
	vmbo-kl	205	28.4
	vmbo-gl	23	3.2
	vmbo-tl	360	49.9
	Totaal	721	100.0
Sector vooropleiding	Economie	290	40.4
	Landbouw	30	4.2
	Techniek	99	3.8
	Zorg & Welzijn	299	41.6
	Totaal	718	100.0
Betrokkenheid		N	%
Spijbelgedrag*	Nooit of vrijwel nooit	329	77.2
	Maandelijks een keer	64	15.0
	Wekelijks een keer of vaker	33	7.8
	Totaal	426	100.0

\* Om als over tijd variërende variabele in de analyse te worden meegenomen, moet een variabele op alle meetmomenten gemeten zijn. Spijbelgedrag is enkel op meetmoment 1 en 4 gemeten. Er is daarom gebruikt gemaakt van de gemiddelde spijbelrequentie per student

opgebouwd, waaronder bijvoorbeeld “*ik heb in deze opleiding veel praktische vaardigheden opgedaan die ik in mijn latere beroep kan gebruiken.*” Deze vragen zijn bij alle meetmomenten gesteld. Er is gebruik gemaakt van een vierpunts-antwoordschaal, lopend van ‘dat is helemaal niet zo’ (1) tot ‘dat is echt zo’ (4). Een tweede reden voor het gebruiken van de perceptie van leerlingen was dat het door de inclusie van verschillende inhoudelijke opleidingen van verschillende niveaus, met daarbij horende verschillen in vaardigheden, eindtermen,

toetsvormen en –momenten, niet mogelijk bleek een relevante meer objectieve prestatie-indicator te verkrijgen die op de vier meetmomenten beschikbaar was.

#### *Tevredenheid/inzicht*

De mate waarin studenten het idee hebben dat zij inzicht in hun opleiding hebben en zij zelf sturing aan hun leertraject kunnen geven, is op alle meetmomenten gemeten aan de hand van twee variabelen: ‘traject’ en ‘sturing’. Bij alle vragen onderliggend aan deze twee variabelen is gebruik gemaakt van een

Tabel 2

Beschrijvende statistiek van variabelen horend bij 'capaciteiten', 'tevredenheid' en 'betrokkenheid'

	Meetmomenten								
	Leerjaar 1				Leerjaar 2				
	November		Mei		November		April		
	N	M (s)	N	M (s)	N	M	N	M (s)	
<b>Capaciteiten</b>									
Leerwinst	531	2.81 (.48)	364	2.88 (.51)	321	2.85 (.50)	267	2.86 (.49)	
<b>Tevredenheid/inzicht</b>									
Traject	528	2.95 (.41)	364	2.88 (.45)	317	2.86 (.46)	266	2.87 (.44)	
Sturing	530	2.48 (.48)	367	2.42 (.43)	324	2.44 (.46)	268	2.41 (.42)	
<b>Betrokkenheid</b>									
Motivatie	523	2.97 (.48)	347	2.88 (.49)	317	2.92 (.44)	260	2.91 (.42)	
Welbevinden	527	3.08 (.47)	347	2.98 (.52)	320	2.95 (.44)	265	2.98 (.44)	
Verwachtingen	532	1.93 (.56)	360	2.06 (.65)	323	1.94 (.55)	270	1.97 (.55)	
	N	%	N	%	N	%	N	%	
Kwalificatieplicht*	Ja	227	31.5	291	40.4	440	61.0	598	82.9
	Nee	494	68.5	430	59.6	281	39.0	123	17.1

\* In verband met afwijkend meetniveau apart van de overige betrokkenheidsindicatoren weergegeven

vierpunts-antwoordschaal, lopend van 'dat is helemaal niet zo' (1) tot 'dat is echt zo' (4). De variabele 'traject' geeft een beeld van de mate waarin studenten tevreden zijn met de mate van begeleiding en de mate waarin de opleiding bijdraagt aan hun toekomstplannen. De variabele omvat de antwoorden op acht vragen, zoals bijvoorbeeld "Ik zie precies voor me hoe ik door deze opleiding kom" en "Je wordt hier goed begeleid." De variabele 'sturing' betreft de mate waarin studenten het idee hebben zelf invloed op hun leertraject te hebben. De variabele is aan de hand van vijf vragen gemeten, zoals "Leraren bepalen hier voor je wat je moet doen."

#### *Betrokkenheid*

In hoeverre studenten zich intrinsiek betrokken voelen bij hun opleiding, is gemeten aan de hand van vijf variabelen 'motivatie', 'welbevinden', 'verwachtingen', 'spijbelgedrag' en 'kwalificatieplicht'. De variabelen 'motivatie', 'welbevinden' en 'verwachtingen' zijn op alle meetmomenten gemeten. Daarbij is een vierpunts-antwoordschaal, lopend van 'dat is helemaal niet zo'

(1) tot 'dat is echt zo' (4) gebruikt. De variabele 'motivatie' geeft een indicatie van de mate waarin een student gemotiveerd is om naar school te gaan en met schoolwerk bezig te zijn. De variabele is opgebouwd uit negen vragen, waaronder "Na het weekend heb ik wel zin om weer naar school te gaan" en "Ik vind deze opleiding niet zo boeiend." De variabele 'welbevinden' geeft een beeld van de mate waarin een student zich op zijn plek voelt binnen de opleiding en het opleidingsinstituut. De variabele omvat de antwoorden op tien vragen, zoals "Ik voel me goed op deze opleiding" en "Ik voel me thuis op deze school." De variabele 'verwachtingen' geeft aan de hand van zes vragen, zoals "Deze opleiding is heel anders dan ik dacht", een indicatie van de mate waarin de opleiding aan verwachtingen van de student voldoet en of de student denkt de opleiding af te maken. Ook is op het eerste en laatste meetmoment informatie verzameld op basis waarvan de variabele 'spijbelgedrag' is vormgegeven. Deze variabele bevat informatie over of studenten gedurende de twee leerjaren gemiddeld nooit of een enkele keer, elke maand wel

een keer, of elke week wel een keer hebben gespijbeld.

De variabele ‘kwalificatieplicht’, een maat voor door de overheid opgelegde betrokkenheid, is op basis van de geboortedatum van de student uit de administratiegegevens in kaart gebracht. Kwalificatieplicht is op 1 augustus 2007 ingevoerd. Van de 721 studenten in de dataset, waren 127 bij aanvang van hun opleiding reeds achttien jaar of ouder, zij waren niet kwalificatieplichtig. Daarnaast is er sprake van een groep van 71 studenten die tussen 1 augustus 2006 en 31 juli 2007 zeventien werden. Zij vallen onder een overgangsregeling waardoor ook zij bij aanvang van de mbo-opleiding niet kwalificatieplichtig waren. De overige studenten in de dataset waren allen bij aanvang van de opleiding verplicht tot hun achttiende levensjaar of tot het behalen van een startkwalificatie onderwijs te blijven volgen. Op ieder meetmoment is vastgesteld welke studenten kwalificatieplichtig waren.

### 2.3 Analysestrategie

#### *Survivalanalyse*

Om een antwoord te geven op de vraag waarom studenten op een bepaald moment na aanvang hun eerst gekozen mbo-opleiding afbreken en anderen dat eerder, later of niet doen, gebruiken we een time survival analysis (Cox, 1972), ook wel *time hazard analysis*

of *event history analysis* genoemd.

In onze studie worden secundaire data-analyses uitgevoerd op data die op vier meetmomenten gedurende de eerste twee leerjaren verzameld zijn bij alle studenten van de acht niveau 2 of niveau 3/4 opleidingen. Omdat niet alle studenten hun niveau 2 opleiding binnen twee jaar weten af te ronden en de niveau 3/4 opleidingen langer duren dan de twee jaren dat de studenten gevolgd zijn, is er sprake van gecensureerde data: niet alle studenten die hun opleiding afbreken, zullen dat doen in de periode dat zij gevolgd zijn. Het is echter niet realistisch om te veronderstellen dat het ook in de volgende leerjaren niet zal gebeuren. De gekozen analysetechniek biedt de mogelijkheid hier rekening mee te houden en zo een vertekend beeld van de invloed van verklarende variabelen op het (moment van) uitvallen te voorkomen (Singer & Willet, 1993).

De variabele die informatie over het moment van uitval bevat is categorisch van aard, dientengevolge voeren we een discrete-time survival analysis uit. Omdat de uitkomst binair is (wel of niet uitval per interval tussen twee meetmomenten) schatten we een logistisch regressiemodel met daarin tijdsconstante en tijdsvariërende verklarende variabelen (Allison, 1982). Eerst wordt een model zonder voorspellers geschat. Daarna wordt het model aangevuld met de voorspellers. De discrete-time survival analysis is met behulp van het softwarepakket R, versie 3.2.3. uitgevoerd (R core Team, 2013). Daarbij is gebruik gemaakt van het Survival-package (Therneau, 2015).

#### *Meervoudige imputatie*

De hierboven uiteengezette survivalanalysetechniek verwijdert, net als veel andere statistische technieken, gehele meetmomenten van studenten wanneer er sprake is van missende waarden op één of meer van de variabelen die tijdens dat meetmoment gemeten zijn. Het missen van gegevens is deels inherent aan deze studie omdat studenten na het afbreken van een opleiding niet meer bij de opleiding stonden ingeschreven en dus niet meer aan metingen konden deelnemen. Zonder ingrijpen blijft een sample van 840 van de 2884

Tabel 3  
*Interne consistentie van diverse variabelen op meetmoment 1*

	Aantal items	Cronbach's $\alpha$
<b>Capaciteiten</b>		
Leerwinst	6	.700
<b>Tevredenheid/inzicht</b>		
Traject	8	.692
Sturing	5	.765
<b>Betrokkenheid</b>		
Motivatie	9	.791
Welbevinden	10	.793
Verwachtingen	6	.844

mogelijke meetmomenten (721 maal 4 meetmomenten) over en zal statistische analyse een vertekend beeld opleveren (Graham, 2009).

De afhankelijke variabele, die het moment van het afbreken van de eerst gekozen mbo-opleiding omvat, is compleet maar de meerderheid van de andere variabelen in de dataset bevat missende waarden. De percentages missende waarden per variabele variëren van 0 tot 63.9 procent. Variabelen zonder missende waarden zijn vooropleidingskenmerken en mbo-opleidingskenmerken. Ook weten we van alle studenten de sekse en wanneer hun kwalificatieplicht verviel. Van de variabelen etniciteit en opleidingsniveau van de ouders missen respectievelijk 29.1 en 29.7 procent. Gegevens over de spijbelfrequentie van studenten mist in 39.5 procent van de gevallen. De overige variabelen, motivatie, welbevinden, verwachtingen, leerwinst, traject en sturing, zijn op vier momenten gemeten. Aangezien een deel van de studenten op enig moment de opleiding verlaat, stijgt het percentage missende waarden gedurende de tijd. Op het eerste meetmoment missen de genoemde variabelen tussen 26.2 en 27.5 procent van de waarden. Op meetmoment twee varieert het percentage missende waarden van die variabelen tussen 49.1 en 51.9. Op meetmoment drie is de range 55.1 tot 56.0 procent en op meetmoment vier mist 62.6 tot 63.9 procent van de waarden.

Om dit probleem te verhelpen zijn de missende waarden in de dataset (in brede vorm) geïmputeerd met behulp van het MICE-pakage (multiple imputation through chained equations) in het softwarepakket R, versie 3.2.3. (Van Buuren & Groothuis-Oudshoorn, 2011). Daarbij is per variabele een geschikte imputatiemethode geselecteerd: logistische regressie (logreg) voor binaire variabelen, multinomiale logistische regressie (polyreg) voor categorische variabelen met meer dan twee categorieën en predictive mean matching (pmm) voor continue variabelen. Bij het imputeren van missende waarden is per te imputeren variabele de restrictie in acht genomen dat variabelen minimaal  $r = 0.30$  met de te imputeren variabelen samen moeten hangen om als voorspeller voor de missende waarden gebruikt te kunnen worden. Een aan-

vullende hulpvariabele is meegenomen bij het imputeren, namelijk de variabele 'opzet'. Deze variabele zelf is niet geïmputeerd omdat deze variabele geen onderdeel uitmaakt van het opgestelde model, maar is, mits aan de minimale correlatie is voldaan, meegenomen als voorspeller voor andere variabelen. Met deze variabele is in kaart gebracht of de opzet van de vakken en opleiding voor studenten duidelijk was. 'Opzet' is op alle vier meetmomenten gemeten door middel van zes items als "Ik weet precies wat me in deze opleiding het komend jaar te wachten staat" (cronbach's  $\alpha$  bij eerst meting = .765).

Meerdere factoren zijn van invloed op de keuze voor het optimale aantal geïmputeerde datasets (Bodner, 2008). Rekening houdend met de aanzienlijke *fraction of missing information* (gemiddeld 0.51) en de zeer redelijke steekproefomvang ( $n = 721$ ) zijn vijftientig geïmputeerde datasets geconstrueerd. Convergentie van het imputatiemodel kwam binnen vijftien iteraties tot stand. De verdelingen van scores van de variabelen in de vijftientig geïmputeerde datasets waren vergelijkbaar met de verdelingen binnen de oorspronkelijke dataset. Op ieder van deze datasets is een survivalanalyse uitgevoerd. De resultaten zijn, gebruikmakend van Little en Rubin's regels (zoals weergegeven in Snijders & Bosker, 2012), gepoold om tot één schatting van de modelparameters te komen. De meetonzekerheid binnen iedere geïmputeerde dataset (*within dataset variance*) is klein: maximaal 0.08. De onzekerheid tussen de geïmputeerde datasets (*between imutations variance*) is eveneens klein: maximaal 0.07. Dit wijst er op dat het imputeren van data geen grote inflatie van meetonzekerheid tot gevolg heeft gehad.

### 3 Resultaten

Allereerst is het aantal studenten dat gedurende de eerste twee leerjaren van het mbo hun opleiding verlaat in kaart gebracht. De gegevens zijn weergegeven in Tabel 4 en zijn afkomstig uit de survivalanalyse waarin nog geen voorspellende variabelen zijn meegenomen. In totaal beëindigt 42.2 procent ( $n = 304$ ) gedurende de eerste twee leerjaren hun

Tabel 4

Uitval gedurende de eerste twee leerjaren van het mbo

Meetmoment		Aantal ingeschreven studenten bij aanvang van tijdsinterval tot meting	Aantal uitgeschreven studenten in tijdsinterval tot meting	Hazard	Survival
Leerjaar 1	November	721	28	0.0388	0.9612
	Mei	693	66	0.0952	0.8660
Leerjaar 2	November	627	179	0.2855	0.5805
	April	448*	31	0.0692	0.5113

\* 417 studenten stonden na april in het tweede leerjaar nog ingeschreven

mbo-opleiding. De voorlaatste kolom in de tabel geeft voor ieder tijdsinterval tussen twee meetmomenten het risico op het afbreken van een opleiding weer voor studenten die aan het begin van het betreffende tijdsinterval nog bij hun opleiding ingeschreven stonden. In survivalanalyse wordt deze kans *hazard rate* genoemd. Uit de gegevens in de kolom blijkt bijvoorbeeld dat de kans op het afbreken van een opleiding tussen aanvang van de opleiding en het eerste meetmoment in november 3.9 procent bedraagt en dat de kans op het afbreken van de opleiding het grootst is onder studenten die bij aanvang van het tijdsinterval dat loopt van het laatste meetmoment in het eerste leerjaar tot het eerste meetmoment in het tweede leerjaar nog ingeschreven stonden, namelijk 28.6 procent. De laatste kolom van de tabel geeft de kans op het voortzetten van een opleiding tot een bepaald meetmoment (ook wel *survival rate*) weer. Uit de gegevens blijkt bijvoorbeeld dat de survivaalkans tot aan het meetmoment in november van het eerste leerjaar 96.1 procent bedraagt en dat deze kans tussen mei in het eerste leerjaar en november in het tweede leerjaar drastisch daalt, van 86.6 naar 58.1 procent. Dit houdt in dat het overgrote deel van de studenten die gedurende de eerste twee leerjaren hun opleiding afbreken, dit rond of vlak na de overgang naar het tweede leerjaar doet.

De resultaten uit het survivalmodel met voorspellers, die allen simultaan zijn toegevoegd, zijn weergegeven in Tabel 5. De derde kolom van de tabel geeft de schatting van de regressiecoëfficiënt voor iedere variabele. In

de kolom daarna wordt de hazard ratio weer gegeven. De hazard ratio is het exponent van de regressiecoëfficiënt. Waarden groter dan één geven een verhoogd risico op het afbreken van een opleiding aan. Een verlaagd risico gaat samen met een hazard ratio kleiner dan één.

De eerste groep variabelen in de tabel betreft student-/familiekenmerken. In de steekproef is het risico op het afbreken van de eerst gekozen mbo-opleiding door studenten waarvan de ouders hoger of universitair onderwijs hebben afgerond 0.58 keer zo groot als het risico op het afbreken van de opleiding door studenten waarvan de ouders maximaal basisonderwijs hebben afgerond (referentiecategorie). Het bijbehorende betrouwbaarheidsinterval (0.34; 0.99) wijst er op dat de groep studenten met hoger opgeleide ouders in de populatie waarschijnlijk een behoorlijk tot minimaal verlaagd risico loopt op het afbreken van de eerst gekozen mbo-opleiding in vergelijking met studenten met laagopgeleide ouders ( $b = -0.55$ ;  $p = .04$ ).

Er is onvoldoende bewijs voor een substantieel verschil in risico op het afbreken van de opleiding tussen studenten waarvan de ouders basis- en middelbaar onderwijs hebben afgerond en studenten waarvan de ouders maximaal basisonderwijs hebben afgerond.

Voor een relatie tussen het afbreken van de opleiding en de sekse of etniciteit van studenten is onvoldoende bewijs gevonden (sekse:  $b = -0.19$ ;  $p = .13$ ; etniciteit:  $b = 0.25$ ;  $p = .16$ ).

De capaciteiten van studenten zijn in beeld gebracht aan de hand van de perceptie

Tabel 5  
Resultaten survivalanalyse

		$\beta$ schatting (SE)	p – waarde*	Hazard ratio On- dergr.	95% bhi** Bovengr.	
<b>Student-/familiekenmerken</b>						
Sekse	Vrouw	-0.19 (0.13)	0.13	0.82	0.64	1.06
Etniciteit	Tenminste één in NL geboren ouder	0.25 (0.18)	0.16	1.28	0.91	1.82
Opleidingsniveau ouders	Basis- & middelbaar onderwijs	-0.53 (0.28)	0.06	0.59	0.34	1.02
	Hoger of universitair onderwijs	-0.55 (0.27)	0.04	0.58	0.34	0.99
<b>Capaciteiten</b>						
Niveau vooropleiding	kl	0.16 (0.15)	0.28	1.18	0.88	1.58
	gl	-0.49 (0.33)	0.13	0.61	0.32	1.16
	tl	-0.08 (0.16)	0.59	0.92	0.68	1.25
Sector vooropleiding	Landbouw	0.02 (0.23)	0.93	1.02	0.65	1.60
	Techniek	-0.13 (0.16)	0.41	0.87	0.64	1.20
	Zorg & Welzijn	0.27 (0.13)	0.03	1.31	1.02	1.69
Niveau mbo-opleiding	Niveau 3/4	-0.11 (0.15)	0.45	0.89	0.67	1.19
Sector mbo-opleiding	Zorg & Welzijn	-0.19 (0.15)	0.19	0.82	0.62	1.10
	Techniek	-0.49 (0.25)	0.05	0.61	0.37	1.00
Leerwinst		-0.36 (0.25)	0.15	0.69	0.42	1.14
<b>Tevredenheid/Inzicht</b>						
Traject		-0.12 (0.28)	0.68	0.89	0.51	1.55
Sturing		0.05 (0.15)	0.76	1.05	0.78	1.40
<b>Betrokkenheid</b>						
Spijbelen	Maandelijks	0.33 (0.16)	0.04	1.39	1.02	1.90
	Wekelijks	0.75 (0.16)	<0.01	2.13	1.55	2.92
Kwalificatieplicht	Wel verplicht	-0.29 (0.10)	<0.01	0.75	0.61	0.91
Motivatie		-0.21 (0.16)	0.21	0.81	0.59	1.12
Welbevinden		0.31 (0.22)	0.17	1.36	0.88	2.10
Verwachtingen		-0.18 (0.27)	0.52	0.84	0.49	1.40
Modelfit: R-kwadraat (maximaal mogelijk 0.955) = 0.198; Likelihood ratio = 639.3						

\* Verkregen op basis van de z-verdeling

\*\* Grenzen van het betrouwbaarheidsinterval (bhi): exponent van  $\beta$  schatting  $\pm 1.96 * SE_{\beta}$  schatting

van leerwinst van studenten en kenmerken van de vmbo-vooropleiding en de gekozen mbo-opleiding. Er zijn enkel aanwijzingen gevonden voor een lager risico op het afbreken van de mbo-opleiding door studenten die een vmbo-opleiding in de sector Zorg & Welzijn volgden ten opzichte van studenten in de vmbo-sector Economie. In de steekproef is het risico op het afbreken van een opleiding

van Zorg & Welzijn-studenten 1.31 keer zo groot als het risico dat studenten met een vooropleiding in de sector Economie (referentiecategorie) lopen ( $b = 0.27$ ;  $p = .03$ ). Het bijbehorende betrouwbaarheidsinterval (1.02; 1.69) suggereert een minimaal tot behoorlijk hoger risico op het afbreken van de eerst gekozen mbo-opleiding van Zorg & Welzijn-studenten ten opzichte van Economie-studen-

ten in de populatie. Voor de overige (voor) opleidingskenmerken en de variabele leerwinst is het gevonden bewijs niet sterk genoeg om een relatie met het afbreken van een opleiding op enig moment gedurende de eerste twee leerjaren te veronderstellen (zie tabel 5).

De motivatie, het welbevinden, de verwachtingen, het spijbelgedrag en de kwalificatieplicht van studenten zijn gebruikt om de betrokkenheid van studenten in kaart te brengen. Er is onvoldoende bewijs voor een relatie tussen de motivatie, het welbevinden of de verwachtingen en het afbreken van een mbo-opleiding gedurende de eerste twee leerjaren gevonden (motivatie:  $b = -0.21$ ;  $p = .21$ ; welbevinden:  $b = 0.31$ ;  $p = .17$ ; verwachtingen:  $b = -0.18$ ;  $p = .52$ ). Het spijbelgedrag van studenten lijkt daarentegen wel indicatief te zijn. In de steekproef is het risico op het afbreken van een opleiding van studenten die hebben aangegeven hooguit maandelijks een keer te spijbelen 1.39 keer zo groot als het risico van de referentiegroep, studenten die vrijwel nooit spijbelen ( $b = 0.33$ ;  $p = .04$ ). Het risico op het afbreken van de eerst gekozen mbo-opleiding van studenten die aangeven wekelijks een keer te spijbelen is in de steekproef 2.13 keer zo groot als het risico van studenten die vrijwel nooit spijbelen ( $b = 0.75$ ;  $p < .01$ ). De betrouwbaarheidsintervallen wijzen op een minimaal tot behoorlijk groter risico van maandelijks spijbelende studenten (bhi: 1.02; 1.90) en een gemiddeld tot veel groter risico van wekelijks spijbelende studenten (bhi: 1.55; 2.92) in vergelijking tot niet spijbelende studenten in de populatie. Ook het vervallen van de kwalificatieplicht op het moment dat de studenten achttien jaar oud worden lijkt te zijn gerelateerd aan een verhoogd risico op het afbreken van de opleiding. In de steekproef hebben studenten die kwalificatieplichtig zijn een 0.75 keer groter (dus een lager) risico op afbreken van de opleiding dan studenten die niet meer kwalificatieplichtig zijn. Het betrouwbaarheidsinterval (0.61; 0.91) wijst op een minimaal tot redelijk verlaagd risico in de populatie voor jongeren die nog kwalificatieplichtig zijn.

Voor een relatie tussen het afbreken van de eerst gekozen mbo-opleiding gedurende

de eerste twee leerjaren en de tevredenheid van studenten of de mate waarin zij inzicht in hun opleiding zeggen te hebben is niet voldoende bewijs gevonden (traject:  $b = -0.12$ ;  $p = .68$ ; sturing:  $b = 0.05$ ;  $p = .76$ ).

## 4 Conclusie en Discussie

Om te onderzoeken hoe het risico op het afbreken van de eerst gekozen mbo-opleiding zich ontwikkelt gedurende de eerste twee leerjaren in het mbo en welke factoren daarop van invloed zijn, is een model getoetst dat is samengesteld uit elementen van de theoretische modellen van Tinto (1975), Finn (1989) en Battin-Pearson et al. (2000). Op basis van gegevens van een steekproef van 721 studenten die rechtstreeks vanuit het vmbo naar het mbo zijn doorgestroomd, blijkt het afbreken van opleidingen rond of vlak na de overgang van het eerste naar het tweede leerjaar een hoogtepunt te kennen. Deze bevinding impliceert dat een groot deel van de studenten die uitvallen de opleiding wel degelijk een kans hebben gegeven, waardoor de processen van betrokkenheid en identificatie die leiden tot uitval zoals beschreven in de theoretische modellen van Finn (1989) en Tinto (1975) wel degelijk een rol zouden kunnen spelen.

Aan de hand van een survivalanalyse is in kaart gebracht welke student-/familiekenmerken en capaciteits-, betrokkenheids- en tevredenheidsindicatoren een rol lijken te spelen bij het al dan niet verlaten van de eerst gekozen mbo-opleiding. De kans op het voortzetten van de gekozen mbo-opleiding, lijkt hoger a) voor studenten met een hogere sociaaleconomische status, b) voor studenten die een vmbo-opleiding in de sector Economie hebben gevolgd (in vergelijking met studenten met een vooropleiding in de sector Zorg & Welzijn), c) voor studenten die kwalificatieplichtig zijn, en d) naarmate studenten minder spijbelen.

Van de sociaaldemografische kenmerken uit de modellen van Tinto (1975) en Battin-Pearson en anderen (2000) is alleen voor de sociaaleconomische status een relatie gevonden met het verlaten van de eerst gekozen mbo-opleiding. Deze resultaten zijn in lijn

met de resultaten uit eerder onderzoek (bijv., Alexander, Entwisle, & Kabbani, 2001; Rumberger & Lim, 2008), waarin eveneens aanwijzingen zijn gevonden dat studenten met een hogere economische status minder risico op uitval lopen. De bevindingen omtrent de variabele sekse sluiten daarentegen niet aan bij eerder onderzoek, waarin gevonden werd dat jongens een groter risico op het afbreken van een opleiding lopen (Stage, 1989; Battin-Pearson et al., 2000; Elffers, 2011; Theunissen, de Man, Verdonk, Bosma, & Feron, 2015). Mogelijk is er op bepaalde momenten gedurende de eerste twee leerjaren wel sprake van een voorspellende waarde van sekse, maar is dat effect zo tijdelijk dat het over langere tijd niet meer waarneembaar is. Dit vermoeden wordt gesterkt door onderzoek van Elffers (2011) waaruit bleek dat sekse een rol speelt in het eerste half jaar na de overgang naar het mbo, maar dat effect werd niet meer gevonden in het tweede semester of rond de overgang naar het tweede leerjaar. De reden voor onverwachte afwezigheid van bewijs voor een relatie tussen het afbreken van een opleiding en de etniciteit van studenten is mogelijk te vinden in de manier waarop etniciteit gemeten is: aan de hand van het geboorteland van de ouders. Andere factoren die aan etniciteit gelieerd zijn, zoals het geboorteland van de student, verblijfstijd in Nederland, en de mate waarin ouders de Nederlandse taal machtig zijn, spelen mogelijk een rol (Herweijer, 2008).

Prestaties van studenten komen voor in ieder van de drie theoretische modellen op basis waarvan het in deze studie getoetste model is samengesteld, zij het op verschillende manieren. In de modellen van Tinto (1975) en Battin-Pearson en anderen (2000) stonden prestaties in de vooropleiding centraal. In lijn met deze modellen zien we dat de door de studenten gevolgde vooropleiding gerelateerd is aan de uitval in het mbo. In het bijzonder speelt de sector van de vooropleiding een rol. In het model van Finn (1989) als wel in het model van Tinto (1975) wordt verondersteld dat ook de huidige prestaties van studenten een rol spelen in het proces dat uiteindelijk tot uitval zou kunnen leiden. Omdat in de huidige studie werd aangenomen dat het

effect van prestaties op het afbreken van opleidingen sterk afhankelijk is van de mate waarin studenten waarde hechten aan het behalen van goede prestaties (bijv., McKeown, Macdonell, & Bowman, 1993), zijn op verschillende momenten in de eerste twee leerjaren de prestaties gemeten aan de hand van de perceptie die studenten van hun leerwinst hebben. Voor een relatie tussen deze perceptie en het afbreken van de eerst gekozen opleiding is onvoldoende bewijs gevonden. Deze bevinding suggereert dat lage cijfers mogelijk doorslaggevend zijn in het afbreken van een opleiding dan ontevredenheid van de student over de behaalde leerwinst. Ook is het mogelijk dat percepties wel een rol spelen bij studenten die zelf het initiatief nemen om hun opleiding af te breken, maar dat dit in mindere mate het geval is voor studenten die door de onderwijsinstelling of hun ouders worden aangeraden te stoppen.

Alle modellen waarop het samengestelde model is gebaseerd bevatten factoren met betrekking tot integratie en betrokkenheid. Naast extern opgelegde betrokkenheid in de vorm van kwalificatieplicht is intrinsieke betrokkenheid in deze studie gemeten aan de hand van spijbelgedrag, tevredenheid, motivatie, welbevinden en verwachtingen van studenten. De gevonden aanwijzingen voor de minimale tot zeer sterke negatieve relatie tussen spijbelen en het afbreken van opleidingen komen overeen met de bevindingen van onder meer Wehlage en Rutter (1986), Garry (1996) en Henry (2007). Daarentegen komen de bevindingen met betrekking tot motivatie, welbevinden en verwachtingen van studenten niet overeen met wat doorgaans wordt gevonden, namelijk dat de betrokkenheid van studenten een sleutelrol vervult in het al dan niet voortzetten van een opleiding (bijv., Fredricks, Blumenfeld, & Paris, 2004; Rumberger & Lim, 2008). Mogelijk gaat het effect van betrokkenheid in de huidige studie op in het veronderstelde negatieve effect van spijbelgedrag. Daarentegen werd wel gevonden dat een verhoogd risico op het afbreken van een opleiding samenhangt met het vervallen van de kwalificatieplicht wanneer jongeren achttien jaar worden. Dat zowel spijbelgedrag en kwalificatieplicht gerelateerd zijn aan



het uitvallen van leerlingen uit de eerst gekozen mbo-opleiding suggereert dat zowel extrinsieke als intrinsieke factoren een rol spelen bij de keuze van het voortzetten dan wel afbreken van een opleiding.

Voor verschillende kenmerken die in het samengestelde model zijn opgenomen is geen evidentie is gevonden in de huidige studie. Voor specifieke variabelen zijn hiervoor reeds potentiële verklaringen gegeven. Daarnaast geldt dat de oorspronkelijke modellen destijds zijn ontwikkeld voor een andere context, namelijk high school en higher education in de Verenigde Staten, welke verschilt van de context van het Nederlandse beroepsonderwijs. Ook hierin zou een mogelijke verklaring kunnen liggen. Deels gaat het echter ook om factoren die in de Nederlandse context eerder zijn onderzocht, zoals betrokkenheid (Elffers, 2011) en verschillende kenmerken uit de sociaal-demografische achtergrond (Luyten et al., 2003; Stevens et al., 2011). Wat deze kenmerken betreft kunnen verschillen in resultaten worden verklaard doordat in het huidige onderzoek is getracht een model te schatten waarin in vele variabelen simultaan zijn geanalyseerd. Zo zou het kunnen zijn dat eventuele kleine effecten van sociaal-demografische achtergrond, de perceptie van leerwinst, en intrinsieke betrokkenheid zijn tenietgedaan worden door de sterkere effecten van spijbelgedrag en kwalificatieplicht.

## 5 Beperkingen en Aanbevelingen

De longitudinale aanpak van het bestuderen van uitval die in deze studie gebruikt is, heeft als grootste voordeel dat het proces tot uitval, en dus niet alleen uitval zelf, in kaart kan worden gebracht. Deze aanpak maakt het bijvoorbeeld mogelijk om veranderingen in gedragingen van studenten aan hun uiteindelijke keuze tot het al dan niet afbreken van hun opleiding te relateren of het effect van invloedrijke gebeurtenissen in de familiecontext (bijv. een scheiding) te bestuderen. In de huidige studie is voornamelijk aandacht besteed aan factoren die in de modellen van Tinto, Finn en Battin-Pearson voorkomen, welke vrijwel allemaal meer gradueel veran-

derende factoren zijn (bijv. welbevinden of motivatie). Informatie over bijvoorbeeld een veranderende gezinssamenstelling hadden wij niet ter beschikking en gezien de steekproefgrootte had het includeren van dit soort factoren waarschijnlijk niet tot robuuste bevindingen geleid, maar het is aannemelijk dat dit soort gebeurtenissen wel degelijk van belang zijn. Voor vervolgonderzoek bevelen wij daarom aan om niet enkel gradueel veranderende factoren mee te nemen, maar om ook abrupt veranderende factoren zoals bijvoorbeeld slecht-nieuws gesprekken met een docent of een zwangerschap van de student in ogenschouw te nemen. Een nadeel van de longitudinale aanpak is dat het aantal missende waarden naarmate de tijd vordert steeds groter wordt, want studenten die een opleiding verlaten, missen alle nog ophanden zijnde meetmomenten. Dat resulteert in additionele meetonzekerheid bij het imputeren van missende waarden. Dit is voornamelijk bij de laatste meetmomenten het geval, alhoewel ook gemeld mag worden dat de additionele meetonzekerheid in deze studie niet bijzonder groot lijkt te zijn.

Ook de gekozen invalshoek is mogelijk van invloed op de conclusies die kunnen worden getrokken. We hebben ons gericht op het afbreken van de eerst gekozen mbo-opleiding. Hoewel het afbreken van de eerst gekozen opleiding mogelijk indicatief is voor een verkeerde studiekeuze, betekent het niet dat studenten naderhand geen opleiding op een hoger, lager, dan wel gelijk niveau volgen en afronden. In de huidige studie beschikten wij niet over informatie om vervolgkeuzes goed in kaart te brengen. Desondanks is het zinvol hier in vervolgonderzoek aandacht aan te besteden, omdat bijvoorbeeld spijbelgedrag en motivatie voor uitvallers mogelijk een andere rol spelen bij hun keuze tot uitval dan voor studenten die kiezen voor het veranderen van opleiding (switchers) het geval is. Het is bijvoorbeeld mogelijk dat een student die overweegt om van opleiding te veranderen wel gemotiveerd is om te leren en ook relatief weinig spijbelt, maar gewoonweg het gevoel heeft binnen een andere opleiding meer of relevantere dingen te kunnen leren. In dat geval is afwezigheid van samenhang of zelfs

een negatieve samenhang tussen spijbelfrequentie en de keuze om te stoppen met de huidige opleiding denkbaar en hangt motivatie mogelijk positief samen met het afbreken van de opleiding. Daarentegen is een student die overweegt zijn opleidingstraject af te breken waarschijnlijk minder gemotiveerd voor school en is daarom mogelijk ook minder aanwezig. In dat geval is het waarschijnlijker dat het risico op het afbreken van de huidige opleiding positief samenhangt met de spijbelfrequentie en negatief met motivatie. We verwachten echter geen verschil tussen de twee groepen studenten wat betreft hun perceptie van leerwinst, omdat het aannemelijk is dat zowel bij uitval als bij switchen het niveau en de inhoud van de opleiding negatief samenhangen met het komen tot de gemaakte keuze.

Een andere focus voor vervolgonderzoek is de reden voor het afbreken van een opleiding. In deze studie is daar, evenals in de theoretische modellen van Tinto, Finn en Battin-Pearson en collega's, geen rekening mee gehouden en ook weten we niet of studenten hun opleiding op eigen initiatief of op aandringen van de onderwijsinstelling hebben afgebroken. Het strekt tot aanbeveling om dat in vervolgonderzoek aan bod te laten komen, want het is bijvoorbeeld mogelijk dat de motivatie van studenten wel een rol speelt voor studenten die op eigen initiatief hun opleiding afbreken, maar dat dit bij studenten die door de onderwijsinstelling worden verzocht de opleiding te verlaten in mindere mate het geval is en dat het eventuele effect in de huidige studie verdwijnt doordat hier geen onderscheid naar gemaakt is (Brunsdon, Davies, Shevlin, & Bracken, 2000).

De data waarvan in deze studie gebruik is gemaakt, zijn verzameld bij een ROC in het Noorden van Nederland. De studentenpopulatie in het Noorden van het land wijkt bijvoorbeeld qua etniciteit (relatief veel studenten met tenminste 1 in Nederland geboren ouder) en de gekozen sector in de vooropleiding (meer jongeren in de landbouwsector) af van de studentenpopulatie in andere delen van Nederland. Mogelijk zijn de gevonden effecten van student- en familiekenmerken dus niet volledig representatief voor Nederlandse mbo-studenten. Er is echter geen

reden om aan te nemen dat studenten in het Noorden van het land ook afwijkende psychologische processen doormaken. Het is dus redelijk om aan te nemen dat de bevindingen omtrent spijbelgedrag en kwalificatieplicht representatief zijn.

Van alle variabelen die in deze studie zijn meegenomen, bleken spijbelgedrag en kwalificatieplicht de grootste voorspellende waarde te hebben voor het voortzetten of afbreken van de eerst gekozen mbo-opleiding gedurende de eerste twee leerjaren in het mbo. Het gevonden effect van spijbelgedrag is niet uniek, maar voor zover bij ons bekend is dit de eerste studie die kwalificatieplicht als mogelijke voorspeller heeft meegenomen. De uitkomsten van deze studie doen vermoeden dat leer- of kwalificatieplicht onder jongeren in het mbo wel degelijk van belang is. Wij raden derhalve aan om leer-/kwalificatieplicht in toekomstig onderzoek naar uitval mee te nemen en de werking van dit beleidsinstrument nader te onderzoeken.

De resultaten omtrent spijbelgedrag en kwalificatieplicht bevestigen dat het aanpakken van spijbelen in een poging tot het terugdringen van uitval, zoals recentelijk in het Nederlandse onderwijs gedaan wordt, een zinvolle actie is. Daarnaast bevestigen deze resultaten dat de geldende kwalificatieplicht bevorderlijk is. Veel jongeren breken hun opleiding pas af op het moment of nadat de kwalificatieplicht vervallen is. Door professionals in de onderwijspraktijk mogen de uitkomsten van dit onderzoek worden gezien als extra motivatie om studenten op hun spijbelgedrag aan te spreken. Voor beleidsmakers bevestigt dit onderzoek de huidige regels rondom leerplicht.

Meerdere factoren waarvoor in eerder onderzoek evidentie werd gevonden, lijken er op basis van het huidige onderzoek niet toe te doen. Voorbeelden daarvan zijn motivatie, tevredenheid en welbevinden van studenten. Enerzijds is dit voor de onderwijspraktijk een geruststelling, want deze resultaten lijken er op te wijzen dat studenten hun opleiding niet afbreken op basis van een gebrek aan motivatie, tevredenheid en welbevinden. Anderzijds zijn deze resultaten zorgwekkend, want een hoge motivatie, een hoge mate van tevreden-

heid en een hogere mate van welbevinden lijken ook geen beschermende factoren. Het lijkt er dus op dat studenten zich in hun keuze tot het voortzetten van hun opleiding en dat studenten slechts in beperkte mate laten leiden door intrinsieke factoren. Op dat vlak lijkt in de onderwijspraktijk dus nog wel winst te behalen.

## Literatuur

- Alexander, K. L., Entwisle, D. R., & Kabbani, N. S. (2001). The dropout process in life course perspective: early risk factors at home and school. *Teachers College Record*, 103, 760-822. doi: 10.1097/00004583-199007000-00014
- Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. In Leinhardt, S. (Ed.), *Sociological Methodology*. San Francisco: Jossey-Bass. doi: 10.2307/270718
- Anderson, D. K. (1993). *Effects of pregnancy, childbirth, and motherhood on high school dropout* (Discussion Paper No. 1027-93). Madison, WI: Madison Institute for Research on Poverty.
- Archambault, I., Janosz, M., Fallu, J. S., & Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence*, 32, 651-670. doi: 10.1016/j.adolescence.2008.06.007
- Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of Educational Psychology*, 92, 568-582. doi: 10.1037/0022-0663.92.3.568
- Bers, T. H., & Smith, K. E. (1991). Persistence of community college students: the influence of student intent and academic and social integration. *Research in Higher Education*, 32, 539-556.
- Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural Equation Modeling: A Multidisciplinary Journal*, 15, 651-675. doi: 10.1080/10705510802339072
- Brunsdon, V., Davies, M., Shevlin, M., & Bracken, M. (2000). Why Do HE Students Drop Out? A Test of Tinto's Model. *Journal of Further and Higher Education*, 24, 301-310. doi: 10.1080/030987700750022244
- Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society, Series B*, 34, 187-202.
- De Witte, K., & Csillag, M., 2014. Does anybody notice? On the impact of improved truancy reporting on school dropout. *Education Economics*, 22, 549-568. doi: 10.1080/09645292.2012.672555
- Driessen, G., van Langen, A., & Vierke, H. (2006). *Basisonderwijs: Veldwerkverslag, leerlinggegevens en oudervragenlijsten. Basisrapportage PRIMA-cohortonderzoek. Zesde meting 2004/05*. Nijmegen, the Netherlands: ITS.
- Dupin-Bryant, P. A. (2004). Pre-entry variables related to retention in online distance education. *The American Journal of Distance Education*, 18, 199-206. doi: 10.1207/s15389286ajde1804\_2
- Elffers, L. (2011). The transition to post-secondary vocational education: students' entrance, experiences, and attainment. Amsterdam: Universiteit van Amsterdam.
- Ensminger, M.E., & Slusarcick, A.L. (1992). Paths to high school graduation or dropout: A longitudinal study of a first-grade cohort. *Sociology of Education*, 65(2), 95-113. doi: 10.2307/2112677
- Ethington, C. A. (1990). A psychological model of student persistence. *Research in Higher Education*, 31, 279-293. doi: 10.1007/BF00992313
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, 59, 117-142.
- Finn, J. D. (1993). *School engagement and students at risk*. Washington, DC: National Center for Education Statistics. doi: 10.3102/00346543059002117
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59-109. doi: 10.3102/00346543074001059
- Freudenberg, N. & Ruglis, J. (2007). Reframing school dropout as a public health issue. *Prevention Chronic Disease*, 4(4), 1-11.
- Garry, E. M. (1996). *Truancy: First step to a lifetime of problems*. Washington, DC: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.
- Graham, J. A. (2009). Missing data analysis: Making it in the real world. *Annual Review of Psychology*, 60, 549-533. doi: 10.1146/annurev.psych.58.110405.085530

- Hausman, L., Schofield, J. & Woods, R. (2007). Sense of belonging as a predictor of intentions to persist among african american and white first-year college students. *Research in Higher Education*, 48(7), 803-839. doi: 10.1007/s11162-007-9052-9
- Henry, K. L. (2007). Who's skipping school: characteristics of truants in 8th and 10th grade. *Journal of School Health* 77, 29-35. doi: 10.1111/j.1746-1561.2007.00159.x
- Herweijer, L. (2008). *Gestruikeld voor de start: de school verlaten zonder startkwalificatie*. Den Haag: Sociaal Cultureel Planbureau.
- Hirschfield, P. (2009). Another way out: The impact of juvenile arrests on high school dropout. *Sociology of Education*, 82, 368-393. doi: 10.1177/003804070908200404
- Janosz, M., Archambault, I., Morizot, J., & Pagani, L. S. (2008). School engagement trajectories and their differential predictive relations to dropout. *Journal of Social Issues*, 64, 21-40. . doi: 10.1111/j.1540-4560.2008.00546.x
- Johnson, W., McGue, M., & Iacono, W. G. (2006). Genetic and environmental influences on academic achievement trajectories during adolescence. *Developmental Psychology*, 42, 514-532. . doi: 10.1037/0012-1649.42.3.514
- Karszen, M., Van der Veen, I. & Volman, M. (2015). Educational outcomes and functioning of bi-ethnic Dutch children in school. *Educational research*, 57, 354 - 367. doi: 10.1080/00131881.2015.1076994
- Kuyper, H., & Van der Werf, M.P.C. (2003). *VOCL'99-1: Technisch rapport*. Groningen, The Netherlands: GION.
- Liu, R., & Liu, E. (2000). *Institutional integration: An analysis of Tinto's theory*. Paper presented at the Annual Forum of the Association of Institutional Research, Cincinnati, OH.
- Luyten, H., Bosker, R.J., Dekkers, H., & Derks, A. (2003). Dropout in the lower tracks of Dutch secondary education: predictor variables and variation among school. *School Effectiveness and School Improvement*, 14, 373-411. doi: 10.1076/sesi.14.4.373.17158
- McKeown, B., Macdonell, A., & Bowman, C. (1993). The point of view of the student in attrition research. *The Canadian Journal of Higher Education*, 23, 65-85.
- Ministerie van Onderwijs, Cultuur en Wetenschap. (2002). *Brochure: Iedereen een startkwalificatie*. Den Haag: auteur.
- Ministerie van Onderwijs, Cultuur en Wetenschap. (2016). *Nieuwe voortijdig schoolverlaters. Convenantjaar 2014-2015. Voorlopige cijfers*. Den Haag: Auteur.
- National Center for Education Statistics (z.j.). *Table 5.1. Compulsory school attendance laws, minimum and maximum age limits for required free education, by state: 2015*. Retrieved from: [https://nces.ed.gov/programs/statereform/tab5\\_1.asp](https://nces.ed.gov/programs/statereform/tab5_1.asp)
- O'Keeffe, P. (2013). A sense of belonging: Improving student retention. *College Student Journal*, 47, 605-613.
- Pascarella, E. T., & Chapman, D. W. (1983). A multi-institutional, path analytic validation of Tinto's model of college withdrawal. *American Educational Research Journal*, 20, 87-102. doi: 10.3102/00028312020001087
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, <http://www.R-project.org/>.
- Rumberger, R. W. (2011). *Dropping out: Why students drop out of high school and what can be done about it*. Cambridge, MA: Harvard University Press.
- Rumberger, R. W., & Lim, S. A. (2008). *Why students drop out of school: a review of 25 years of research*. Santa Barbara: California Dropout Research Project.
- Singer, J. D., & J. B. Willett. (1993). It's about time: using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics*, 18, 155-195. doi: 10.2307/1165085
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multi-level analysis: An introduction to basis and advanced multilevel modeling*. Los Angeles, CA: Sage.
- Stage, F.K. (1989). Motivation, academic and social integration, and the early dropout. *American Educational Research Journal*, 26, 385-402. doi: 10.3102/00028312026003385
- Stevens, P.A.J, Clycq, N., Timmerman, C., & Van Houtte, M. (2011). Researching race/ethnicity and educational inequality in the Netherlands: A critical review of the research

literature between 1980 and 2008. *British Educational Research Journal*, 37, 5-43. doi: 10.1080/01411920903342053

Sweeten, G. (2006). Who will graduate? Disruption of high school education by arrest and court involvement. *Justice Quarterly*, 23, 462-480. doi: 10.1080/07418820600985313

Terenzini, P. T. & Pascarella, E. T. (1977). Voluntary freshman attrition and patterns of social and academic integration in a university: a test of a conceptual model. *Research in Higher Education*, 6, 25-43. doi: 10.1007/BF00992014

Ter Bogt, T., Van Lieshout, M., Doornwaard, S., & Eijkemans, Y. (2009). *Middelengebruik en voortijdig schoolverlaten*. Utrecht: Trimbos Instituut.

Therneau, T. M. (2015). *A Package for Survival Analysis in S. Versie 2.39-4*. URL: <http://CRAN.R-project.org/package=survival>.

Theunissen, M.-J., De Man, I., Verdonk, P., Bosma, H. & Feron, F. (2015). Are Barbie and Ken too cool for school? A case-control study on the relation between gender and dropout. *European Journal of Public Health*, 25, 57-62. doi: 10.1093/eurpub/cku097

Thomas, S. I. (2000). Ties that bind: A social network approach to understanding student integration and persistence. *The Journal of Higher Education*, 71, 591-615. doi: 10.2307/2649261

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45, 89-125. doi: 10.3102/00346543045001089

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1-67. doi: 10.18637/jss.v045.i03

Vugteveen, J., Timmermans, A. C., Korpershoek, H., van Rooijen, M., & Opendakker, M.-C. (2016). Overgangen en aansluitingen in het onderwijs: Deelrapportage 3: empirische studie naar de cognitieve en niet-cognitieve ontwikkeling van leerlingen rondom de vmbo-mbo overgang. Groningen: GION onderzoek/onderwijs.

Wehlage, G. G., & Rutter, R. A. (1986). Dropping out: How much do schools contribute to the problem? *Teachers College Record*, 87, 374-392.

Zijlsing, D. H., Keuning, J., Keizer-Mittelhaeuser, M.-A., Naaijer, H. M., & Timmermans, A. C. (2017). *Cohortonderzoek COOL5-18 Technisch rapport meting VO-3 in 2014*. GION onderwijs/onderzoek.

## Auteurs

**Jorien Vugteveen** werkt als onderzoeker bij het Heymansinstituut voor Psychologisch Onderzoek, Rijksuniversiteit Groningen. **Anneke C. Timmermans** werkt als universitair docent bij GION Onderwijs/Onderzoek, Rijksuniversiteit Groningen.

**Correspondentieadres:** Jorien Vugteveen, Grote Kruisstraat 2/1, 9712 TS Groningen. Email: [j.vugteveen@rug.nl](mailto:j.vugteveen@rug.nl)

## Abstract

### The risk at drop-out during the first two years of vocational education

Approximately ninety percent of students move on to vocational education after completing pre-vocational education. In vocational education, however, thousands of students drop out each year. Over the past five decades several theoretical models to describe drop out were suggested (bijv., Tinto, 1975; Finn, 1984; Battin-Pearson et al., 2000). A number of 721 students were followed during the first two years in Dutch vocational education in order to investigate the effect of factors from these theoretical models at the students' drop out risk. Forty-two percent of the students ( $n = 304$ ) discontinued their studies at one point during this time. A survival analysis brought to light that dropout peaks around or shortly after the transition from the first to the second year of vocational education. Higher educational levels in parents, pre-vocational education in an economy track (compared to a health care track), a lower truancy frequency, and compulsory school attendance seems to be associated with reduced risk at dropping out. Apart from being obligated by law to go to school, not enough evidence was found for the potential role of other time varying factors (e.g., motivation, expectations and the degree to which student have control over their own learning trajectory). This finding might be caused by the use of operationalizations that deviate from the original ones.

**Keywords:** dropout, vocational education, risk factors