

Vergelijkingsvoorkeuren na het behalen van een voldoende of onvoldoende¹

M. J. P. W. Wehrens, H. Kuyper, A. P. Buunk en M. P. C. van der Werf

Samenvatting

Het halen van een onvoldoende kan bedreigend zijn voor de eigenwaarde van leerlingen. Mogelijk fungeert het kiezen van een lager vergelijkingsniveau in dergelijke situaties als een zelfbeschermende strategie. In deze studie werd onderzocht of leerlingen op de middelbare school een lager absoluut vergelijkingsniveau kozen na het inbeelden van een onvoldoende op een toets dan na het inbeelden van een voldoende op een toets. Na controle voor het gemiddelde rapportcijfer, klasstype en geslacht werd inderdaad gevonden dat leerlingen gemiddeld een lager absoluut vergelijkingsniveau kozen na een onvoldoende, maar dat dit niveau wel nog steeds relatief hoog was. Het blijkt dus dat zelfs een lichte manipulatie van het inbeelden van het halen van een onvoldoende een zelfbeschermingsreactie teweegbrengt bij leerlingen.

1 Inleiding

Zowel het daadwerkelijk halen van een onvoldoende als het inbeelden van het halen van een onvoldoende worden vaak door leerlingen ervaren als bedreigend. Dit kan verklaard worden vanuit de *self-worth theory of achievement motivation* (Covington, 1992) waarin wordt beweerd dat “het zoeken naar zelfacceptatie de hoogste menselijke prioriteit is” (p. 74) en dat “op school zelfacceptatie afhankelijk is van iemands capaciteit om in een competitieve omgeving te presteren” (p. 74). Volgens deze theorie hangt de academische eigenwaarde van leerlingen dus grotendeels af van hun academische vaardigheden. Aangezien het halen van een onvoldoende onder bepaalde omstandigheden (bijvoorbeeld na hard studeren) beschouwd kan worden als een teken van zwakke academische vaardigheden, is het voor de academische eigenwaarde van leerlingen erg belangrijk om onvoldoendes te vermijden of

verklaringen voor onvoldoendes te genereren. Leerlingen passen veel verschillende strategieën toe om hun eigenwaarde na een onvoldoende te beschermen (zie bijvoorbeeld Covington, 1992; Martin, Marsh, & Debus, 2001, 2003). Twee bekende strategieën zijn defensief pessimisme – dat wil zeggen dat leerlingen voor een toets lage verwachtingen hebben en nadenken over mogelijke negatieve uitkomsten voordat die daadwerkelijk optreden (Norem & Cantor, 1986) – en het van tevoren creëren van mogelijke excuses voor falen – dat wil zeggen uitstelgedrag, onrealistische doelen stellen, chronisch onderpresteren, een minder belangrijke persoonlijke zwakte toegeven, minder tijd steken in het leren, etcetera. In het huidige onderzoek was de assumptie dat de bescherming van de eigenwaarde na een onvoldoende ook zou plaatsvinden door het veranderen van de sociale-vergelijkingsstandaard, dat wil zeggen iemands voorkeur voor een vergelijkingsniveau.

In het onderzoek naar sociale vergelijking is veel aandacht besteed aan het *relatieve* vergelijkingsniveau van mensen. Binnen de context van schoolprestaties verwijst een relatief vergelijkingsniveau naar het prestatieniveau van een vergelijkingssander ten opzichte van het eigen prestatieniveau van een leerling. Het relatieve vergelijkingsniveau wordt gekwalificeerd als opwaarts wanneer de vergelijking plaatsvindt met een beter presterende ander, en neerwaarts wanneer de vergelijking plaatsvindt met een slechter presterende ander. In zijn bekende theorie over neerwaartse vergelijking stelde Wills (1981) dat mensen zich na een bedreiging van hun eigenwaarde of fysieke welzijn neerwaarts zouden vergelijken, dus met een slechter presterende ander, om zich beter te voelen over zichzelf.

De theorie van Wills over neerwaartse vergelijking werd bevestigd door studies van onder andere Levine en Green (1984) waarin werd aangetoond dat kinderen meer interesse

hadden in neerwaartse vergelijkingen op het moment dat hun eigen prestaties achteruit gingen dan wanneer hun eigen prestaties niet achteruit gingen. Wills (1991) beschouwde het falen op een toets inderdaad als een bedreigende situatie waarop zijn theorie over neerwaartse vergelijking van toepassing was.

Gibbons, Benbow en Gerrard (1994) introduceerden als aanvulling op het relatieve vergelijkningsniveau het *absolute* vergelijkningsniveau. In de context van schoolprestaties houdt het absolute vergelijkningsniveau het prestatieniveau in van de klasgenoten met wie een leerling zich over het algemeen vergelijkt. Als leerling A bijvoorbeeld een 6 haalt en zich vergelijkt met leerling B die een 8 haalt, dan is 8 het absolute vergelijkningsniveau van leerling A. Uit onderzoek blijkt dat leerlingen na het halen van een onvoldoende een daling in het absolute vergelijkningsniveau laten zien. Gibbons e.a. (1994) vonden bijvoorbeeld een daling in het absolute vergelijkningsniveau over een periode van drie tot zes weken onder begaafde jongens die slechter presteerden dan ze zelf verwacht hadden in een academisch programma voor begaafde leerlingen. In een tweede studie vonden Gibbons e.a. (1994) dat zowel mannelijke als vrouwelijke universitaire studenten die dachten dat ze slecht hadden gepresteerd, een daling in het absolute vergelijkningsniveau rapporteerden. Hierop volgend toonden Gibbons, Blanton, Gerrard, Buunk en Eggleston (2000) aan dat een verandering in het absolute vergelijkningsniveau gerelateerd was aan toekomstige studieprestaties. Meer specifiek vonden ze dat, na controle voor eerdere prestaties, onder universitaire studenten een daling in het vergelijkningsniveau een negatief effect had op studieprestaties een semester later.

Opgemerkt moet worden dat een daling in het absolute vergelijkningsniveau niet per definitie gelijk is aan een relatieve neerwaartse vergelijking. Zoals Gibbons e.a. (1994) aangaven, staat een daling van het absolute vergelijkningsniveau los van het feit of het vergelijkningsniveau daalt tot een niveau onder het eigen prestatieniveau (en dus gelijk is aan een relatieve neerwaartse vergelijking) of niet. Met name in situaties waarin mensen streven naar zelfverbetering na falen (bij-

voorbeeld in een schoolklas) ligt het niet erg voor de hand dat een daling van het absolute vergelijkningsniveau gelijk is aan een daadwerkelijke neerwaartse vergelijking (vgl. Gibbons et al., 2002). Een verlaging van het absolute vergelijkningsniveau kan beschouwd worden als een zelfbeschermende strategie. Lane, Gibbons, Gerrard, Blanton en Buunk (2002) menen dat deze strategie bedoeld is om de frequentie van toekomstige negatieve uitkomsten van sociale vergelijking te reduceren. Dat wil zeggen dat iemand met een daling van het vergelijkningsniveau de kans verkleint op ongunstige uitkomsten van sociale vergelijking.

In een onderzoek naar de invloed van het verlagen van het absolute vergelijkningsniveau op toekomstige academische prestaties en optimisme als mogelijke moderator van deze relatie vonden Gibbons e.a. (2000) dat het relatieve vergelijkningsniveau een veel minder goede voorspeller was van prestaties dan het absolute vergelijkningsniveau (voor een uitgebreidere discussie van de verschillen tussen het absolute en relatieve vergelijkningsniveau, zie Gibbons et al., 2002). Op basis hiervan stelden Gibbons e.a. (2000) dat studenten hun keuze voor bepaalde vergelijkingssandiden niet zozeer baseren op het relatieve prestatieverschil met deze vergelijkingssandiden, maar dat ze bepaalde vergelijkingssandiden kiezen omdat zij het absoluut gezien goed of slecht hebben gedaan. Bovendien stelden Gibbons e.a. (2000) dat methodologisch gezien het absolute vergelijkningsniveau gevoeliger is voor veranderingen in motieven voor en gewoontes in sociale vergelijking dan het relatieve vergelijkningsniveau. Daarom raadden Gibbons e.a. (2000) aan dat in onderzoek naar sociale vergelijking, met name met betrekking tot (school)prestaties, gebruik gemaakt wordt van het absolute vergelijkningsniveau. Een bijkomend voordeel van het gebruik van het absolute vergelijkningsniveau is dat de keuzemogelijkheden voor het absolute vergelijkningsniveau niet afhankelijk zijn van iemands eigen positie op een bepaalde dimensie (Buunk & Gibbons, 2007). In een schoolklas heeft een slecht presterende leerling bijvoorbeeld meer *kans* om opwaartse in plaats van neerwaartse vergelijkingssandiden

te kiezen vanwege zijn of haar eigen positie op de prestatiedimensie. Deze afhankelijkheid geldt niet voor de keuzemogelijkheden voor het absolute vergelijkingsniveau.

In dit artikel is onderzocht hoe een academische dreiging van invloed was op het absolute vergelijkingsniveau van leerlingen. De verwachting was dat het absolute vergelijkingsniveau van leerlingen lager zou zijn na het inbeelden van een onvoldoende dan na het inbeelden van een voldoende op een toets. Er werd echter ook verwacht dat het absolute vergelijkingsniveau in de onvoldoendeconditie nog steeds relatief hoog zou zijn, omdat werd aangenomen dat het feit dat de huidige studie plaatsvond in de *schoolcontext* en het feit dat aan leerlingen werd gevraagd hun *antwoorden* op een toets te vergelijken ervoor zouden zorgen dat leerlingen zouden streven naar zelfverbetering. In de schoolcontext is het belangrijk om prestaties te verbeteren en leerlingen hebben de mogelijkheid om een onvoldoende op een toets te compenseren met een voldoende op een andere toets; het vergelijken van antwoorden op een toets kan studenten informatie geven die bruikbaar is voor deze zelfverbetering. De antwoorden van een goed presterende klasgenoot geven uiteraard meer bruikbare informatie voor zelfverbetering dan de antwoorden van een slecht presterende klasgenoot.

Schoolprestaties vormen een zeer bruikbare dimensie voor het inbeelden van een onvoldoende of een voldoende en voor het aangeven van een absoluut vergelijkingsniveau aangezien prestatiecriteria op school erg helder zijn. Leerlingen weten heel goed wat een goede en een slechte prestatie is op een toets. Aan leerlingen werd gevraagd om zich een onvoldoende (een 4) of een voldoende (een 8) op een toets in te beelden en om het cijfer aan te geven van de klasgenoot waarmee ze hun antwoorden op de toets het liefst zouden willen vergelijken, dat wil zeggen hun absolute vergelijkingsniveau. Het is aannemelijk dat leerlingen in de onvoldoendeconditie wisten hoe slecht het halen van een 4 was en dat leerlingen in de voldoendeconditie wisten hoe goed het halen van een 8 was. Bovendien was het rapporteren van hun absolute vergelijkingsniveau eenduidig; leerlingen kenden de 'betekenis' van het cijfer dat zij noemden.

Een ingebeeld onvoldoende of voldoende zal voor sommige leerlingen meer realistisch zijn dan voor andere leerlingen, afhankelijk van hun eigen schoolprestaties. Voor goed presterende leerlingen zal het niet moeilijk zijn om zich voor te stellen dat ze een 8 halen, maar het inbeelden van het halen van een 4 zal wel moeilijker zijn. Voor slecht presterende leerlingen zal het makkelijker zijn om zich voor te stellen dat ze een 4 halen dan een 8. Om te controleren voor deze invloed van schoolprestaties werd het gemiddelde rapportcijfer meegenomen in de analyses. Bovendien werd gecontroleerd voor klastype en geslacht, omdat leerlingen uit de hogere klastypen wellicht hogere absolute vergelijkingsniveaus zouden rapporteren dan leerlingen uit de lagere klastypen, en jongens en meisjes wellicht zouden verschillen qua absoluut vergelijkingsniveau.

2 Methode

2.1 Data

De studie vormde onderdeel van het groot-schalige longitudinale onderzoek "Voortgezet Onderwijs Cohort Leerlingen – 1999" (VOCL'99), waarin 19.391 leerlingen werden gevolgd vanaf de brugklas totdat zij hun opleiding op de middelbare school hadden afgerond. De dataverzameling voor de huidige studie vond plaats voor de zomervakantie in 2002 onder de leerlingen die de eerste twee jaar op de middelbare school succesvol hadden doorlopen en dus op dat moment in de derde klas zaten. Het Nederlandse middelbareschoolsysteem kent vijf klastypen waarin leerlingen eindexamen kunnen doen: vwo (hoogste niveau), havo, mavo, vbo en ivbo (laagste niveau). De huidige studie had betrekking op de drie hoogste klastypen.

De resultaten in het artikel zijn gebaseerd op 3.564 leerlingen van vwo, havo en mavo. De vragenlijst in de voldoendeconditie was ingevuld door 50,3% van deze groep; de vragenlijst in de onvoldoendeconditie was ingevuld door 49,7% van de leerlingen. De totale steekproef bestond voor 46,8% uit jongens. Met betrekking tot de drie klastypen zat 33,1% van de leerlingen op vwo, 27,7% op havo en 39,2% op mavo. De gemiddelde leef-

tijd zoals gemeten voor de zomervakantie in 2002 toen de dataverzameling plaatsvond, was 15,3 jaar ($Sd. = 0,4$). Door het ontbreken van het gemiddelde rapportcijfer bestond de uiteindelijke steekproef die voor de analyses is gebruikt uit 2.885 leerlingen.

2.2 Procedures

De variabelen die in dit onderzoek zijn gebruikt zijn verzameld met behulp van een leerlingvragenlijst die is afgenomen toen de leerlingen in de derde klas zaten. Deze vragenlijst bestond uit vier, elkaar deels overlappende versies. Elke leerling vulde één van de vier versies van de vragenlijst in. Deze versies werden willekeurig uitgedeeld aan de leerlingen. Een versie had betrekking op een denkbeeldige situatie waarin de leerling een onvoldoende haalde. Een andere versie had betrekking op een denkbeeldige situatie waarin de leerling een voldoende haalde. De andere twee versies waren niet relevant voor de huidige studie. In de voldoendeconditie kregen de leerlingen de volgende instructie: “Stel je voor dat je net een 8 op een toets hebt gehaald.” In de onvoldoendeconditie kregen de leerlingen een vergelijkbare instructie: “Stel je voor dat je net een 4 op een toets hebt gehaald.” Vervolgens beantwoordden de leerlingen een vraag over hun vergelijkingsniveau. Informatie over geslacht en klastype werd via de schooladministraties verkregen. De dataverzameling werd gecoördineerd door het Centraal Bureau voor de Statistiek (voor een beschrijving van het cohortonderzoek zie bijvoorbeeld Kuiper, Lubbers, & Van der Werf, 2003).

2.3 Variabelen

Gemiddelde rapportcijfer

Leerlingen werd gevraagd om hun rapportcijfers aan te geven voor verschillende vakken (Zijsling, Kuiper, Lubbers, & Van der Werf, 2005). Het gemiddelde rapportcijfer werd berekend over minimaal acht van twaalf ‘theoretische’ vakken, namelijk Nederlands, Engels, Duits, Frans, aardrijkskunde, geschiedenis, economie, maatschappijleer, biologie, wiskunde, natuurkunde en scheikunde. De meer ‘praktische’ vakken (informatica, techniek, tekenen, verzorging en gymnastiek)

werden niet meegenomen in de berekening van het gemiddelde rapportcijfer. Leerlingen waren niet verplicht om alle twaalf theoretische vakken te volgen. Het gemiddelde rapportcijfer was 6,7 ($Sd. = 0,7$, min. = 4,3, max. = 9,7). Door het ontbreken van het gemiddelde rapportcijfer viel 32,0% van de vwo-leerlingen in de steekproef af (16,8% jongens, 15,2% meisjes), 13,7% van de havo-leerlingen (7,3% jongens, 6,4% meisjes) en 11,9% van de mavo-leerlingen (7,3% jongens, 4,6% meisjes).

Vergelijkingsniveau

Na het lezen van de voldoende- of onvoldoendemanipulatie werd de leerlingen gevraagd: “Met wie zou je je antwoorden het liefst willen vergelijken? Met iemand die een heeft gehaald.” Leerlingen noteerden niet de naam van een klasgenoot, maar in plaats daarvan het precieze cijfer op een decimaal nauwkeurig van degene met wie ze hun antwoorden zouden willen vergelijken. Cijfers konden gegeven worden op een schaal van 1,0 tot en met 10,0. De respons was 82,2%.

3 Resultaten

3.1 Descriptieve resultaten en correlaties

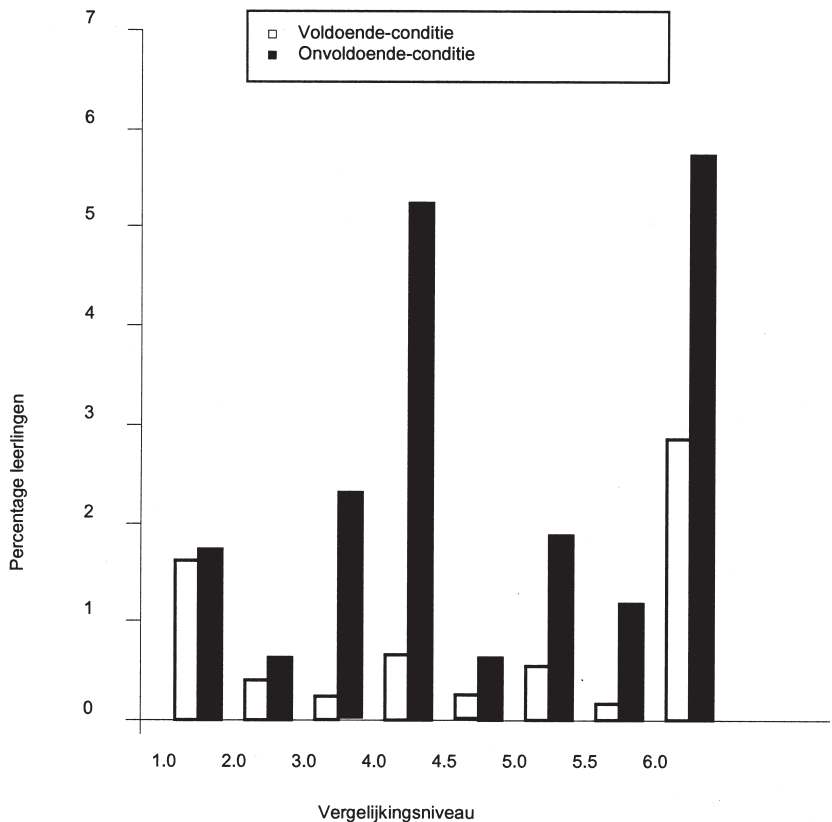
Aangezien de twee versies van de vragenlijst willekeurig waren uitgedeeld aan de leerlingen, zou het gemiddelde rapportcijfer niet moeten verschillen tussen de voldoende- en onvoldoendeconditie. Dit bleek inderdaad het geval te zijn (voldoendeconditie $M = 6,6$, $Sd. = 0,7$; onvoldoendeconditie: $M = 6,7$, $Sd. = 0,7$; $t(2.839) = -0,34$). De correlatie tussen het gemiddelde rapportcijfer en het vergelijkingsniveau was opvallend laag (gemiddeld $r = 0,12$, $p < 0,001$; voor vwo $r = 0,09$, $p < 0,01$; voor havo $r = 0,05$, n.s.; voor mavo $r = 0,10$, $p < 0,01$). Dit betekent dat leerlingen slechts een lichte neiging hadden om een hoger vergelijkingsniveau te kiezen naarmate hun eigen gemiddelde rapportcijfer hoger was. Het gemiddelde rapportcijfer en geslacht waren niet gecorreleerd ($r = -0,03$, n.s.).

3.2 Covariantieanalyse

Om de hypothese te toetsen dat leerlingen een lager absoluut vergelijkingsniveau kiezen na een onvoldoende dan na een voldoende, werd de invloed van de controlevariabelen (gemiddelde rapportcijfer, klastype en geslacht) en de predictorvariabele (conditie) op het vergelijkingsniveau onderzocht in een covariantieanalyse (ANCOVA). Vwo en havo werden als dummyvariabelen gecodeerd met mavo als referentiecategorie. Geslacht werd als dummyvariabele gecodeerd met jongens als referentiecategorie. De twee experimentele condities werden als dummyvariabele gecodeerd met de voldoendeconditie als referentiecategorie.

Zoals verwacht had *conditie* een significant effect op het vergelijkingsniveau ($F(1,2.379) = 108,5, p < 0,001, \eta_p^2 = 0,044$). Leerlingen in de voldoendeconditie rapporteerden een hoger vergelijkingsniveau dan leerlingen in de onvoldoendeconditie ($M =$

8,1, $Sd. = 1,6$ vs. $M = 7,4, Sd. = 2,1$), resulterend in een Cohens d (Cohen, 1988) van 0,39. Met betrekking tot de controlevariabelen bleek dat het gemiddelde rapportcijfer een significant effect op het vergelijkingsniveau had ($F(1,2.379) = 16,9, p < 0,001, \eta_p^2 = 0,007$). Hoe hoger het gemiddelde rapportcijfer was, des te hoger was het vergelijkingsniveau. De vwo-dummyvariabele had ook een significant effect op het vergelijkingsniveau ($F(1,2.379) = 28,0, p < 0,001, \eta_p^2 = 0,012$). Vwo-leerlingen rapporteerden een hoger vergelijkingsniveau dan mavo-leerlingen ($M = 8,1, Sd. = 1,9$ vs. $M = 7,5, Sd. = 1,9$). De havo-dummyvariabele had geen significant effect op het vergelijkingsniveau ($F(1,2.379) = 4,0$). Havo-leerlingen rapporteerden geen hoger vergelijkingsniveau dan mavo-leerlingen ($M = 7,6, Sd. = 1,9$ vs. $M = 7,5, Sd. = 1,9$). *Geslacht* had geen significant effect ($F(1,2.379) = 0,0$; jongens: $M = 7,7, Sd. = 2,0$; meisjes: $M = 7,7, Sd. = 1,7$). Voor



Figuur 1. Spreiding van vergelijkingsniveaus van 1,0 tot en met 6,0 in de voldoende- en onvoldoende-conditie.

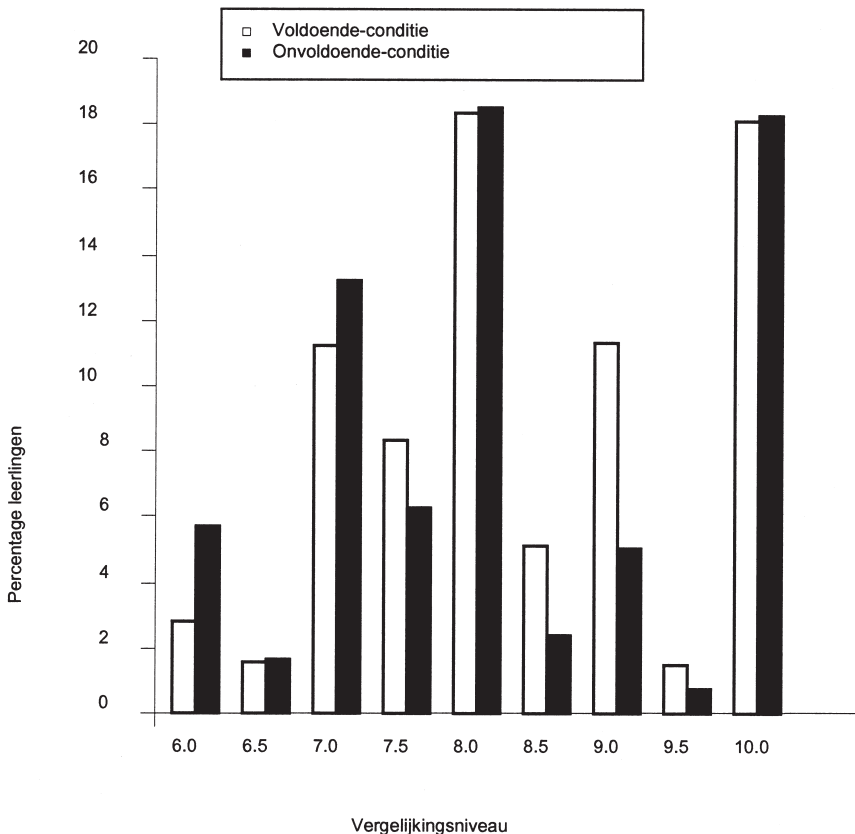
de volledigheid zijn in aanvullende analyses de tweeweginteracties getoetst tussen de controle- en predictorvariabelen, maar geen van de interacties was significant.

3.3 Aanvullende beschrijvende resultaten

De mediaan van het vergelijkingsniveau in de voldoendeconditie was 8,0; in de onvoldoendeconditie was dit 7,8. Figuur 1 toont de verdeling van de vergelijkingsniveaus van 1,0 tot en met 6,0 in de voldoende- en onvoldoendeconditie. Figuur 2 toont de verdeling van de vergelijkingsniveaus van 6,0 tot en met 10,0 in beide condities.

Uit de Figuren 1 en 2 blijkt dat leerlingen een duidelijke voorkeur hadden voor een 8,0 of een 10,0 als vergelijkingsniveau: in beide condities koos ongeveer 18,5% van de leerlingen een 8,0 als vergelijkingsniveau en ongeveer 18,0% koos ervoor om hun antwoorden te vergelijken met die van een klasgenoot met een 10,0 op de toets. De verschillen tus-

sen de beide condities, meisjes en jongens en de klastypen met betrekking tot het percentage leerlingen dat een 4,0 of lager koos als vergelijkingsniveau of een 10,0 koos als vergelijkingsniveau zijn weergegeven in Tabel 1. Met betrekking tot het vergelijkingsniveau gelijk aan of lager dan 4,0 werd het duidelijkste verschil gevonden tussen de twee condities: het percentage leerlingen dat een vergelijkingsniveau van 4,0 of lager koos was hoger in de onvoldoendeconditie dan in de voldoendeconditie (12,5% vs. 3,2%). Een ander duidelijk verschil was dat tussen jongens en meisjes: het percentage jongens dat een 4,0 of lager koos was hoger dan het percentage meisjes (9,1% vs. 6,7%). Met betrekking tot het vergelijkingsniveau van 10,0 werd het grootste verschil gevonden tussen de drie klastypen: het percentage leerlingen dat een vergelijkingsniveau van 10,0 koos daalde van vwo naar mavo (25,5% voor vwo, 17,5% voor havo en 12,1% voor mavo). Er



Figuur 2. Spreiding van vergelijkingsniveaus van 6,0 tot en met 10,0 in de voldoende- en onvoldoendeconditie.

was ook een verschil tussen jongens en meisjes: het percentage jongens dat een 10,0 koos was hoger dan het percentage meisjes (21,5% vs. 15,4%).

4 Conclusie

In het huidige onderzoek werd verwacht dat het inbeelden van een onvoldoende op een schooltoets zou leiden tot een daling in het absolute vergelijkingsniveau. Er werd echter ook verwacht dat het absolute vergelijkingsniveau na het inbeelden van een onvoldoende nog steeds relatief hoog zou zijn, omdat het streven naar zelfverbetering zou worden geactiveerd door het feit dat de huidige studie plaatsvond in een prestatiecontext waarin zelfverbetering na een onvoldoende mogelijk was en het feit dat aan leerlingen werd gevraagd hun antwoorden op een toets te vergelijken – het vergelijken van antwoorden op een toets is bruikbaar voor zelfverbetering. Bovendien zouden leerlingen na een onvoldoende nog steeds een hoog absoluut vergelijkingsniveau kunnen kiezen om vast te kunnen stellen of hun eigen cijfer terecht was. De verwachtingen van de huidige studie werden bevestigd. Leerlingen die zich hadden ingebeeld dat ze een onvoldoende hadden gehaald op een toets, rapporteerden gemiddeld een

lager absoluut vergelijkingsniveau dan leerlingen die zich hadden ingebeeld dat ze een voldoende op een toets hadden gehaald, maar dit gemiddelde absolute vergelijkingsniveau na een onvoldoende was nog steeds relatief hoog. Het lagere absolute vergelijkingsniveau in de onvoldoendeconditie ten opzichte van de voldoendeconditie zou geïnterpreteerd kunnen worden als een zelfbeschermingsstrategie waarmee geprobeerd werd de frequentie van toekomstige negatieve uitkomsten van sociale vergelijking te reduceren.

Opvallend was dat sommige leerlingen zelfs na het inbeelden van een onvoldoende het hoogst mogelijke absolute vergelijkingsniveau kozen. Het percentage leerlingen dat het hoogst mogelijke absolute vergelijkingsniveau koos, was in beide condities ongeveer gelijk. In een prestatiesituatie geeft het vergelijken van antwoorden met iemand met de hoogste toetsscore natuurlijk de meeste informatie voor de verbetering van de eigen toekomstige prestaties. Daarnaast rapporteerde een aanzienlijk percentage van de leerlingen in de onvoldoendeconditie een absoluut vergelijkingsniveau gelijk aan of lager dan 4,0. Deze leerlingen waren wellicht meer geïnteresseerd in zelfverrijking of zelfevaluatie dan in zelfverbetering – drie motieven die een rol kunnen spelen bij het vergelijken van

Tabel 1

Percentage leerlingen in beide condities dat een 4,0 of lager koos als vergelijkingsniveau of een 10,0 koos, zowel gemiddeld als per klastype

| Klas- type | Vergelijkingsniveau ≤ 4 | | | | | | | | |
|---------------|------------------------------|---------|-----|-------------|---------|------|-----------|---------|-----|
| | Voldoende | | | Onvoldoende | | | Gemiddeld | | |
| | Jongens | Meisjes | M | Jongens | Meisjes | M | Jongens | Meisjes | M |
| Vwo | 4,3 | 1,1 | 2,7 | 13,7 | 10,0 | 11,6 | 8,7 | 5,7 | 7,1 |
| Havo | 3,8 | 0,9 | 2,3 | 14,0 | 13,1 | 13,5 | 8,9 | 7,2 | 8,0 |
| Mavo | 5,8 | 2,9 | 4,2 | 13,4 | 11,7 | 12,5 | 9,5 | 7,3 | 8,3 |
| Gem. | 4,7 | 1,8 | 3,2 | 13,7 | 11,5 | 12,5 | 9,1 | 6,7 | 7,8 |

| Klas- type | Vergelijkingsniveau = 10 | | | | | | | | |
|---------------|--------------------------|---------|------|-------------|---------|------|-----------|---------|------|
| | Voldoende | | | Onvoldoende | | | Gemiddeld | | |
| | Jongens | Meisjes | M | Jongens | Meisjes | M | Jongens | Meisjes | M |
| Vwo | 28,9 | 21,5 | 25,1 | 28,3 | 23,9 | 25,9 | 28,6 | 22,8 | 25,5 |
| Havo | 24,5 | 12,8 | 18,2 | 18,8 | 14,9 | 16,7 | 21,6 | 13,9 | 17,5 |
| Mavo | 14,7 | 9,1 | 11,7 | 15,0 | 10,4 | 12,5 | 14,9 | 9,7 | 12,1 |
| Gem. | 22,4 | 14,3 | 18,1 | 20,6 | 16,4 | 18,3 | 21,5 | 15,4 | 18,2 |

schoolprestaties (Wood, 1989). Uit de resultaten van Studie 3 van Gibbons e.a. (2002) bleek inderdaad dat hoge absolute vergelijkingsniveaus samenhangen met een interesse in zelfverbetering en lage absolute vergelijkingsniveaus gerelateerd waren aan een interesse in zelfverrijking. Op basis hiervan is het aannemelijk dat de leerlingen die in de huidige studie het hoogst mogelijke absolute vergelijkingsniveau kozen een dermate sterk zelfverbeteringsmotief hadden dat dit niet werd beïnvloed door het inbeelden van een onvoldoende op een toets. Dit wordt ondersteund door het gelijke percentage leerlingen dat in beide condities het hoogst mogelijke vergelijkingsniveau koos.

Het is de vraag of het kiezen van het hoogst mogelijke absolute vergelijkingsniveau puur om toegang te hebben tot alle juiste antwoorden gekwalificeerd kan worden als *sociale vergelijking*. Wood (1996) definieerde sociale vergelijking als “het nadenken over informatie over een of meer andere personen in relatie tot zichzelf” (p. 520), waar *in relatie tot zichzelf* betekent dat iemand zoekt naar een, al is het maar minimale, gelijkenis of verschil tussen de ander en zichzelf op een bepaalde dimensie. Volgens deze definitie zou de bovenstaande situatie uit de huidige studie niet gekwalificeerd worden als sociale vergelijking, omdat leerlingen bij het kiezen van het hoogst mogelijke absolute vergelijkingsniveau niet lijken te zoeken naar enige overeenkomst of verschil tussen hun vergelijkingsander en zichzelf: een leerling kan eenvoudigweg een klasgenoot met een 10,0 kiezen om de antwoorden mee te vergelijken, omdat de juiste antwoorden niet op een andere, meer ‘objectieve’, manier te krijgen zijn (bijvoorbeeld van de leraar of via Google). Aan de andere kant kan aangevoerd worden dat een vergelijkingsproces noodzakelijk is om er überhaupt achter te komen wie een 10,0 heeft. Bovendien vindt volgens Festinger (1954) sociale vergelijking plaats wanneer er geen objectieve, niet-sociale bronnen zijn. Het feit dat een leerling de antwoorden van een klasgenoot wil zien omdat de antwoorden niet te verkrijgen zijn via meer objectieve bronnen, lijkt in lijn te zijn met Festingers (1954) definitie van sociale vergelijking, hoewel Festinger verwees naar

het *type* informatie dat mensen zoeken, terwijl het in de huidige studie ging om de *manier waarop* informatie wordt verkregen.

Behalve dat is onderzocht of het inbeelden van een onvoldoende zou leiden tot een verlaging van het absolute vergelijkingsniveau zijn ook de verschillen bekeken tussen de condities, geslachten en klastypen met betrekking tot het percentage leerlingen dat een vergelijkingsniveau van 4,0 of lager koos, dan wel een vergelijkingsniveau van 10,0. Het bleek dat het percentage leerlingen dat een vergelijkingsniveau van 4,0 of lager koos, hoger was in de onvoldoendeconditie dan in de voldoendeconditie en dat het percentage jongens dat een 4,0 of lager koos, hoger was dan het percentage meisjes. Het percentage jongens dat een 10,0 koos was ook hoger dan het percentage meisjes. Daarnaast daalde het percentage leerlingen dat een 10,0 koos van vwo naar mavo. Hoewel er geen hypothesen waren geformuleerd over verschillen tussen de condities, geslachten en klastypen met betrekking tot het kiezen van het hoogst mogelijke vergelijkingsniveau, ligt het voor de hand dat het percentage leerlingen dat een 10,0 kiest, daalt van vwo naar mavo, ervan uitgaande dat zelfverbetering het best wordt bereikt door de antwoorden te vergelijken met die van een klasgenoot met een 10,0 en dat dit streven naar zelfverbetering sterker is in de hogere klastypen dan in de lagere klastypen. Dit verschil tussen de klastypen werd inderdaad gevonden in het huidige onderzoek. Verwachtingen over verschillen tussen jongens en meisjes met betrekking tot de keuze voor het hoogst mogelijke vergelijkingsniveau zijn minder eenduidig te formuleren, omdat uit onderzoek naar geslacht, prestatiemotivatie en prestatiedoelen wisselende resultaten naar voren zijn gekomen. In sommige studies werden geen verschillen gevonden tussen jongens en meisjes en in andere studies werden stereotype verschillen gevonden (Hyde & Durik, 2005).

In het huidige onderzoek werd gevonden dat hoe hoger leerlingen presteerden op school, geoperationaliseerd als hun gemiddelde schoolprestatie, hoe hoger het vergelijkingsniveau was dat zij kozen. Dit komt overeen met bevindingen van Gibbons e.a. (2002) die hetzelfde, zij het sterkere, positieve

verband aantoonde tussen (school)prestaties en keuze voor een vergelijkingsniveau (Studies 1 en 3). Het lijkt er dus op dat hoewel een behoorlijk aantal leerlingen het hoogst mogelijke absolute vergelijkingsniveau koos, leerlingen ook de neiging hadden om vergelijkingsniveaus dichter bij hun eigen prestatieniveau te kiezen. Zelfevaluatie is natuurlijk een belangrijk motief voor het vergelijken van schoolprestaties. Een vergelijkingsniveau is meer informatief voor zelfevaluatie naarmate het dichter bij iemands eigen prestatie ligt.

Het huidige onderzoek toonde aan dat het absolute vergelijkingsniveau veranderde als gevolg van het inbeelden van het halen van een onvoldoende op een toets. Echter, uit het huidige onderzoek valt niet af te leiden of de *zelfevaluatie* van een leerling ook veranderde als gevolg van de manipulatie. In principe zijn er twee mogelijke verklaringen voor een daling van het absolute vergelijkingsniveau in de onvoldoende-conditie: 1) de zelfevaluatie van leerlingen bleef gelijk, maar leerlingen rapporteerden een absoluut vergelijkingsniveau *dichter* bij hun eigen onveranderde zelfevaluatie, en 2) leerlingen voelden zich slechter over hun eigen capaciteiten na het inbeelden van een onvoldoende dan na het inbeelden van een voldoende en als gevolg daarvan kozen ze een lager absoluut vergelijkingsniveau, terwijl wellicht het verschil tussen hun eigen positie en hun vergelijkingsniveau niet kleiner werd. Voor toekomstig onderzoek zou het interessant zijn om te onderzoeken of het absolute vergelijkingsniveau slechts daalt *als gevolg van* een daling van de zelfevaluatie van leerlingen, of dat de daling van het absolute vergelijkingsniveau een effect op zichzelf is. In het huidige onderzoek was geen controlegroep opgenomen. Het is daarom onduidelijk in welke conditie precies het absolute vergelijkingsniveau veranderde als gevolg van de manipulatie. Toekomstig onderzoek zou een controlegroep op moeten nemen zodat duidelijk wordt onder welke omstandigheden de veranderingen in het vergelijkingsniveau optreden.

Concluderend kan gesteld worden dat, ondanks enkele beperkingen, het huidige onderzoek laat zien dat zelfs een lichte manipulatie

van het inbeelden van het halen van een onvoldoende een zelfbeschermingsreactie teweegbrengt bij leerlingen. Opvallend genoeg was er desondanks toch een groep leerlingen die het hoogst mogelijke vergelijkingsniveau koos en die dus niet beïnvloed leek door het inbeelden van een onvoldoende, maar deze groep werd kleiner naarmate het klastype daalde. Er is een taak weggelegd voor toekomstig onderzoek om de motieven van deze groep leerlingen bloot te leggen voor hun keuze voor het hoogst mogelijke vergelijkingsniveau, en om de toekomstige prestaties van deze leerlingen te vergelijken met die van andere leerlingen, zodat zal blijken of hun keuze een positieve invloed heeft op hun prestaties. Onderzoek zou ook moeten kijken hoe volhardend leerlingen zijn in het kiezen van het hoogst mogelijke absolute vergelijkingsniveau na (het inbeelden van) herhaaldelijk falen.

Noot

- 1 Dit onderzoek is gesubsidieerd door het Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), subsidienummer 411-21-702.

Literatuur

- Buunk, A. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes*, 102, 3-21.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Covington, M. V. (1992). *Making the grade. A self-worth perspective on motivation and school reform*. New York: Cambridge.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Gibbons, F. X., Benbow, C. P., & Gerrard, M. (1994). From top dog to bottom half: Social comparison strategies in response to poor performance. *Journal of Personality and Social Psychology*, 67, 638-652.
- Gibbons, F. X., Blanton, H., Gerrard, M., Buunk, A. P., & Eggleston, T. (2000). Does social com-

- parison make a difference? Optimism as a moderator of the relation between comparison level and academic performance. *Personality and Social Psychology Bulletin*, 26, 637-648.
- Gibbons, F. X., Lane, D. J., Gerrard, M., Reis-Bergan, M., Lautrup, C. L., Pexa, N. A., et al. (2002). Comparison-level preferences after performance: Is downward comparison theory still useful? *Journal of Personality and Social Psychology*, 83, 865-880.
- Hyde, J. S., & Durik, A. M. (2005). Gender, competence, and motivation. In A. Elliot & C.S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 375-391). New York: Guilford.
- Kuyper, H., Lubbers, M. J., & Werf, M. P. C. van der. (2003). *VOCL 99-I: Technisch Rapport. Bijlage bij VOCL 99; de resultaten in het eerste leerjaar*. Groningen, Nederland: GION.
- Lane, D. J., Gibbons, F. X., Gerrard, M., Blanton, H., & Buunk, B. P. (2002). Comparison strategies in response to threat: When does social comparison make a difference? In S.P. Shohov (Ed.), *Advances in psychology research: Vol. 10* (pp. 71-97). New York: Nova Science.
- Levine, J. M., & Green, S. M. (1984). Acquisition of relative performance information: The roles of intrapersonal and interpersonal comparison. *Personality and Social Psychology Bulletin*, 10, 385-393.
- Martin, A. J., Marsh, H. W., & Debus, R. L. (2001). Self-handicapping and defensive pessimism: Exploring a model of predictors and outcomes from a self-protection perspective. *Journal of Educational Psychology*, 93, 87-102.
- Martin, A. J., Marsh, H. W., & Debus, R. L. (2003). Self-handicapping and defensive pessimism: A model of self-protection from a longitudinal perspective. *Contemporary Educational Psychology*, 28, 1-36.
- Norem, J. K., & Cantor, N. (1986). Anticipatory and post hoc cushioning strategies: optimism and defensive pessimism in "risky" situations. *Cognitive Therapy and Research*, 10, 347-362.
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245-271.
- Wills, T. A. (1991). Similarity and self-esteem in downward comparison. In J. Suls & T.A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 51-78). Hillsdale, NJ: Erlbaum.
- Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, 106, 231-248.
- Wood, J. V. (1996). What is social comparison and how should we study it? *Personality and Social Psychology Bulletin*, 22, 520-537.
- Zijsling, D., Kuyper, H., Lubbers, M. J., & Werf, M. P. C. van der. (2005). *VOCL'99-3: Technisch Rapport*. Groningen, Nederland: GION.

Manuscript aanvaard op: 16 april 2011

Auteurs

Maïke Wehrens is werkzaam aan de Afdeling Sociale en Organisatie Psychologie van de Rijksuniversiteit Groningen. **Bram Buunk** is als KNAW academie-hoogleraar werkzaam aan dezelfde afdeling. **Hans Kuyper** en **Greetje van der Werf** zijn werkzaam aan het Gronings Instituut voor Onderzoek van Onderwijs (GION) van de Rijksuniversiteit Groningen.

Correspondentieadres: Maïke Wehrens, Afdeling Sociale en Organisatie Psychologie, Rijksuniversiteit Groningen, Grote Kruisstraat 2/1, 9712 TS, Groningen. E-mail: m.j.p.w.van.der.vlugt@rug.nl.

Abstract

Comparison preferences after success or failure at school

Getting an insufficient grade at school might be threatening for students. In such a situation the choice of a lower comparison level possibly functions as a self-protective strategy. The present study investigated whether students in high school reported a lower absolute preferred comparison level after imagining a failure on a test than after imagining a success on a test. While controlling for grade point average, track and gender, it was found indeed that students on average displayed a downward shift in preferred comparison level after failure, but that this level was still relatively high. So, it appears that even a modest manipulation of imagining failure causes a self-protective reaction among high school students.