

*L. Hornstra, L. van den Bergh  
en E. Denessen*

#### Samenvatting

In onderzoek naar de effecten van groepsstereotiepe houdingen die leraren van leerlingen hebben, wordt doorgaans gebruik gemaakt van zelfrapportages, veelal in de vorm van vragenlijsten, om deze houdingen te bepalen. Bij het gebruik van zelfrapportages bestaat het gevaar dat sociaal wenselijke antwoorden worden gegeven. Impliciete methoden uit de sociale psychologie zoals de impliciete associatie test en priming kunnen dan uitkomst bieden. We illustreren het gebruik van deze methoden aan de hand van eigen onderzoek naar de effecten van groepsstereotiepe houdingen van leraren op de prestaties van leerlingen. Dit laat zien dat met impliciete metingen van vooroordelen – in tegenstelling tot vragenlijsten – een relatie kan worden vastgesteld met prestaties van leerlingen uit groepen waarover vooroordelen bestaan.

#### 1 Inleiding

In vrijwel elke sociaal wetenschappelijke discipline wordt regelmatig onderzoek gedaan naar houdingen, ook wel aangeduid als attitudes, bijvoorbeeld ten opzichte van personen, groepen of andere fenomenen. Met name in de sociale psychologie wordt veel onderzoek gedaan naar houdingen als determinanten of voorspellers van gedrag. Hierbij wordt een houding opgevat als een onderliggende dispositie die evaluatief van aard is. Het gaat dan om samenvattende evaluatieve oordelen die uit te drukken zijn in dimensies als goed-slecht, fijn-vervelend, of aardig-onaardig (Ajzen, 2001; 2005). Houdingen kunnen echter meerdere dimensies omvatten. Deels kunnen houdingen automatisch geactiveerd worden en leiden tot spontane gedragingen, maar houdingen en het daaruit

voortvloeiende gedrag kunnen ook deels gebaseerd zijn op bewuste overwegingen (Gawronski & Bodenhausen, 2006). Door de jaren heen zijn er binnen de sociale psychologie steeds betere technieken ontwikkeld om houdingen te kunnen meten. Vooral de ontwikkeling van zogeheten ‘impliciete metingen’ – die houdingen trachten te meten die respondenten mogelijk niet willen of kunnen rapporteren – kan gezien worden als een doorbraak op dit gebied (Wittenbrink & Schwartz, 2007).

Ook in de onderwijskunde spelen houdingen mogelijk een belangrijke rol. Zo wordt binnen de onderwijskunde al enkele decennia onderzoek gedaan naar het effect van verwachtingen van leraren op de academische prestaties van leerlingen. Vooral van specifieke groepen leerlingen waarover groepsstereotyperingen kunnen bestaan, zoals allochtone leerlingen, blijken leraren relatief lage verwachtingen te hebben (Jussim & Harber, 2005). In de literatuur worden deze differentieële verwachtingen ten aanzien van verschillende groepen leerlingen vaak toegeschreven aan stereotiepe houdingen van leraren. Zo stellen McKown en Weinstein (2008) dat verwachtingen deels bepaald lijken te worden door de etniciteit van leerlingen, waardoor leerkrachten hogere verwachtingen hebben van autochtone leerlingen dan van leerlingen die behoren tot etnische minderheidsgroeperingen. Ook Jussim, Eccles en Madon (1996) veronderstellen dat differentieële verwachtingen mogelijk veroorzaakt worden door stereotiepe houdingen van leraren ten aanzien van groepskenmerken van leerlingen, zoals hun etnische achtergrond, gender of sociaal-economische status. Er is echter nog weinig eenduidig bewijs voor het bestaan van negatieve effecten van groepsstereotiepe houdingen van leraren op de verwachtingen die leraren van hun leerlingen hebben.

Effecten van leraarverwachtingen op de prestaties zouden tot stand komen via het gedrag van de leraar (Rosenthal, 1994). Op basis van hun verwachtingen kunnen leraren

leerlingen van wie zij verschillende verwachtingen hebben namelijk ook verschillend benaderen. Verwachtingen uit zich met name in non-verbaal leerkrachtgedrag. Zo blijkt dat leraren bijvoorbeeld meer warmte uitstralen naar een leerling van wie zij hoge verwachtingen hebben, of wachten ze minder lang op een antwoord bij een leerling van wie zij minder hoge verwachtingen hebben (Rosenthal, 1994; 2003). Voor leerlingen blijkt de houding van de leerkracht vooral uit dit non-verbale leerkrachtgedrag (Babad, Bernieri, & Rosenthal, 1991, Van Tartwijk, 1993). Een potentieel gevolg van differentieel leraar-gedrag is dat leerlingen vervolgens gaan presteren op een wijze die overeenstemt met de verwachting die de leraar van hen heeft (Jussim & Harber, 2005). Verwachtingen zouden door het gedrag van de leerkracht dus leiden tot corresponderende prestaties. Hoewel uit de reviewstudie van Jussim en Harber (2005) is gebleken dat lagere verwachtingen die leraren van bepaalde groepen leerlingen hebben in veel gevallen zijn gerechtvaardigd vanwege de eerdere prestaties van leerlingen uit deze groepen, blijft het de vraag in hoeverre verwachtingen die leraren van leerlingen hebben ook deels het gevolg zijn van hun houdingen tegenover bepaalde groepen leerlingen.

In onderwijskundig onderzoek wordt nog nauwelijks gebruik gemaakt van de eerder geschetste ontwikkelingen in de sociale psychologie op het gebied van instrumentatie om dergelijke groepsstereotiepe houdingen te meten. In dit artikel zullen we het gebruik van impliciete metingen in onderwijskundig onderzoek illustreren.

## 2 Expliciete en impliciete metingen van houdingen

### 2.1 Het gebruik van zelfrapportages

Groepsstereotiepe houdingen zijn, evenals veel persoonlijkheidseigenschappen, latente en hypothetische constructen die niet rechtstreeks zijn vast te stellen. Veruit de meest gebruikte methode om houdingen te meten is zelfrapportage, vaak door middel van vragenlijsten. Hierbij wordt een respondent gevraagd om expliciet op een *rating scale* te be-

oordelen in hoeverre hij of zij positief, dan wel negatief over een bepaalde groep denkt (Wittenbrink & Schwarz, 2007). Ook binnen de onderwijskunde is zelfrapportage een veelgebruikte methode om houdingen vast te stellen (Cohen, Manion, & Morrison, 2007). Bij zelfrapportage bestaat echter altijd het risico dat een respondent geen valide antwoorden geeft. Een vraag kan bijvoorbeeld te moeilijk zijn, de respondent probeert zichzelf 'rooskleuriger' voor te doen dan daadwerkelijk het geval is, of het bevroegde betreft iets waarvan de respondent zich mogelijk niet volledig bewust is (Ajzen, 2005; Jobe, 2000; Quillian, 2008). Niet alleen bij het gebruik van vragenlijsten, ook bij andere vormen van zelfrapportage, zoals interviews, kunnen deze problemen optreden (Quillian, 2008).

Het gebruik van zelfrapportage blijkt dan ook niet altijd tot valide uitkomsten te leiden. Het antwoorden op een vragenlijst vereist dat de respondent de vraag begrijpt, bepaalt welke informatie van hem of haar gevraagd wordt, de informatie ophaalt uit het geheugen, een beslissing neemt over de accuraatheid van deze informatie en een antwoord formuleert op de wijze die gevraagd wordt (Jobe, 2000; Schwarz, 2008). Meetfouten kunnen bij elk van deze stappen optreden. Bij onderzoek met vragenlijsten is er uiteraard altijd in zekere mate van meetfouten sprake, waarvoor gecorrigeerd kan worden (correctie voor attenuatie). Echter, wanneer het systematische meetfouten betreft – bijvoorbeeld als gevolg van sociaal wenselijk antwoordgedrag – kan dit de validiteit van het meetinstrument aantasten.

Dat vragenlijstscores met betrekking tot groepsstereotiepe houdingen een relatief geringe predictieve waarde hebben, kan onder andere veroorzaakt worden door strategisch antwoordgedrag van de respondent (Jobe, 2000). Hiervan blijkt vooral sprake te zijn wanneer het gaat om controversiële en sociaal gevoelige onderwerpen (Gawronski, LeBel, & Peters, 2007; Schwarz & Oyserman, 2001). Centraal daarbij staat het concept 'sociale wenselijkheid', dat omschreven kan worden als de neiging van respondenten om zodanig te antwoorden dat ze zichzelf zo positief mogelijk doen overkomen voor anderen (Mitchell, Patel, & MacDorman, 2010).

Sociaal wenselijk gedrag treedt op wanneer de onderzoeker vragen stelt die door de respondent als bedreigend ervaren worden. Dit kan het geval zijn wanneer het 'ware' antwoord van de respondent ingaat tegen algemeen geaccepteerde waarden. Hierdoor kunnen gedragingen of houdingen die sociaal afgekeurd worden, ondergerapporteerd worden, terwijl gedragingen of houdingen die als wenselijk beschouwd worden, benadrukt of overdreven worden (Schaeffer, 2000). Naast sociale wenselijkheid, vereist zelfrapportage dat de respondent zich volledig bewust is van datgene dat bevraagd wordt. Groepsstereotiepe houdingen blijken echter vaak deels onbewust te zijn en hierdoor moeilijk te doorgronden (Quillian, 2008). Ook dit kan de geringe predictieve waarde van zelfrapportages van groepsstereotiepe houdingen verklaren.

Volgens het MODE model van Fazio en Towles-Schwen (1999) hangt de mate waarin een houding gedrag kan voorspellen af van de mate waarin een persoon de motivatie en mogelijkheden heeft om zijn reacties bewust te overwegen. De afkorting MODE staat dan ook voor 'Motivation and Opportunity as Determinants of behaviour'. Wanneer een domein meer gevoelig of controversieel is, dan is het waarschijnlijker dat de persoon zijn gedrag zal controleren, mits diegene daar ook de mogelijkheid toe heeft. Dit model is ook van toepassing op het antwoordgedrag bij expliciete metingen van houdingen. Groepsstereotiepe houdingen zullen vaak zeer gevoelig liggen, waardoor respondenten een hoge motivatie zullen hebben om een mogelijk negatieve houding niet te rapporteren (Wittenbrink & Schwarz, 2007).

## **2.1 De ontwikkeling van impliciete metingen**

Bovengenoemde problemen zijn aanleidingen geweest voor de ontwikkeling van alternatieve methoden om houdingen in kaart te brengen (Schaeffer, 2000). In de sociale psychologie zijn zogeheten 'impliciete metingen' ontwikkeld. Deze impliciete metingen zijn de laatste jaren enorm populair geworden in bijna elk gebied van de psychologie (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005). Er zijn inmiddels veel ver-

schillende impliciete meetinstrumenten ontwikkeld (zie voor een overzicht Gawronski, 2009). Wat deze instrumenten met elkaar gemeen hebben is dat zij een construct meten zonder de respondent om een directe, verbale respons te vragen. De houding wordt dus indirect, en daarmee dus impliciet gemeten (Fazio & Olson, 2003). De term 'impliciet' heeft dan ook niet zozeer betrekking op het feit of de houding impliciet is, maar op het feit dat de houding impliciet gemeten wordt. De respondent is zich hierbij niet bewust van datgene dat gemeten wordt (Petty, Fazio, & Briñol, 2008) of kan geen bewuste controle uitoefenen op datgene dat gemeten wordt (Steffens, 2004), al zijn er aanwijzingen dat respondenten die de onderliggende principes van een IAT kennen in staat zijn om de uitkomsten van de test te beïnvloeden (zie Fiedler & Bluemke, 2005).

De meest gebruikte impliciete metingen van houdingen hebben de vorm van computertests voor automatisch optredende evaluatieve reacties, waarbij reactietijden worden bepaald als een indicatie voor de houding van de respondent (Wittenbrink & Schwarz, 2007). Veelgebruikte impliciete metingen van houdingen zijn de Impliciete Associatie Test (IAT) ontwikkeld door Greenwald, McGhee en Schwarz (1998) en 'Evaluative Priming' ontwikkeld door Fazio, Sabonmatsu, Powell en Kardes (1986). We zullen in de volgende paragrafen de principes van deze twee typen meetinstrumenten bespreken.

### *De Impliciete Associatie Test (IAT)*

De Impliciete Associatie Test (IAT) is een computertest waarbij de respondent wordt gevraagd stimuli die op het scherm worden aangeboden te sorteren in twee categorieën. De stimuli hebben betrekking op de concepten die worden geëvalueerd (bijvoorbeeld insecten, bloemen, mannen, vrouwen, hetero's of homo's, autochtonen of allochtonen) of op de evaluaties van deze concepten (bijvoorbeeld mooi, lelijk, goed, slecht, positief, negatief). De concepten worden in een andere tekstkleur aangeboden dan de evaluaties. De indelingscategorieën worden links en rechts van het computerscherm geplaatst (bijvoorbeeld insecten links, bloemen rechts in groe-

Blok1

Positief	Negatief
Geluk	

Positief	Negatief
Ramp	

Blok 2

<u>Bloemen</u>	<u>Insecten</u>
<u>Mier</u>	

<u>Bloemen</u>	<u>Insecten</u>
<u>Tulp</u>	

Blok 3

<u>Bloemen</u>	<u>Insecten</u>
of	of
Positief	Negatief
<u>Mier</u>	

<u>Bloemen</u>	<u>Insecten</u>
of	of
Positief	Negatief
Geluk	

Blok 4

<u>Insecten</u>	<u>Bloemen</u>
<u>Tulp</u>	

<u>Insecten</u>	<u>Bloemen</u>
<u>Mier</u>	

Blok 5

<u>Insecten</u>	<u>Bloemen</u>
of	of
Positief	Negatief
<u>Mier</u>	

<u>Insecten</u>	<u>Bloemen</u>
of	of
Positief	Negatief
Geluk	

Figuur 1. Impressie van een IAT taak; blok 1 tot en met 5 (tekst in groene letters is in deze figuuronderstreept).

ne tekst; positief links en negatief rechts in zwarte tekst, zie Figuur 1). Vervolgens worden de stimuli in de corresponderende kleur in het midden van het scherm aangeboden (mier of roos in groen, ramp of geluk in zwart). De respondent sorteert met twee toetsen van het toetsenbord (bijvoorbeeld 'A' en 'L') de stimuli in de bijbehorende categorie.

De IAT bestaat naast één of meer oefenblokken uit twee testblokken. In het voorbeeld van houdingen ten aanzien van insecten en bloemen worden in het eerste testblok (Blok 3 in Figuur 1) de categorieën 'bloemen' en 'positief' aan de linkerkant van het computerscherm geplaatst (en is de 'A' de corresponderende sorteertoets) en 'insecten' en 'negatief' rechts (met de 'L' als corresponderende sorteertoets). In het tweede testblok (Blok 5 in Figuur 1) worden de categorieën 'bloemen' en 'insecten' omgekeerd, zodat 'insecten' en 'positieve' woorden met dezelfde knop dienen worden gesorteerd. De onderliggende assumptie van de IAT is dat een gedeelde responseknop tot kortere reactietijden leidt als een respondent beide categorieën sterker associeert. Iemand met een sterk positieve associatie met bloemen en/of een sterk negatieve associatie met insecten zal in het voorbeeld kortere reactietijden laten zien bij het eerste testblok dan bij het tweede testblok (Greenwald et al., 1998). Een houdingscore wordt bepaald door de gemiddelde reactietijden van de responses van beide blokken van elkaar af te trekken.

#### *Evaluatieve Priming*

Een priming-test is gebaseerd op een principe dat vergelijkbaar is met de IAT. Ook bij een priming-test worden stimuli met twee toetsen van het toetsenbord gesorteerd. In tegenstelling tot de IAT worden echter alleen de evaluaties gesorteerd (bijvoorbeeld 'positief' en 'negatief'). Voordat het woord verschijnt dat moet worden gesorteerd (bijvoorbeeld 'hemel' of 'ziekte') wordt een concept (bijvoorbeeld 'spin') kort – enkele miliseconden – in het midden van het beeldscherm getoond (ge-primed). Niet aan alle stimuli gaat deze prime vooraf. De prime 'spin' wordt afgewisseld met een neutrale prime, die doorgaans uit een betekenisloze reeks van letters bestaat (bijvoorbeeld 'bbbbbb'). Bij deze

priming-test is de veronderstelling dat iemand met een negatieve houding ten aanzien van spinnen langere reactietijden laat zien als een positief woord voorafgegaan wordt door de 'spin' prime dan wanneer een positief woord voorafgegaan wordt door de neutrale prime (zie Wittenbrink & Schwartz, 2007). De houdingscore wordt bepaald door de gemiddelde reactietijden van de responses op negatieve woorden na de 'spin'-prime af te trekken van de gemiddelde reactietijden van de responses op negatieve woorden na de neutrale prime.

Hieronder zullen we enkele psychologische onderzoeken beschrijven waarin met dergelijke instrumenten is gewerkt. Vervolgens zullen we twee onderzoeken presenteren als voorbeelden van toepassingen van deze instrumenten in onderzoek naar houdingen van leraren.

### 3. Uitkomsten van onderzoek met impliciete metingen

#### **3.1 Psychologisch onderzoek met impliciete metingen**

In de sociale psychologie is inmiddels veel onderzoek verricht waarbij houdingen impliciet zijn gemeten. Enkele willekeurige voorbeelden van onderzoek zijn die naar impliciete houdingen ten aanzien van religies (Rowatt, Franklin, & Cotton, 2005), bepaalde merken van levensmiddelen (Maison, Greenwald, & Bruin, 2004), roken (Sherman, Rose, Koch, Presson, & Chassin, 2003), en etnische groepen (Dovidio, Kawakami, & Gaertner, 2002; DeSteno, Dasgupta, Bartlett, & Caidric, 2004). Met name voor onderzoek naar vooroordelen blijken impliciete meetinstrumenten veelvuldig te worden ingezet.

De waarde van impliciete metingen van houdingen blijkt voornamelijk uit de mate waarin deze instrumenten het gedrag of oordelen van mensen kunnen voorspellen (Greenwald, Poehlman, Uhlmann, en Banaji, 2009). Fazio en Olson (2003) bespreken verschillende empirische onderzoeken waarbij gebruik gemaakt is van impliciete metingen van houdingen, waaronder promotieonderzoek van Jackson (1997) dat liet zien dat respondenten een essay, waarvan zij dachten dat

het door een Afro-Amerikaanse student was geschreven, significant negatiever beoordeelde wanneer zij een negatievere impliciet gemeten houding hadden ten opzichte van mensen met een andere huidskleur (Jackson, 1997). Towles-Schwen en Fazio (2006) onderzochten de kwaliteit van interetnische relaties tussen kamergenoten op een universiteit in de VS. Daar is het op universiteiten gebruikelijk dat 1<sup>e</sup>-jaars studenten op basis van willekeurige selectie samen een kamer toegewezen krijgen. De uitkomsten van dit onderzoek lieten zien dat de impliciet gemeten houding van blanke studenten ten opzichte van mensen met een andere huidskleur voorspellend was voor de duur van het samenwonen. Bij blanke studenten met een negatieve houding werd het samenwonen met een Afro-Amerikaanse student eerder beëindigd dan bij blanke studenten met een positievere houding.

Niet alleen op het gebied van interetnische houdingen blijken impliciete metingen voorspellers te zijn van gedrag. Zo onderzochten Bessenoff en Sherman (2000) de houding van respondenten ten opzichte van dikke mensen. Uit dit onderzoek bleek dat hoe negatiever de impliciet gemeten houding ten opzichte van dikke mensen was, des te verder weg van een dik persoon de respondenten hun stoel plaatsten. Arcuri, Castelli, Galdi, Zogmaister en Amadori (2008) lieten zien dat een impliciete maat om politieke houdingen te meten in staat was het stemgedrag een maand later te voorspellen.

Een fraai voorbeeld van het gebruik van de IAT in de pedagogiek is het onderzoek van Castelli, Zogmaister en Tomelleri (2009). Zij onderzochten in de VS in hoeverre houdingen van kinderen ten aanzien van blanken en Afro-Amerikanen gerelateerd waren aan de impliciet gemeten etnische vooroordelen van ouders. Uit hun onderzoek bleken de houdingen van de kinderen significant gecorreleerd te zijn met de vooroordelen van hun moeders en niet met die van hun vaders en evenmin met expliciet gemeten vooroordelen van beide ouders.

Het hierboven aangehaalde onderzoek van Castelli et al. (2009) is een voorbeeld van onderzoek waaruit gebleken is dat impliciete metingen een sterkere predictieve waarde

hebben dan expliciete metingen van houdingen. Hoewel vaker is geconstateerd dat de predictieve validiteit van impliciete metingen sterker is dan die van expliciete metingen (zie bijvoorbeeld Dovidio, Kawakami, & Gaertner, 2002), blijkt dit zeker niet in alle onderzoeken eenduidig het geval te zijn (zie bijvoorbeeld Rowatt et al., 2005). In een meta-analyse naar de predictieve validiteit van IAT's rapporteren Greenwald et al. (2009) een gemiddeld effect in termen van correlaties tussen scores op impliciete metingen en criteriumscores ( $r$ ) van .274. De grootte van effecten bleek afhankelijk te zijn van het type attitude dat werd gemeten en het type uitkomstmaat dat als criterium werd gebruikt. Zo bleek uit deze meta-analyse dat in onderzoek naar politiek stemgedrag en consumentengedrag de effecten van expliciete metingen op het gedrag groter zijn dan bij sociaal gevoeliger onderwerpen, zoals groepsstereotyperingen. Echter, bij onderzoek naar groepsstereotiepe houdingen bleek de predictieve validiteit van impliciete metingen tweemaal groter ( $r = \pm .20$ ) te zijn dan van expliciete metingen ( $r = \pm .10$ ). Greenwald et al. constateren tevens dat rapportages van effectmaten in veel onderzoeksrapportages ontbreken en dat er weinig onderzoek is uitgevoerd waarin de predictieve validiteit van zowel expliciete als impliciete metingen is getoetst. Om beter zicht te krijgen op de validiteit van metingen van houdingen, wordt er dan ook voor gepleit om impliciete en expliciete metingen in combinatie te gebruiken in onderzoek (zie bijvoorbeeld Maison et al., 2004).

### **3.2 Samenhang tussen impliciete en expliciete meetinstrumenten**

IAT's en zelfrapportage kunnen beide worden ingezet om houdingen van mensen ten aanzien van verschillende onderwerpen te meten. Wanneer men de houding ten aanzien van hetzelfde onderwerp zou meten met een vragenlijst en een IAT zou men ook samenhang tussen beide typen metingen kunnen verwachten, aangezien men dezelfde houding tracht te meten. Dit blijkt echter niet altijd het geval te zijn: er worden erg hoge, maar ook erg lage correlaties gevonden tussen scores op IAT's en vragenlijsten. In eerder Amerikaans onderzoek werd een erg ster-

ke samenhang ( $r = .75$ ) gevonden tussen een IAT en een vragenlijst over de voorkeur van mensen voor democraten of republikeinen en tevens werd een sterke samenhang tussen beide meetinstrumenten gevonden bij de voorkeur voor Coca Cola of Pepsi ( $r = .65$ ) (Nosek, 2007). Er was echter nauwelijks een samenhang te vinden tussen impliciete en expliciete metingen van de voorkeur voor Aziaten of blanken ( $r \approx .00$ ). Verschillen in de samenhang tussen scores op vragenlijsten en IAT's op diverse onderwerpen zouden deels verklaard kunnen worden door de sociaal gevoeligheid van het onderwerp waarover de houding gemeten wordt. Bij sociaal gevoelige onderwerpen denken mensen langer na over het geven van hun antwoord op een vragenlijst (Hofmann, et al., 2005). Hierdoor kan het voorkomen dat men meer sociaal wenselijk gaat antwoorden. Aangezien de sociaal wenselijkheid door het gebruik van impliciete metingen geminimaliseerd wordt, kan het dus voorkomen dat de scores van een persoon op de vragenlijst nauwelijks samenhangen met de scores op de impliciete meting. In een meta-analyse van correlaties tussen scores op impliciete metingen en zelfrapportages (Hofman et al., 2005) bleek dat naast strategisch of sociaal wenselijk antwoordgedrag andere factoren van invloed kunnen zijn op de correlaties tussen beide typen scores, zoals de spontaniteit van responses of verschillen in de conceptualisering van expliciete en impliciete metingen van attitudes (zie ook Gawronski, 2009).

### **3.3 Impliciete metingen van groepsstereotiepe houdingen van leraren**

Om na te gaan in hoeverre impliciete metingen van toegevoegde waarde kunnen zijn binnen het onderwijskundig onderzoek hebben wij recentelijk twee studies verricht (Van den Bergh, Denessen, Hornstra, Voeten, & Holland, 2010; en Hornstra, Denessen, Bakker, Van den Bergh, & Voeten, 2010). In beide studies is onderzocht in hoeverre impliciet gemeten houdingen van leraren in staat waren de verwachtingen van leerkrachten en hiermee indirect de prestaties van leerlingen te voorspellen.

Het onderzoek van *Van den Bergh et al.* (2010) betrof etnische vooroordelen van 41 leraren in het basisonderwijs en de relatie hiervan met de leerkrachtverwachtingen en de Cito-scores rekenen en begrijpend lezen van de autochtone en Turkse of Marokkaanse leerlingen in hun klassen. Ten behoeve van dit onderzoek is een IAT ontwikkeld, waarbij jongensnamen moesten worden gesorteerd in de categorieën 'Nederlands' (bijvoorbeeld Tim of Niels) en 'Turks/Marokkaans' (bijvoorbeeld Mohammed of Ahmed). Daarnaast moesten positieve (bijvoorbeeld plezier en vrede) en negatieve woorden (bijvoorbeeld pijn en ramp) worden gesorteerd. De met de IAT verkregen houdingscores bleken nauwelijks te correleren met de houdingscores die met een vragenlijst waren verkregen ( $r = -.06$ ).

Uit multilevel analyses bleek dat de met behulp van een IAT gemeten houding van leraren ten opzichte van allochtonen voorspellend waren voor de prestaties van allochtone en Turkse of Marokkaanse leerlingen. In klassen van leraren met sterkere impliciet gemeten vooroordelen jegens Turkse of Marokkaanse kinderen was het verschil in Cito-scores op rekenen en begrijpend lezen tussen Turkse of Marokkaanse en autochtone leerlingen significant groter dan in klassen van leraren met minder sterke vooroordelen jegens Turkse of Marokkaanse kinderen. Vanwege het multilevel karakter van het onderzoek is de grootte van de effecten van de impliciete houding geëvalueerd door de verschillen in scores tussen de Turkse of Marokkaanse en autochtone leerlingen in klassen van verschillende leerkrachten met elkaar te vergelijken. Wanneer de leerkracht met de meest negatieve houding vergeleken werd met de leerkracht die de meest positieve houding had, verschilde het voorspelde verschil in prestaties tussen Turkse of Marokkaanse en autochtone leerlingen (hun gestandaardiseerde scores op de Cito-toets rekenen) 1.09 punten. Voor de Cito-scores begrijpend lezen werd een vergelijkbaar verschil gevonden, namelijk een verschil van 1.05 punten. In vergelijking met leraren met een gemiddelde score op de IAT was het verschil in prestaties van autochtone en Turkse of Marokkaanse leerlingen in klassen van leraren met de laagste score op de IAT 0.58 groter voor de

gestandaardiseerde rekentoetscore en 0.56 groter voor de gestandaardiseerde score op de begrijpend leestoets. Oftewel; zelfs bij min of meer gemiddelde impliciete houdingsscores, zijn er middelmatige effecten van de houding op de prestatieverschillen tussen Turkse of Marokkaanse leerlingen en autochtone leerlingen.

Het effect van impliciet gemeten vooroordelen op de prestaties van Turkse en Marokkaanse leerlingen bleek volledig gemedieerd te worden door de verwachtingen die leerkrachten van deze leerlingen hadden. De vragenlijstscores met betrekking tot etnische vooroordelen van de leraren bleken geen enkele relatie met de verwachtingen en de prestaties te hebben.

Het onderzoek van *Hornstra et al.* (2010) was gericht op kinderen met dyslexie. Hoewel het bij deze groep minder voor de hand lijkt te liggen, kunnen kinderen met leerproblemen ook met groepsstereotypering te maken krijgen. Zo kunnen leerkrachten verschillen in hun opvattingen ten aanzien van inclusief onderwijs en hun bereidheid om deze kinderen te onderwijzen (Avramidis & Norwich, 2002; Lindsay, 2007). Opvattingen over de aard van leerproblemen en hun eigen rol in het schoolsucces van deze leerlingen kunnen tevens verschillen, wat van invloed blijkt op de wijze waarop zij omgaan met deze leerlingen (Jordan, Schwarz, & McGhie-Richmond, 2009; Jordan & Stanovich, 2001). Ook bij deze leerlingen speelt de houding van de leerkracht dus mogelijk een rol. In het onderzoek van Hornstra et al. (2010) werd bij 30 leraren in het reguliere basisonderwijs een priming test afgenomen om zicht te krijgen op negatieve associaties van leraren met de term 'dyslexie'. In deze test moesten leraren positieve en negatieve woorden (dezelfde woorden als in het onderzoek van Van den Bergh et al., 2010) categoriseren. Een deel van de woorden werd gedurende 250 miliseconden voorafgegaan met de prime 'dyslexie'. De overige woorden werden voorafgegaan met de neutrale prime 'bbbbbb'. Ook in deze studie is met een vragenlijst expliciet naar de houding van leraren ten aanzien van dyslexie gevraagd. De impliciete en expliciete metingen van de houding van leraren ten aanzien van dyslexie bleken nauwe-

lijks met elkaar samen te hangen ( $r = .05$ ).

Multilevel analyses lieten zien dat het met de impliciet gemeten houding van leraren ten aanzien van dyslexie mogelijk was de prestaties van de leerlingen te voorspellen. In klassen van leraren met een meer negatieve houding ten aanzien van dyslexie was het verschil in Cito-scores op spelling tussen leerlingen met en zonder dyslexie groter dan in klassen met een meer positieve leraar. Afhankelijk van de impliciet gemeten houding, kon de Cito-score tot maximaal 1.90 standaarddeviaties verschillen tussen de dyslectische leerlingen en de niet-dyslectische leerlingen; bij de leerkracht met de meest negatieve houding was de voorspelde waarde voor de Cito-scores van de dyslectische leerlingen 1.82 standaarddeviatie lager dan die van de niet-dyslectische leerlingen, terwijl de voorspelde Cito-score van dyslectische leerlingen van de leerkracht met de meest positieve houding 0.08 standaarddeviatie hoger was dan die van de niet-dyslectische leerlingen.

In dit onderzoek werden ook de verwachtingen die leerkrachten hadden van hun leerlingen opgenomen. In tegenstelling tot het onderzoek van Van den Bergh, et al. (2010), bleken leerkrachtverwachtingen de relatie tussen de impliciet gemeten houding en leerlingprestaties niet te mediëren. Er werd echter wel een direct effect van de leerkrachtverwachtingen op de prestaties van de leerlingen gevonden. Dit effect gold zowel voor leerlingen met als zonder dyslexie. Ook in dit onderzoek bleek de expliciete maat niet voorspellend te zijn voor de uitkomsten bij leerlingen.

## 4 Discussie

Als groepsstereotiepe houdingen impliciet worden gemeten blijken ze vaker voorspellend te zijn voor later gedrag dan als ze expliciet worden gemeten. Met de introductie van impliciete metingen in onderzoek naar de effecten van groepsstereotiepe houdingen van leraren zijn we in staat gebleken om de predictieve waarde van deze houdingen voor de prestaties van leerlingen empirisch vast te stellen, zowel waar het gaat om allochtone



leerlingen (Van den Bergh et al., 2010) als om leerlingen met dyslexie (Hornstra et al., 2010). In vergelijking met expliciete metingen, bleken impliciete metingen die betrekking hebben op oordeelsvorming van leerkrachten betere voorspellers te zijn van academische prestaties van leerlingen uit groepen waarover voordoordelen kunnen bestaan. Dit sluit tevens aan bij de assumpties van het MODE model (Fazio & Towles-Schwen, 1999). Leerkrachten ervaren het zeer waarschijnlijk als sociaal onwenselijk om eventuele negatieve houdingen te rapporteren, waardoor zij gemotiveerd zullen zijn om een meer positieve houding te rapporteren als zij daartoe de mogelijkheid hebben. Impliciete metingen van houdingen meten automatisch optredende reacties die moeilijk bewust te controleren zijn (Fazio & Olson, 2003). In de klas zullen automatisch optredende reacties van leerkrachten ten opzichte van bepaalde groepen leerlingen ook een rol spelen. Volgens het MODE model zou dit veroorzaakt kunnen worden doordat de leraar voor de klas niet altijd kan beschikken over de mogelijkheid of de motivatie om eventuele negatieve houdingen te controleren. De uitkomsten van beide onderzoeken duiden er dus op dat leerkrachten in de klas tegenover hun leerlingen mogelijk negatieve houdingen – die ze niet rapporteren op een vragenlijst – wel uiten. Dit gebeurt waarschijnlijk onbewust en onbedoeld, maar het lijkt wel een invloed te hebben op de prestaties van de leerlingen die het betreft.

Naast het MODE model zijn de gevonden resultaten ook te verklaren vanuit de dimensie van de houding die gemeten wordt en daarmee corresponderend het type gedrag dat mogelijk beter voorspeld wordt door impliciete dan door expliciete metingen (Gawronski & Bodenhausen, 2006) Dovidio, Kawakami, en Gaertner (2002) hebben aangetoond dat expliciete metingen van houdingen betere voorspellers zijn van verbaal gedrag – dat in grote mate bewust te controleren is – terwijl impliciete metingen van houdingen juist betere voorspellers zouden zijn van relatief automatisch optredend non-verbaal gedrag. Juist uit het non-verbale gedrag valt iemands houding af te leiden (Dovidio, et al., 2002). Leerlingen merken

voornamelijk door het non-verbale gedrag van hun leerkracht een mogelijk negatieve houding op, waardoor hun prestaties beïnvloed kunnen worden. Ook Rosenthal (2003) heeft laten zien dat non-verbaal leerkrachtgedrag dé mediërende factor is die het verband verklaart tussen oordeelsvorming van leerkrachten en prestaties van leerlingen. Mogelijk zijn expliciete metingen dus wel in staat om verbaal leerkrachtgedrag ten opzichte van leerlingen uit bepaalde groepen te voorspellen, maar heeft het non-verbale leerkrachtgedrag, wat beter voorspeld zou kunnen worden door impliciete metingen, een groter effect op de prestaties van de leerlingen.

In de twee genoemde studies zijn alleen de verbanden tussen impliciete en expliciete metingen van houdingen, leerkrachtverwachtingen en prestaties van leerlingen direct onderzocht. Het gedrag van de leerkracht in de klas is in beide studies buiten beschouwing gelaten. Met behulp van observatiestudies zou verder onderzocht kunnen worden in hoeverre impliciete en expliciete houdingen inderdaad voorspellers zijn van verbaal en non-verbaal gedrag en het effect hiervan op de prestaties van de leerlingen. Hoewel er dus nog meer onderzoek nodig is naar de mediërende factoren die de samenhang tussen impliciete metingen van houdingen en leerprestaties van leerlingen uit bepaalde groepen verklaren, geven de twee studies die in dit artikel besproken zijn aanwijzingen dat impliciete metingen van houdingen betere voorspellers zijn van verschillen tussen prestaties van groepen leerlingen dan expliciete metingen.

In de besproken studies wordt verondersteld dat de houdingen van de leerkracht effect hebben op hun verwachtingen en op de prestaties van de leerlingen. Het is echter niet geheel uit te sluiten dat bevooroordeelde houdingen van leerkrachten het gevolg zijn van hun dagelijkse ervaringen met de laag presterende allochtone of dyslectische leerlingen die in het onderzoek waren betrokken. In dat geval zouden de prestaties van deze leerlingen de houding van de leerkracht vormen en niet andersom. Echter, onderzoek naar de professionele ontwikkeling van docenten suggereert dat houdingen van leerkrachten het resultaat zijn van een lange

geschiedenis van opgedane ervaringen in verschillende socialisatiecontexten, zoals de persoonlijke achtergrond van de leerkracht, ervaringen tijdens de opleiding en ervaringen met verschillende groepen leerlingen tijdens de professionele loopbaan (Beijaard, Verloop, & Vermunt, 2000). Longitudinaal onderzoek zou meer zicht kunnen geven op de mogelijk wederkerige relaties tussen houdingen en leerprestaties.

Opvallend was dat leerkrachtverwachtingen in de twee onderzoeken naar allochtone leerlingen en dyslectische leerlingen een zeer verschillende rol innamen. De houding ten opzichte van allochtone leerlingen bleek bepalend te zijn voor de verwachtingen die leerkrachten van deze leerlingen hadden. Bij de houding ten opzichte van dyslexie bleek er echter geen samenhang met de verwachtingen van leerlingen met dyslexie. Zowel de houding als de verwachtingen hadden een direct verband met de prestaties van de leerlingen. Verder onderzoek zou meer inzicht kunnen bieden in mogelijke verklaringen voor de gevonden verschillen tussen beide studies met betrekking tot de rol van verwachtingen. Desalniettemin bleek er voor zowel Turkse en Marokkaanse leerlingen als voor leerlingen met dyslexie een verband tussen prestaties en de impliciete meting van de houding van de leerkracht. Vervolgonderzoek zou inzicht kunnen bieden in de vraag in hoeverre dergelijke groepsstereotypering ook speelt voor andere groepen 'risico'-leerlingen.

Ook binnen andere richtingen van onderwijskundig onderzoek zouden impliciete metingen van houdingen een belangrijke bijdrage kunnen leveren. Hierbij valt bijvoorbeeld te denken aan onderzoek naar de houding van leerkrachten ten opzichte van bepaalde groepen ouders. Daarnaast zouden impliciete metingen toegepast kunnen worden in onderzoek naar opvattingen van leerkrachten over instructievormen of in onderzoek naar oordeelsvorming van leerlingen ten opzichte van klasgenoten uit bepaalde groepen, zoals klasgenoten met een andere etnische achtergrond, een leerprobleem, of bijvoorbeeld een lichamelijke handicap.

Naast de toegevoegde waarde die impliciete metingen kunnen hebben binnen het onderwijskundig onderzoek, kunnen deze in-

strumenten in de praktijk ook een belangrijke bijdrage leveren. Impliciete metingen zoals de IAT en Priming kunnen relatief eenvoudig en snel worden afgenomen en leerkrachten zouden met deze instrumenten kunnen achterhalen in hoeverre zij zelf, wellicht onbewust, bepaalde vooroordelen hebben. Bewustwording van vooroordelen kan een eerste stap zijn in het verminderen van de negatieve effecten ervan (Stephan & Stephan, 2005).

Samengevat kan gesteld worden dat de twee besproken studies aanwijzingen hebben gegeven dat impliciete metingen van houdingen voorspellers kunnen zijn van prestaties van bepaalde groepen leerlingen en dat impliciete metingen daarom wellicht een bijdrage kunnen leveren aan onderwijskundig onderzoek, met name wanneer het onderwerp van onderzoek sociaal gevoelig van aard is.

## Literatuur

- Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology*, 52(1), 27-58.
- Ajzen, I. (2005). *Attitudes, personality, and behavior* (2nd ed.). Milton-Keynes, England: Open University Press (McGraw-Hill).
- Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C., & Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology*, 29, 367-389.
- Avramidis, E., & Norwich, B. (2002). Teachers' attitudes towards integration/inclusion: A review of the literature. *European Journal of Special Needs Education*, 17, 129-147.
- Babad, E., Bernieri, F., & Rosenthal, R. (1991). Students as judges of teachers' verbal and nonverbal behavior. *American Educational Research Journal*, 28, 211-234.
- Beijaard, D., Verloop, N., & Vermunt, J. D. (2000). Teachers' perceptions of professional identity: An exploratory study from a personal knowledge perspective. *Teaching and Teacher Education*, 16, 749-764.
- Bessenoff, G. R. en Sherman, J. W. (2000) Automatic and controlled components of prejudice toward fat people: evaluation versus stereotype activation. *Social Cognition*, 18, 329-53.
- Castelli, L., Zogmaister, C., & Tomelleri, S. (2009).

- The transmission of racial attitudes within the family. *Developmental Psychology*, 45 (2), 586-591.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6nd ed.) New York: Routledge.
- DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Caidric, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science*, 15 (5), 319-324.
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82, 62-68.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, 54, 297-327.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229-238.
- Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (Eds.), *Dual-process Theories in Social Psychology* (pp. 97-116). New York: Guilford Press.
- Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the implicit association tests. *Basic and Applied Social Psychology*, 27 (4), 307-316.
- Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology*, 50 (3), 141-150.
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692-731.
- Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, 2, 181-193.
- Greenwald, A. G., McGhee, D. E., & Schwarz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464-1480.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97 (1), 17-41.
- Hofmann, W., Gawronski, B., Gschwendner, T. Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, 31, 1369-1385.
- Hornstra, L., Denessen, E., Bakker, J., van den Bergh, L., & Voeten, M. (2010). Teacher Attitudes toward Dyslexia: Effects on Teacher Expectations and the Academic Achievement of Students with Dyslexia. *Journal of Learning Disabilities*, 43, 515-529. doi: 10.1177/0022219409355479.
- Jackson, J. R. 1997. *Automatically activated racial attitudes*. PhD thesis, Indiana University.
- Jobe, J. B. (2000). Cognitive processes in self-report. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The science of self-report: Implications for research and practice* (pp. 25-29). Mahwah, NJ: Lawrence Erlbaum.
- Jordan, A. & Stanovich, P. (2001). Patterns of teacher-student interaction in inclusive elementary classrooms and correlates with student self-concept. *International Journal of Disability, Development and Education*, 48, 33-52.
- Jordan, A., Schwartz, E., & McGhie-Richmond, D. (2009). Preparing teachers for inclusive classrooms. *Teaching and Teacher Education*, 25, 535-542.
- Jussim, L., Eccles, J., & Madon, S. J. (1996). Social perception, social stereotypes, and teacher expectancies: Accuracy, and the quest for the powerful self-fulfilling prophecy. *Advances in Experimental Social Psychology*, 29, 281-388.
- Jussim, L., & Harber, K. D. (2005). Teacher expectancies and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9, 131-155.
- Lindsay, G. (2007). Educational psychology and the effectiveness of inclusive education/mainstreaming. *British Journal of Educational Psychology*, 77, 1-24.
- Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive validity of the implicit asso-

- ciation test in studies of brands, consumer attitudes, and behavior. *Journal of Consumer Psychology*, 14 (4), 405-415.
- McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology*, 46, 235-261.
- Mitchell, W. J., Ho, C. C., Patel, H., & MacDorman, K. F. (2010). Does social desirability bias favor humans? Explicit-implicit evaluations of synthesized speech support a new HCI model of impression management. *Computers in Human Behavior*, 7, 402-412.
- Nosek, B. A. (2007). Implicit – explicit relations. *Current Directions in Psychological Science*, 16, 65-69.
- Petty, R. E., Fazio, R. H., & Briñol, P. (Eds) (2008) *Attitudes: Insights from the new implicit measures*. Mahwah, NJ: Erlbaum.
- Quillian, L. (2008). Does unconscious racism exist? *Social Psychology Quarterly*, 71, 6-11.
- Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science*, 3, 176-179.
- Rosenthal R. (2003). Covert communication in laboratories, classrooms and the truly real world. *Current Directions in Psychological Science*, 12, 151-154.
- Rowatt, W. C., Franklin, L. M., & Cotton, M. (2005). Patterns and personality correlates of implicit and explicit attitudes toward christians and muslims. *Journal for the Scientific Study of Religion*, 44 (1), 29-43.
- Schaeffer, N. C. (2000). Asking questions about threatening topics: A selective overview. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The Science of self-report: Implications for research and practice* (pp. 25-29). Mahwah, NJ: Lawrence Erlbaum.
- Schwarz, N. (2008). Attitude measurement. In: W. D. Crano & R. Prislin, *Attitudes and attitude change* (pp. 41-60). New York: Psychology Press.
- Schwartz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation*, 22, 127-160.
- Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social and Clinical Psychology*, 22 (1), 13-39.
- Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology*, 51 (3), 165-179.
- Stephan, W. G., & Stephan, C. W. (2005). Intergroup relations program evaluation. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.) *On the nature of prejudice. Fifty years after Allport*. pp. 431-446). Oxford: Blackwell Publishing.
- Towles-Schwen, T., & Fazio, R. H. (2006). Automatically activated racial attitudes as predictors of the success of interracial roommate relationships. *Journal of Experimental Social Psychology*, 42, 698-705.
- Van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M. & Holland, R. (2010). The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap. *American Educational Research Journal*, 47, 497-527. doi: 10.3102/0002831209353594
- Van Tartwijk, J. (1993). *Docentgedrag in beeld: De interpersoonlijke betekenis van onverbal gedrag van docenten in de klas*. Utrecht, The Netherlands: W.C.C.
- Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling*. Cambridge, MA: Harvard University Press.
- Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit Measures of Attitudes*. New York: Guilford Press.

Manuscript aanvaard op 7 september 2011

## Auteurs

**Lisette Hornstra** is werkzaam als promovendus en docent bij de afdeling Pedagogiek, Onderwijskunde en Lerarenopleiding (POWL) van de Universiteit van Amsterdam.

**Linda van den Bergh** is werkzaam als promovendus en docent bij Fontys Pedagogische Opleidingen 's-Hertogenbosch en is verbonden aan Eindhoven School of Education van de Technische Universiteit Eindhoven.

**Eddie Denessen** is universitair hoofddocent aan de opleidingen Pedagogische Wetenschappen en Onderwijskunde en lid van het Behavioural Science Institute van de Radboud Universiteit Nijmegen.

*Correspondentieadres:* Lisette Hornstra, Universiteit van Amsterdam, Afdeling POWL, Nieuwe Prinsengracht 130, 1018 VZ Amsterdam. Email: T.E.Hornstra@uva.nl.

## Abstract

### **Implicit measures of prejudiced attitudes of teachers**

Research on the effects of teachers' prejudiced attitudes about students typically involves the use of self-reports to assess these attitudes. However, these self-report measures have some flaws, especially because social desirability is induced when teachers are to rate their level of prejudice. In this contribution, we want to illustrate the use of implicit measures in this type of research. Two methods that were developed by social psychologists, the implicit association test and priming, will be introduced. Research on the effects of teachers' prejudiced attitudes on student achievement shows that implicit measures of prejudiced attitudes – contrary to self-report measures – are related to achievement outcomes of students from specific groups.