

## **Verantwoording Checklist Studievaardigheid**

C.A. de Wit

A.P.J.M. Heuvelmans

C. Sluiter



15644

3.4  
95-1  
110

OPD Memorandum

95-1

## Verantwoording Checklist Studievaardigheid

C.A de Wit

A.P.J.M. Heuvelmans

C. Sluiter

Cito

Arnhem, december 1995

Cito Instituut voor Toetsontwikkeling  
Postbus 1034 6801 MG Arnhem

Bibliotheek

8501 011 9699



Prijs f 7,50

© Cito Arnhem 1995

Uit deze uitgave mogen zonder toestemming van de auteurs korte aanhalingen met volledige bronvermelding worden overgenomen.

- de daarbij in te zetten voorwaardelijke en vakinhoudelijke kennis die nodig is om gevraagde leerprestaties, de realisatie van onderwijs-leerdoelstellingen, te kunnen leveren.

Bij onze begripsbepaling van de term studievoordigheid is grotendeels aangesloten bij de concretisering zoals die door Van Dyck et al. (1991) is gegeven. Hieronder volgt een kleine uitweiding ten aanzien van de begrippen 'sturing' en 'voorwaardelijke en vakinhoudelijke kennis' aangezien deze begrippen een belangrijke rol spelen bij het identificeren van de mate van studievoordigheid van leerlingen.

Voor het benoemen van kenmerken van leerlingen die aanleiding kunnen geven tot problemen in het vervolgtraject van een studie wordt een onderscheid gemaakt tussen kenmerken die betrekking hebben op sturing van leeractiviteiten en kenmerken die te maken hebben met voorwaardelijke en vakinhoudelijke kennis.

Sturing heeft te maken met werkplanning, voorbereiding en oriëntatie op wat er geleerd of gedaan moet worden. Een studievoordige leerling zorgt voor een planmatige aanpak en een correcte werkwijze, en zorgt voor controle, revisie en bezinning op de gevolgde werkwijze en het bereikte resultaat. Sturing veronderstelt dat de leerling een overzicht heeft van wat er aan leeractiviteiten (onderstrepen, samenvatten, herlezen, memoriseren) nodig is om leerprocessen te activeren en leerdoelen te bereiken, en dat hij weet hoe die leeractiviteiten moeten verlopen. Een leerling die in staat is een juiste sturing aan zijn leeractiviteiten te geven hanteert bijvoorbeeld een goede tijdsindeling, heeft goede leesgewoonten, zoekt hulp bij studieproblemen en durft toe te geven dat iets niet begrepen wordt.

Voor de voorwaardelijke en vakinhoudelijke kennis die nodig is voor zelfstandig leren, geldt in de eerste plaats dat een leerling tenminste een globaal idee moet hebben van datgene waarop zijn leren zich zou moeten richten. De leerdoelen moeten voor de leerling bekend zijn en hij moet weten *welke prestatie* verlangd wordt. Voorts moet hij kennis en vaardigheidselementen van meer algemeen voorwaardelijke aard bezitten: goed kunnen memoriseren, leesvaardig zijn en informatie kunnen ordenen en analyseren.

Voor elk van de hierboven genoemde kenmerken is bestaand itemmateriaal te vinden dat geschikt is om op te nemen in een *leerlingvragenlijst*. Hier is er echter expliciet voor gekozen om de checklist door *docenten* te laten invullen, aangezien het niet te verwachten valt dat leerlingen een objectief beeld van zichzelf zullen geven op een instrument dat een rol kan spelen bij het bepalen van hun verdere schoolcarrière.

Om een idee te krijgen van de soort vragen die in leerlingvragenlijsten worden gebruikt, zijn de volgende vragenlijsten aan een nader onderzoek onderworpen (Heuvelmans, intern werkdocument):

- MSQL (Pintrich et al, 1991)
- De HAVLOT methode (Gerrits et al, 1988)
- VHVO, vragenlijst huiswerk voorgezet onderwijs (Neuwahl et al, 1988).
- Vragenlijst propaedeuse zelfonderzoek (Hogeschool Eindhoven. 1989)
- Leerwijzer (Hogeschool Enschede, 1992)
- Studeergedragvragenlijst (TH Eindhoven, 1987)
- Diagnostische vragenlijst studiemethoden (Oosterhuis-Geers, 1992).
- Studie-ondersteuning propaedeuse HES-vt (Hanzehogeschool Groningen, 1992)
- Studievaardigheid en de kwalificatie voor HBO en WO (Van Dyck et al, 1991)

Met de inhoud van deze vragenlijsten als inspiratiebron is het prototype van een checklist studievaardigheid voor docenten ontwikkeld. Bij de constructie van de checklist is het uitgangspunt gehanteerd dat de aspecten van studievaardigheid die in de checklist aan de orde komen, betrekking moeten hebben op kenmerken van leerlingen die, in samenhang met andere informatie over de leerling (toetsscores en/of rapportcijfers), bepalend zijn voor succes in de bovenbouw van havo en vwo. Bovendien moeten de leerlingkenmerken in de checklist 'uiterlijk waarneembaar' of 'plausibel afleidbaar' zijn, omdat de checklist door docenten ingevuld moet worden. Vragen waarover een school of docent naar verwachting geen zinvolle of geldige informatie kan verschaffen, horen niet in de checklist thuis. Zo mogen er geen vragen gesteld worden over onderwerpen die zich doorgaans aan de waarneming van de school of de docent onttrekken, bijvoorbeeld vragen over de omstandigheden waaronder een leerling thuis zijn huiswerk maakt.

Het ontwikkelde prototype van de checklist bestond uit 35 items die elk betrekking hadden op een bepaald aspect van studievaardigheid. Docenten konden bij ieder 'beoordelingsaspect' kiezen uit drie antwoordcategorieën, te weten '—', '□', en '+', verwijzend naar een oordeel met een negatieve, neutrale en positieve strekking. De 35 beoordelingsaspecten zijn opgenomen in bijlage 1.



## 2 De proefafname

### 2.1 Opzet en uitvoering

Om informatie te verkrijgen over de meettechnische eigenschappen van de checklist is in mei 1994 een proefafname gehouden in het derde leerjaar van de havo en het vwo. Klassen namen deel aan de proefafname voor tenminste één van de vakken Nederlands, Engels en wiskunde. Aan docenten die lesgeven aan de klassen in de vakken waarvoor deze klassen deelnamen, is gevraagd om voor de eerste tien leerlingen de checklist in te vullen. Voor een uitgebreide beschrijving van de opzet van de proefafname wordt verwezen naar het verslag van de constructie van de toetsen (Sluijter, in voorbereiding).

In totaal is voor 1940 van de leerlingen die aan de proefafname hebben meegedaan het prototype van de checklist studievoordigheid door tenminste één docent ingevuld. Tabel 2.1 geeft een overzicht van het aantal leerlingen dat respectievelijk door een, twee of drie docenten is beoordeeld. In deze tabel is het type van de klas waarvan leerlingen deel uitmaakten buiten beschouwing gelaten.

Tabel 2.1  
Aantal leerlingen beoordeeld door docenten  
Nederlands, Engels en/of wiskunde

| Nederlands | Engels | Wiskunde | Aantal |
|------------|--------|----------|--------|
| X          | X      | X        | 240    |
| X          | X      | -        | 166    |
| X          | -      | X        | 149    |
| -          | X      | X        | 411    |
| X          | -      | -        | 236    |
| -          | X      | -        | 415    |
| -          | -      | X        | 323    |
|            |        |          | 1940   |

In tabel 2.2 staat het aantal leerlingen dat per klastype door docenten Nederlands, Engels of wiskunde is beoordeeld. Een voorbeeld: van de  $240 + 166 + 149 + 236 = 791$  leerlingen voor wie door een docent Nederlands een checklist is ingevuld, zaten er 272 in een 3 havo-klas, 437 in een 3 vwo-klas en 82 in een 3 havo/vwo klas.

Tabel 2.2  
Aantal leerlingen uit klastype havo, vwo en havo/vwo  
beoordeeld door docenten Nederlands, Engels of wiskunde

|            | Klastype |     |          | Totaal |
|------------|----------|-----|----------|--------|
|            | havo     | vwo | havo/vwo |        |
| Nederlands | 272      | 437 | 82       | 791    |
| Engels     | 488      | 602 | 142      | 1232   |
| Wiskunde   | 411      | 585 | 127      | 1123   |

Bij de meeste analyses die naar aanleiding van de proefafname zijn verricht, is geen onderscheid gemaakt tussen leerlingen uit verschillende klastypen. In de literatuur zijn namelijk geen aanwijzingen gevonden die wijzen op het bestaan van een per klastype specifieke studievoordigheid. Er is bij dit onderzoek vanuit gegaan dat er hoogstens sprake zal zijn van een verschil in de *mate van* studievoordigheid die leerlingen uit verschillende klastypen (havo, havo/vwo en vwo) bezitten, en niet in de *soort* studievoordigheid.

## 2.2 Onderzoeksvragen

De belangrijkste vraag van dit onderzoek is of de checklist een geschikt instrument is om de studievoordigheid van leerlingen mee te beoordelen. Voor het beantwoorden van deze vraag moeten twee aspecten in beschouwing worden genomen. Ten eerste moet de kwaliteit van de afzonderlijke items goed zijn, en ten tweede moet de checklist in zijn geheel aan bepaalde eisen voldoen. In de volgende drie paragrafen worden beide 'kwaliteitscriteria' nader gespecificeerd. Tenslotte

wordt in de vierde paragraaf nader ingegaan op de vraag of er verschillen zijn in studievoordigheid, zoals gemeten door de checklist, tussen leerlingen uit havo-, havo/vwo- en vwo-klassen.

### **2.2.1 Kwaliteit van de items**

Om in de uiteindelijke versie van de checklist te worden opgenomen, moeten de items in het prototype van de checklist aan verschillende kwaliteitseisen voldoen. Zo mogen de items niet overbodig zijn, of gebrekkig geformuleerd en moet elk item wat betreft inhoud en de manier waarop het gescoord moet worden, passen bij de rest van de items. Om te bepalen of de items aan de bovengenoemde criteria voldoen, zijn verschillende kwalitatieve en kwantitatieve itemanalyses verricht. Deze analyses worden hieronder kort toegelicht.

Om de kwaliteit van de items te onderzoeken is in de eerste plaats nagegaan of er items zijn die bij nader inzien overbodig zijn, of waarvan de formulering verbeterd zou kunnen worden. Hierbij is gebruik gemaakt van het door de deelnemende docenten gegeven commentaar.

In de tweede plaats is van ieder item in de checklist de bijdrage aan Cronbach's alpha berekend, een maat uit de klassieke testtheorie die aangeeft hoe betrouwbaar - dat wil zeggen nauwkeurig, consistent en reproduceerbaar - de scores op een meetinstrument tenminste zijn. Het is de bedoeling dat elk item in de checklist een positieve bijdrage aan alpha levert.

Om een indruk te krijgen van het onderscheidend vermogen van de items is vervolgens de item-rest correlatie ( $R_{ir}$ ) van de items berekend. De  $R_{ir}$  van een item geeft aan in hoeverre dat item onderscheid maakt tussen personen met een hoge score en personen met een lage score op een test. Van de items in de checklist wordt geëist dat zij een  $R_{ir}$  van 0.4 of hoger hebben.

Om gedetailleerde informatie te verkrijgen over de kwaliteit van de items is ten slotte onderzocht of de items van de checklist door een itemresponsmodel beschreven kunnen worden. Hiertoe is bepaald of de items van de checklist binnen het algemene

eenparameter logistisch itemresponsmodel (OPLM; Verhelst en Eggen, 1989; Verhelst en Glas, 1995) op één schaal te plaatsen zijn. Bij de bespreking van de resultaten zal nader op dit model worden ingegaan.

Eerder werd reeds vermeld dat docenten bij het invullen van de checklist steeds uit drie antwoordcategorieën konden kiezen, te weten '—', '□', en '+', verwijzend naar een oordeel met een negatieve, neutrale en positieve strekking. Aan deze categorieën werd een numerieke score van respectievelijk 1, 2 en 3 toegekend. Hierbij is er vanuit gegaan dat het zinvol is om drie antwoordcategorieën te onderscheiden. Met behulp van een homogeniteitsanalyse is onderzocht of deze aanname gegrond is. Om het scoringsvoorschrift bij de checklist eenvoudig te houden is het nodig dat bij alle items dezelfde manier van scoretoekenning gehanteerd kan worden. Items waarvoor op basis van de homogeniteitsanalyse een afwijkend scoringsvoorschrift nodig blijkt te zijn, komen minder snel in aanmerking voor opname in de definitieve checklist.

### **2.2.2 Beoordelaarsovereenstemming over de items**

De checklist is door docenten ingevuld, omdat niet te verwachten valt dat leerlingen een objectief beeld van zichzelf zullen geven op een instrument dat een rol kan spelen bij het bepalen van hun verdere schoolcarrière. Een vraag die zich nu onmiddellijk voordoet, is in hoeverre het voor de beoordeling van de studievoordigheid van een leerling uitmaakt door *welke* docent de checklist wordt ingevuld. Om deze vraag te beantwoorden zou een onderzoek naar beoordelaars-overeenstemming moeten worden verricht.

In het geval van de checklist treedt hierbij echter een probleem op. Voor het invullen van de checklist is het noodzakelijk dat een docent de leerling kent. Voor elke leerling is er *binnen een vak* daarom slechts één docent beschikbaar die de studievoordigheid van de leerling kan beoordelen. Er kan bijvoorbeeld niet aan vier docenten Engels gevraagd worden de checklist voor een bepaalde leerling in te vullen, omdat deze leerling slechts één docent Engels heeft. Het is, met andere woorden, niet mogelijk om *binnen één bepaald vak* te onderzoeken in hoeverre

verschillende docenten het met elkaar eens zijn bij het beoordelen van de studievoordigheid van hun leerlingen.

Het is uiteraard wèl mogelijk om de checklist door docenten van *verschillende vakken* te laten invullen. Per item kan dan worden onderzocht in hoeverre docenten van verschillende vakken het met elkaar eens zijn als het gaat over het beoordelen van de studievoordigheid van hun leerlingen.

Indien dit onderzoek als resultaat oplevert dat docenten van verschillende vakken het sterk met elkaar eens zijn, kunnen we twee conclusies trekken, namelijk dat studievoordigheid een eigenschap is die *niet specifiek is voor een bepaald vak*, èn dat het voor het beoordelen van de studievoordigheid van een leerling niet uitmaakt *door welke docent* de checklist wordt ingevuld.

Indien zou blijken dat verschillende vakdocenten *niet* met elkaar overeenstemmen, is het niet mogelijk om te bepalen wat daarvan de oorzaak is. De oorzaak zou kunnen zijn dat studievoordigheid een vakgebonden eigenschap is, of dat studievoordigheid niet objectief te beoordelen is. Op beide manieren is te verklaren dat docenten van verschillende vakken de studievoordigheid van een leerling anders beoordelen.

Er is voor gekozen om voor elk item in de checklist te onderzoeken in hoeverre docenten van de vakken Nederlands, Engels en wiskunde tot een zelfde oordeel komen over de studievoordigheid van hun leerlingen. Dit is alleen onderzocht voor leerlingen die door drie docenten zijn beoordeeld.

### **2.2.3 Kwaliteit van de checklist**

Nadat die items uit het prototype van de checklist zijn verwijderd die niet aan de in paragraaf 2.2.1 beschreven eisen voldoen, is onderzocht of de resulterende checklist ook in zijn geheel aan bepaalde psychometrische eisen voldoet. Verder is nagegaan of deze checklist door een enkele, willekeurige, docent ingevuld kan worden, of dat er sprake is van zulke grote verschillen tussen beoordelaars van de verschillende vakken dat de checklist door meer dan één docent moet worden ingevuld.

Om vast te stellen of de definitieve checklist in zijn geheel een goed meetinstrument is, is de betrouwbaarheid (in de vorm van Cronbach's alpha) ervan bepaald na het verwijderen van de items die niet aan de boven beschreven eisen voldeden.

Om te bepalen door hoeveel vakdocenten de studievoordigheid van een leerling beoordeeld zou moeten worden, is de mate van overeenstemming tussen de docenten bepaald voor wat betreft de totaalscore op de checklist. Op basis van de gevonden maat van overeenstemming is vervolgens geschat welke beoordelaarsovereenstemming zou worden verkregen indien niet drie maar slechts twee of zelfs maar één beoordelaar zou worden gebruikt. Op basis van deze schattingen kan bepaald worden hoeveel docenten de checklist in moeten vullen om tot een objectieve beoordeling van de studievoordigheid van een leerling te komen.

Om een indruk te krijgen van de validiteit van de checklist is de totaalscore op de checklist per vak in verband gebracht met de volgende operationalisering van het studiesucces van leerlingen:

- het cijfer op het eindrapport van 3 havo of 3 vwo van het schooljaar '93-'94 van de verschillende vakken;
- de verwachting van de docent dat de leerling wat betreft prestaties in het schooljaar '94-'95 tot de slechtere, middelmatige, of betere leerlingen zal behoren;
- schattingen van de vaardigheid van de leerling voor de vakken Nederlands, Engels en wiskunde, zoals bepaald middels het in het kader van de proefafname aan de leerling voorgelegde opgavenmateriaal voor deze vakken.

#### **2.2.4 Verschillen tussen havo- en vwo-leerlingen**

Zoals in paragraaf 2.1 al is opgemerkt, is er geen reden om aan te nemen dat het begrip studievoordigheid iets anders inhoudt bij havo-, havo/vwo- of vwo-leerlingen. Het zou echter wel zo kunnen zijn dat leerlingen uit verschillende klastypen verschillen in de *mate* waarin zij studievoordig zijn. In dat kader is onderzocht of er significante verschillen zijn tussen de gemiddelde totaalscore op de definitieve checklist voor havo-, havo/vwo- en vwo-leerlingen.

### **3. Onderzoeksresultaten**

#### **3.1 Kwaliteit van de items**

Tabel 3.1 geeft een overzicht van de resultaten van de analyses die zijn uitgevoerd om de kwaliteit van de items te bepalen. In de tabel staan alle 35 items uit het prototype van de checklist in oorspronkelijke volgorde vermeld. Een kruisje in een kolom geeft aan dat het betreffende item, bij de analyse waarop die kolom betrekking heeft, op een negatieve manier is opgevallen. De getallen in de zesde kolom hebben betrekking op de mate van beoordelaarsovereenstemming. Hieronder wordt voor elk van de zes kolommen aangegeven welke analyses zijn verricht, en welke criteria gehanteerd werden bij het bepalen van de kwaliteit van de items.

##### **3.1.1 Inhoudelijke kwaliteit**

In de eerste kolom van tabel 3.1 is aangegeven welke items bij nader inzien op inhoudelijke gronden uit de checklist verwijderd moeten worden. Het betreft de volgende items: item 7, 9, 10, 30, 33 en 34.

Pas na de proefafname bleek dat item 7 en item 9 inhoudelijk gezien niet correct zijn. Bij deze vragen moet de docent beoordelen of een leerling voor schriftelijke overhoringen (item 7) en proefwerken (item 9) 'constante cijfers' behaalt. Hiermee wordt ten onrechte geïmpliceerd dat het constant behalen van slechte cijfers een positief aspect van studievaardigheid is.

Uit commentaar van docenten bleek dat het beantwoorden van de items 10, 30, 33 en 34 te veel inzicht in de persoonlijkheid van een leerling vraagt. Een docent merkte op dat hij deze vragen niet eens voor zijn eigen kinderen zou kunnen beantwoorden, laat staan voor een van zijn leerlingen.

Tabel 3.1

Overzicht van resultaten van het onderzoek naar de kwaliteit v/d items  
in de checklist.

| Item | Omschrijving                            | Inh. | Cr. $\alpha$ | $R_{it}$ | OPLM | Hom. | $\hat{\rho}^2$ | Conclusie |
|------|-----------------------------------------|------|--------------|----------|------|------|----------------|-----------|
| 1    | Huiswerk af                             |      |              |          |      |      | 0.67           |           |
| 2    | Huiswerk netheid                        |      |              |          |      |      | 0.74           |           |
| 3    | Motivatie                               |      |              |          |      |      | 0.59           |           |
| 4    | Opletten in de klas                     |      |              |          |      |      | 0.57           |           |
| 5    | Gedrag in de klas                       |      |              |          |      |      | 0.61           |           |
| 6    | Cijfers S.O.'s hoogte                   |      |              |          |      |      | 0.51           |           |
| 7    | Cijfers S.O.'s constantheid             | X    | X            |          |      | X    | 0.32           | Uit       |
| 8    | Cijfers P.W. hoogte                     |      |              |          |      |      | 0.44           |           |
| 9    | Cijfers P.W. constantheid               | X    | X            |          |      | X    | 0.26           | Uit       |
| 10   | Zelfvertrouwen                          | X    |              |          |      | X    | 0.50           | Uit       |
| 11   | Punctualiteit                           |      |              |          |      |      | 0.57           |           |
| 12   | Omgang met leerlingen                   |      | X            |          | X    | X    | 0.34           | Uit       |
| 13   | Omgang met docenten                     |      |              |          |      |      | 0.25           |           |
| 14   | Spreekvaardigheid                       |      |              |          |      |      | 0.55           |           |
| 15   | Leesvaardigheid                         |      |              |          |      |      | 0.35           |           |
| 16   | Lesvoorbereiding                        |      |              |          |      |      | 0.57           |           |
| 17   | Absentie                                |      | X            | X        | X    | X    | 0.52           | Uit       |
| 18   | Zelfstandig werken                      |      |              |          |      | X    | 0.36           |           |
| 19   | Samenwerken                             |      |              |          | X    | X    | 0.39           | Uit       |
| 20   | Schrijfvaardigheid                      |      |              |          |      |      | 0.48           |           |
| 21   | Uitdrukkingsvaardigheid                 |      |              |          |      |      | 0.46           |           |
| 22   | Vragen wanneer iets niet begrepen wordt |      | X            |          |      |      | 0.55           |           |
| 23   | Van buiten leren                        |      |              |          |      |      | 0.45           |           |
| 24   | Nauwkeurig werken                       |      |              |          |      |      | 0.62           |           |
| 25   | Begrijpen opgedragen taken              |      |              |          |      |      | 0.44           |           |
| 26   | Aandacht in de les                      |      |              |          |      |      | 0.53           |           |
| 27   | Zelfcontrole                            |      |              |          |      |      | 0.38           |           |
| 28   | Verstandelijke capaciteit               |      |              |          | X    | X    | 0.51           | Uit       |
| 29   | Theoretische interesse                  |      |              |          |      |      | 0.50           |           |
| 30   | Incasseringsvermogen                    | X    |              |          | X    | X    | 0.08           | Uit       |
| 31   | Doorzettingsvermogen                    |      |              |          |      |      | 0.30           |           |
| 32   | Concentratievermogen                    |      |              |          |      |      | 0.45           |           |
| 33   | Faalangst                               | X    | X            | X        | X    | X    | 0.43           | Uit       |
| 34   | Stressgevoeligheid                      | X    | X            | X        | X    | X    | 0.49           | Uit       |
| 35   | Werktempo/efficiënt werken              |      |              |          |      |      | 0.35           |           |

### **3.1.2 Bijdrage aan Cronbach's alpha**

Voor elk item is berekend welke bijdrage het leverde aan Cronbach's alpha door het verschil te bepalen tussen de waarde van alpha gebaseerd op alle items, en alpha zoals berekend voor het geval het item niet in de checklist wordt opgenomen. Van elk item in de checklist werd geëist dat het een positieve bijdrage aan Cronbach's alpha levert.

### **3.1.3 Discriminerend vermogen**

De item-rest correlatie ( $R_{ir}$ ) geeft aan hoe groot de correlatie is tussen de score op een item en de totaalscore op de test, waarbij deze totaalscore gebaseerd is op de overige items. Deze maat wordt gebruikt om een indruk te krijgen van de mate waarin een item onderscheid maakt tussen personen met een hoge en een lage testscore. Van de items in checklist werd geëist dat de waarde van de  $R_{ir}$  groter moest zijn dan 0.4.

### **3.1.4 Eenparameter logistisch itemresponsmodel (OPLM)**

Om te onderzoeken of de items van de checklist alle dezelfde vaardigheid meten, is getracht om de items binnen het algemene eenparameter logistisch itemresponsmodel (OPLM; Verhelst & Eggen, 1989; Verhelst, Glas & Verstralen, 1993; Verhelst & Glas, 1995) op één schaal te plaatsen. OPLM is gebaseerd op de itemresponsstheorie (IRT). Voor een bespreking van deze theorie wordt verwezen naar Eggen en Sanders (1993). Een van de belangrijkste aannames binnen een itemresponsmodel is dat samenhang tussen antwoorden op de items uitsluitend veroorzaakt wordt door de vaardigheid die de items meten. Wat deze aanname betreft zitten er bij het analyseren van de gegevens uit de proefafname enkele haken en ogen aan de toepassing van het model. In de proefafname kan een leerling namelijk door meerdere docenten van verschillende vakken worden beoordeeld. Dit heeft tot gevolg dat een leerling meermaals in het gegevensbestand kan voorkomen. Bij een gewone toets zou dit onaanvaardbaar zijn, omdat een leerling slechts een keer dezelfde toets voorgelegd mag krijgen. Anders zijn de verschillende waarnemingen niet meer

onafhankelijk van elkaar, en kunnen de itemparameters verkeerd geschat worden. Als men er vanuit gaat dat de docenten de leerlingen onafhankelijk van elkaar hebben beoordeeld, is het echter te verdedigen dat het meervoudig voorkomen van een leerling in het gegevensbestand in dit geval wèl aanvaardbaar is. Voorzichtigheid bij het interpreteren van de analyses blijft echter geboden.

Een ander punt waar bij de analyses volgens het itemresponsmodel rekening mee moet worden gehouden is dat de items in de checklist polytome gescoord zijn. De interpretatie van de itemparameters is voor het polytome geval ingewikkelder dan bij dichotome gescoorde items. Aangezien de bespreking van deze interpretatie vrij uitvoerig is wordt hiervoor verwezen naar Verhelst, Glas en Verstralen (1993).

Ondanks de genoemde bezwaren werden met voornoemd itemresponsmodel enkele analyses verricht. Deze analyses geven namelijk meer gedetailleerde informatie over de vaardigheid van de leerlingen en de kwaliteit van de items in de checklist dan de bij gebruik van het klassieke testtheoriemodel. Aan de hand van de grafische weergave van de items werd ontdekt dat docenten bij sommige items de neiging hebben om de zwakkere leerlingen minder streng te beoordelen dan op grond van het model verwacht zou worden.

Bij de analyses is gebruik gemaakt van 2736 volledige records. Zeven items bleken niet schaalbaar te zijn. Hoewel statistische toetsing aangeeft dat de passing van de resterende achtentwintig items als geheel te wensen overlaat ( $R_{1c}=791.851$   $df=153$   $p=0.000$ ), blijkt uit de grafische modelfit voor ieder item afzonderlijk dat het scoregedrag van docenten op de individuele items toch redelijk door het model beschreven wordt.

De zeven items die niet schaalbaar bleken te zijn binnen de itemresponsanalyses, zijn in tabel 3.1 met een kruisje aangeduid.

### **3.1.5 Homogeniteitsanalyse**

Bij het invullen van de checklist konden docenten steeds kiezen uit drie antwoordcategorieën, waaraan respectievelijk de numerieke scores 1, 2 en 3 werden

toegekend. Met behulp van een homogeniteitsanalyse is onderzocht of deze scoretoekenning in de praktijk zinvol was.

Bij homogeniteitsanalyse wordt per item naar de meest optimale scoretoekenning gezocht, en wel zo dat leerlingen met een vergelijkbaar antwoordpatroon op de test zoveel mogelijk hetzelfde gewicht op de items krijgen, en leerlingen met een verschillend antwoordpatroon juist een verschillend gewicht. Uit de homogeniteitsanalyse bleek dat het voor 11 van de 35 items beter zou zijn om de eerste twee categorieën samen te nemen. Deze items zijn in tabel 3.1 van een kruisje voorzien. Voor de overige items geldt dat de scoretoekenning van 1, 2 en 3 punten aan de drie categorieën goed gekozen is.

### 3.1.6 Beoordelaarsovereenstemming

Om te kijken in hoeverre het voor de beoordeling van de studievoordigheid van een leerling uitmaakt door *welke* vakdocent de checklist wordt ingevuld, is per item onderzocht in welke mate docenten van de vakken Nederlands, Engels en wiskunde tot eenzelfde beoordeling van hun leerlingen komen. Voor elk item zijn op basis van een tweeweg-random-effect ANOVA (personen x beoordelaars) variantiecomponenten geschat. Deze variantiecomponenten werden vervolgens gebruikt om de beoordelaarsovereenstemming te schatten. Deze maat,  $\hat{\rho}^2$ , is gedefinieerd als:

$$\hat{\rho}^2 = \frac{\hat{\sigma}_p^2}{\hat{\sigma}_p^2 + (\hat{\sigma}_b^2 + \hat{\sigma}_{res}^2) / k},$$

waarbij  $k$  staat voor het aantal beoordelaars, en  $\hat{\sigma}_p^2$ ,  $\hat{\sigma}_b^2$  en  $\hat{\sigma}_{res}^2$  voor de variantiecomponenten van respectievelijk de beoordeelde personen, de beoordelaars, en het residu. De overeenstemmingscoëfficiënt loopt van 0 (geen overeenstemming) tot 1 (maximale overeenstemming).

Voor het berekenen van de mate van overeenstemming zijn alleen die leerlingen geselecteerd die door alle drie de vakdocenten waren beoordeeld. Van deze selectie

zijn vervolgens alleen die leerlingen gebruikt voor wie alle items waren ingevuld. Op basis van een volledige matrix van 180 leerlingen bij 3x35 items werd het onderzoek naar beoordelaarsovereenstemming verricht. Hierbij bleek dat de mate van overeenstemming voor alle items uit de checklist laag was.

Om te onderzoeken wat de oorzaak zou kunnen zijn voor het feit dat verschillende vakdocenten zo weinig overeenstemmen is een aantal secundaire analyses uitgevoerd.

Bij de boven beschreven analyse werden de docenten als zogeheten 'generieke beoordelaars' beschouwd, dat wil zeggen dat alle docenten binnen een vak samen als één docent van het betreffende vak werden beschouwd. In werkelijkheid werden de 180 leerlingen echter niet door één enkele docent Nederlands beoordeeld, maar hebben *verschillende* docenten Nederlands hun oordeel over de leerlingen gegeven.

Om de invloed van de gehanteerde opzet op de uitkomsten van de analyse te onderzoeken, is het gegevensbestand opgesplitst in groepjes leerlingen die allemaal door *dezelfde* drie docenten Nederlands, Engels en wiskunde waren beoordeeld. Vervolgens is voor tien van deze groepjes apart een beoordelaarsovereenstemmingsonderzoek verricht. Deze tien analyses leverden vrijwel dezelfde resultaten op als de eerste analyse.

De docenten kunnen ook een gebrek aan overeenstemming vertonen, omdat de leerlingen per vak een verschillend niveau van studievoordigheid bezitten. Om hier een uitspraak over te kunnen doen werden van de 180 leerlingen diegenen geselecteerd die *op alle drie de vakken* in 3 havo of 3 vwo een eindrapportcijfer van 7 of hoger hadden behaald, onder de aanname dat leerlingen die hoge cijfers halen studievoordiger zijn dan leerlingen die lage cijfers behalen. Vervolgens is onderzocht in hoeverre docenten van verschillende vakken het met elkaar eens zijn als het gaat om het beoordelen van de studievoordigheid van deze leerlingen. Ook deze analyse, waarbij de docenten opnieuw als 'generieke beoordelaars' werden beschouwd, leverde vrijwel dezelfde resultaten op: de drie 'soorten' beoordelaars bleken het ook bij deze leerlingen lang niet altijd met elkaar eens te zijn.

In tabel 3.1 staat bij elk item de waarde van de overeenstemmingscoëfficiënt vermeld zoals deze in het eerstgenoemde onderzoek werd gevonden. De grootste mate van overeenstemming werd bereikt bij item 2 ( $\hat{\rho}^2 = 0.74$ ) uit het prototype van de checklist. Item 30 leverde de laagste overeenstemmingscoëfficiënt op ( $\hat{\rho}^2 = 0.08$ ). De gemiddelde waarde van de coëfficiënt bedraagt 0.46.

Op grond van de boven besproken criteria is bepaald welke items uit het prototype van de checklist verwijderd moesten worden. Hierbij is de vuistregel gehanteerd dat items die aan meer dan twee van de gespecificeerde criteria niet voldeden, niet in de definitieve versie van de checklist zouden worden opgenomen. In de laatste kolom van tabel 3.1 wordt aangegeven voor welke items dit geldt. De definitieve versie van de checklist is opgenomen in bijlage 2.

### **3.2 Kwaliteit van de checklist**

De definitieve versie van de checklist studievvaardigheid bestaat uit 25 items. De betrouwbaarheid van de checklist is hoog: Cronbach's alpha bedraagt 0.944 in het geval het vak van de docenten buiten beschouwing wordt gelaten, en 0.943, 0.946 en 0.943 indien alpha voor Nederlands, Engels en wiskunde apart wordt berekend. De standaardmeetfout van de score op de checklist is zowel voor de gehele groep, als apart voor de vakken Nederlands, Engels en wiskunde telkens 2.5.

De totaalscore op de checklist wordt berekend door de som te nemen van de scores op de afzonderlijke items. Het bereik van de totaalscore ligt daarmee tussen de 25 en 75. Voor de huidige steekproef bedraagt de gemiddelde totaalscore op de checklist 57.3 met een spreiding van 10.7 (zie tabel 3.3). Dit gemiddelde is berekend op basis van de beoordeling van die leerlingen waarvoor alle items in de checklist zijn ingevuld.

Bij de proefafname zijn leerlingen door één, twee of drie docenten beoordeeld. Een belangrijke vraag bij de implementatie van de checklist is of het nodig is dat de studievvaardigheid van een leerling door meerdere docenten beoordeeld wordt. Uit het onderzoek naar de mate van beoordelaarsovereenstemming op itemniveau,

besproken in paragraaf 3.1, bleek al dat docenten veelal tot verschillende itemscores kwamen bij dezelfde leerlingen. Nadat de definitieve versie van de checklist was vastgesteld, werd de beoordelaarsovereenstemming nogmaals onderzocht, maar nu met betrekking tot de totaalscore op de checklist. Voor de groep van 180 leerlingen werd onderzocht of docenten van de vakken Nederlands, Engels en wiskunde tot eenzelfde totaalscore op de checklist komen, waarbij deze totaalscore werd bepaald op grond van de scores op de 25 uiteindelijk geselecteerde items. De overeenstemming tussen docenten voor wat betreft de totaalscore op de checklist bleek opnieuw laag te zijn ( $\hat{\rho}^2 = 0.626$ ). Omdat de overeenstemming al laag was bij drie beoordelaars, wordt geen schatting meer gegeven voor het geval er twee of slechts één beoordelaar de checklist in zou vullen.

Ten behoeve van het validiteitsonderzoek van de definitieve versie van de checklist is aan de docenten gevraagd om voor elke leerling de volgende informatie te verstrekken: het eindcijfer van het schooljaar '93-'94 op het betreffende vak, en een globaal oordeel van de docent over de te verwachten prestaties van de leerling in het schooljaar '94-'95. Verder is voor elke leerling de vaardigheid in een bepaald vak geschat op basis van de prestaties op de voor Nederlands, Engels en wiskunde in de proefafname aangeboden opgaven.

Op grond van deze gegevens is voor ieder vak de correlatie tussen de totaalscore op de checklist en het cijfer op het eindrapport van het schooljaar '93-'94 voor het betreffende vak, en tussen de totaalscore en de geschatte vaardigheid op het betreffende vak berekend. Ook is het verband bepaald tussen de totaalscore op de checklist en het globale oordeel van docenten over de prestaties van leerlingen in het schooljaar '94-'95. Deze correlaties worden voor elk van de drie vakken in tabel 3.2 vermeld.

Tabel 3.2

Correlaties tussen de totaalscore op de checklist en enkele achtergrondgegevens  
(Tussen haakjes het aantal observaties op grond waarvan de correlatie berekend is)

|            | Correlatie tussen totaalscore op de checklist en: |                                |                          |
|------------|---------------------------------------------------|--------------------------------|--------------------------|
|            | Eindcijfer<br>leerjaar '93-'94                    | Prestaties<br>leerjaar '94-'95 | Geschatte<br>Vaardigheid |
| Nederlands | .561 (396)                                        | .544 (410)                     | .215 (694)               |
| Engels     | .551 (701)                                        | .471 (755)                     | .288 (1102)              |
| Wiskunde   | .574 (756)                                        | .525 (780)                     | .346 (858)               |

### 3.3 Verschillen tussen havo- en vwo-leerlingen

De checklist is ingevuld voor leerlingen uit 3 havo, 3 havo/vwo en 3 vwo. Een voor de hand liggende vraag is of er met betrekking tot studievoardigheid, zoals deze met de checklist gemeten is, verschillen zijn te ontdekken tussen deze drie groepen. Om deze vraag te beantwoorden is voor elk van de drie klastypen de gemiddelde totaalscore op de checklist berekend. In tabel 3.3 wordt de gemiddelde score op de checklist weergegeven, voor de gehele groep, per klastype en per vak.

Tabel 3.3

Gemiddelde score op de checklist, voor de groep als geheel, en uitgesplitst naar de vakken Nederlands, Engels en wiskunde en de klastypen havo, vwo en havo/vwo

|                  | Klastype |      |          | Totaal |
|------------------|----------|------|----------|--------|
|                  | havo     | vwo  | havo/vwo |        |
| Nederlands       | 54.8     | 59.9 | 57.9     | 57.9   |
| Engels           | 56.1     | 60.3 | 57.9     | 58.4   |
| Wiskunde         | 54.1     | 56.4 | 56.0     | 55.5   |
| Groep als geheel | 55.2     | 58.9 | 57.3     | 57.3   |

Uit tabel 3.3 valt op te maken dat havo-leerlingen over het algemeen als iets minder studievaardig worden beoordeeld dan vwo-leerlingen. Havo/vwo-leerlingen liggen daar tussenin. Verder blijkt dat de gemiddelde scores bij docenten van verschillende vakken enigszins afwijken. Een variantie-analyse uitgevoerd op de afhankelijke variabele 'totaalscore op de checklist' met de verschillende vakken en het klastype als onafhankelijke variabelen, leverde zowel een hoofdeffect voor de variabele 'soort vak' ( $F=21.89$ ,  $df=2$ ,  $p=0.000$ ) als een hoofdeffect voor de variabele 'klastype' ( $F=41.43$ ,  $df=2$ ,  $p=0.000$ ) op. Er trad geen interactie-effect op. Naar aanleiding van de hoofdeffecten werden nog twee analyses verricht. Die toonden aan dat het verschil tussen de (totaalscores zoals bepaald door docenten van de) vakken Nederlands en wiskunde en het verschil tussen de vakken Engels en wiskunde statistisch significant is. Verder bleken alle gevonden verschillen voor wat betreft de totaalscore op de checklist van leerlingen van verschillende klastypen statistisch significant te zijn. (Bij alle bovenstaande resultaten werd  $\alpha=0.05$  gehanteerd).

### 3.4 Normtabellen

Op basis van de in dit onderzoek verzamelde gegevens zijn normtabellen opgesteld die bedoeld zijn om als referentie te dienen voor docenten die de checklist willen gebruiken om er de studievaardigheid van hun leerlingen mee te beoordelen. De totaalscore op de checklist van een bepaalde leerling kan op deze manier afgezet worden tegen de scores van een landelijke groep vergelijkbare leerlingen.

Tabel 3.4 bevat normtabellen die op de volgende manier zijn bepaald: de scores van leerlingen op de checklist zijn van laag naar hoog geordend en vervolgens in tien groepen (decielen) ondergebracht die elk tien procent van de scores bevatten. Scoregroep 1 bevat de tien procent laagste scores, scoregroep 2 de tien procent scores die daarop volgen, enzovoort. Scoregroep 10 bevat dus de tien procent hoogste scores. Tabel 3.4 geeft per vak voor zowel de leerlingen uit 3 havo als de leerlingen uit 3 vwo aan welke scores in welke scoregroepen horen. Met behulp van de tabel kan bepaald worden in welke scoregroep de score van een bepaalde leerling valt, dus hoe studievaardig deze leerling is ten opzichte van een landelijke groep vergelijkbare leerlingen. Valt de score van een leerling bijvoorbeeld in scoregroep 1, dan heeft zeker negentig procent van de leerlingen een hogere score op de checklist

behaald. Heeft een leerling een score die in scoregroep 10 valt, dan hoort zijn of haar score bij de tien procent hoogste scores die behaald zijn.

Tabel 3.4  
Normtabellen voor scores op de checklist studievaardigheid

| Score-<br>groep | Nederlands |       |        | Score-<br>groep | Engels |       |        | Score-<br>groep | Wiskunde |       |        |
|-----------------|------------|-------|--------|-----------------|--------|-------|--------|-----------------|----------|-------|--------|
|                 | Havo       | Vwo   | Totaal |                 | Havo   | Vwo   | Totaal |                 | Havo     | Vwo   | Totaal |
| 1               | 25-42      | 25-45 | 25-43  | 1               | 25-41  | 25-47 | 25-46  | 1               | 25-40    | 25-43 | 25-39  |
| 2               | 43-46      | 46-51 | 44-50  | 2               | 42-46  | 48-51 | 47-49  | 2               | 41-45    | 44-47 | 40-46  |
| 3               | 47-49      | 52-54 | 51-52  | 3               | 47-50  | 52-54 | 50-53  | 3               | 46-48    | 48-50 | 47-49  |
| 4               | 50-51      | 55-58 | 53-54  | 4               | 51-52  | 55-58 | 54-56  | 4               | 49-51    | 51-53 | 50-54  |
| 5               | 52-54      | 59-60 | 55-57  | 5               | 53-56  | 59-61 | 57-59  | 5               | 52-54    | 54-56 | 55-57  |
| 6               | 55-58      | 61-64 | 58-61  | 6               | 57-59  | 62-65 | 60-61  | 6               | 55-57    | 57-59 | 58-60  |
| 7               | 59-61      | 65-67 | 62-63  | 7               | 60-62  | 66-68 | 62-64  | 7               | 58-60    | 60-63 | 61-63  |
| 8               | 62-64      | 68-71 | 64-68  | 8               | 63-68  | 69-71 | 65-67  | 8               | 61-64    | 64-67 | 64-66  |
| 9               | 65-68      | 72-73 | 69-71  | 9               | 69-72  | 72-74 | 68-70  | 9               | 65-68    | 68-71 | 67-72  |
| 10              | 69-75      | 74-75 | 72-75  | 10              | 73-75  | 75    | 71-75  | 10              | 69-75    | 72-75 | 73-75  |



## 4 Discussie en conclusies

De vraag die in dit onderzoek centraal stond was of de checklist studievoordigheid een geschikt instrument is om door docenten gebruikt te worden om de studievoordigheid van hun leerlingen mee te beoordelen.

Voor het beantwoorden van deze vraag is allereerst onderzocht welke van de 35 items uit het prototype van de checklist van voldoende kwaliteit waren om in de definitieve versie van de checklist te worden opgenomen. Op basis van het commentaar van docenten en de resultaten van verschillende analyses zijn in totaal 10 van de 35 items uit het prototype van de checklist verwijderd.

Een aantal van de verwijderde items had betrekking op psychologische eigenschappen van leerlingen, zoals faalangst, stressgevoeligheid en zelfvertrouwen. Wat bij de analyses opviel was dat docenten bij deze items de neiging hadden om leerlingen die op de overige items in de regel laag scoorden een hogere score toe te kennen dan verwacht zou mogen worden op grond van de scores op de andere items. Wellicht zijn docenten bij het beoordelen van minder studievoordige leerlingen wat voorzichtiger geweest met het geven van een negatief oordeel, omdat de betreffende items betrekking hadden op eigenschappen van leerlingen waar de docent weinig zicht op heeft. Deze bevinding is geheel in lijn met het commentaar van docenten die de checklist hebben ingevuld, namelijk dat er bij die items te veel psychologisch inzicht van hen wordt gevraagd.

Nadat de definitieve versie van de checklist was vastgesteld is onderzocht of de 25 items in de checklist samen een betrouwbaar en valide instrument vormen. De betrouwbaarheid van de checklist bleek bijzonder hoog te zijn, zowel voor de vakken afzonderlijk als wanneer er geen onderscheid gemaakt werd tussen vakdocenten. Het ziet er dan ook naar uit dat de items in de checklist alle betrekking hebben op één en hetzelfde construct.

De validiteit van de checklist is onderzocht door de mate van samenhang te bepalen tussen de totaalscore op de checklist en drie verschillende operationalisering van studiesucces.

Er bleek een redelijk positief verband tussen de hoogte van de score op de checklist voor een bepaald vak en het cijfer op het eindrapport van het betreffende vak. Leerlingen met hogere scores op de checklist bleken in de regel ook hogere cijfers op het eindrapport van het schooljaar '93-'94 voor het betreffende vak te behalen. Hetzelfde gold voor het verband met het oordeel van docenten over de prestaties van leerlingen in het schooljaar '94-'95. Leerlingen die volgens docenten slechter zullen presteren scoren in de regel lager op de checklist dan de 'gemiddelde' leerlingen. De 'gemiddelde' leerlingen scoren over het algemeen ook weer lager dan de 'betere' leerlingen'.

Tenslotte is ook het verband onderzocht tussen scores op de checklist voor een bepaald vak en de prestaties van leerlingen op het opgavenmateriaal dat hen voorgelegd is in het kader van de proefafname. De correlaties tussen de vaardigheidsschattingen van leerlingen en hun scores op de checklist bleken vrij laag te zijn. Dit is echter niet onwenselijk. Zouden de checklistscores sterk samenhangen met prestaties op de plaatsingstoetsen, dan zou de checklist weinig kunnen bijdragen aan de voorspellende waarde van de instrumenten als geheel. De checklist kan nu naast de plaatsingstoetsen een rol spelen bij het voorspellen van het toekomstige prestatieniveau van leerlingen.

Aangezien de checklist niet door de leerlingen zelf, maar door docenten is ingevuld, is ernaar gestreefd om alleen die items in de checklist op te nemen die betrekking hebben op eigenschappen en vaardigheden die door docenten beoordeeld kunnen worden. Nu is het bij beoordelingen in het algemeen zo dat men zich af kan en moet vragen in hoeverre het uitmaakt door welke persoon die beoordeling gegeven wordt. Zoals reeds eerder aangegeven, wordt het onderzoek naar beoordelaarsovereenstemming in het geval van de checklist echter door twee aspecten bemoeilijkt. Ten eerste is het zo dat een docent een leerling voor het invullen van de checklist moet kennen. Daarom is het niet mogelijk om, zoals bij onderzoek naar de mate van objectiviteit van correctievoorschriften bij opstellen of proefwerken nog wel eens wordt gedaan, aan een aantal verschillende docenten van hetzelfde vak te vragen of zij hun mening ten aanzien van de studievoordigheid van een leerling kunnen geven.

Het is ook niet zinnig om meerdere docenten van dezelfde school om hun oordeel ten aanzien van de studievoordigheid van een leerling te vragen, aangezien elk van deze docenten de leerling vanuit zijn eigen kader, zijn vakgebied, zal beoordelen. Indien de docenten van mening verschillen kan men in het laatste geval niet zeggen of dit wordt veroorzaakt door een verschil in persoonlijke visie of doordat leerlingen zich binnen verschillende vakken anders gedragen.

Binnen dit onderzoek is ervoor gekozen om de overeenstemming te onderzoeken tussen de docenten van de vakken Nederlands, Engels en wiskunde. Indien namelijk zou blijken dat de docenten het met elkaar eens zijn, zouden we immers mogen concluderen dat het voor de beoordeling van de studievoordigheid niet uitmaakt door *welke docent* van *welk vak* de checklist wordt ingevuld. Uit de resultaten van het overeenstemmingsonderzoek blijkt echter dat docenten van verschillende vakken het niet met elkaar eens zijn als het gaat om het beoordelen van de studievoordigheid van hun leerlingen. De belangrijkste conclusie die op basis van dit onderzoek getrokken kan worden is dat scholen bij het bepalen van de studievoordigheid van een leerling met behulp van de checklist niet af kunnen gaan op het oordeel van een enkele docent. Het maakt blijkbaar uit welke vakdocent de checklist invult. Het is niet vast te stellen wat daar de oorzaak van is. Docenten kunnen tot een verschillende score komen bij een identieke leerling omdat de studievoordigheid van deze leerling verschilt voor de vakken waarin zij lesgeven, maar ook omdat zij vanuit hun persoonlijke visie niet tot een eenduidig oordeel kunnen komen, terwijl de studievoordigheid van de leerling voor beide vakken niet verschilt.

Het feit dat verschillende docenten tot andere scores op de checklist kunnen komen voor een identieke leerling is niet verontrustend. Zo geven Van Dyck et al. (1991) aan dat studievoordigheid per vak of vakgebied kan verschillen. Voor de voorspellende waarde van plaatsingstoetsen en de checklist is dit resultaat zelfs gunstig. Het gebruik van de checklist door docenten Nederlands, Engels en wiskunde leidt tot drie onafhankelijke betrouwbare scores voor een leerling. Elk van die scores geeft aan hoe het gesteld is met de studievoordigheid van deze leerling binnen de betreffende vakken. De drie scores leveren alle een unieke bijdrage aan de voorspelling van het toekomstig prestatieniveau van leerlingen zodat iedere checklistscore kan worden beschouwd als een aparte predictor.

In vervolgonderzoek zal nader moeten worden onderzocht in hoeverre de drie checklistscores in combinatie met de scores op de plaatsingstoetsen, in staat zijn om het toekomstig prestatieniveau van leerlingen te voorspellen.

## Literatuur

- Dyck, M. van, Leijh, G. & Baks, R.A.J. (1991). *Studievaardigheid en de kwalificatie voor HBO en WO*. De Lier: Academisch Boeken Centrum.
- Eggen, T.J.H.M. & Sanders, P.F. (1993). *Psychometrie in de praktijk*. Arnhem: Cito.
- Elshout-Mohr, M. & Van Daalen-Kapteijns, M. (1985a). Het leren van begrippen, in het bijzonder in het eerste stadium van het hoger onderwijs. *Pedagogische Studiën*, 62, 11, pp 459-470.
- Elshout-Mohr, M. & Van Daalen-Kapteijns, M. (1985b). Van volgen van onderwijs naar zelfstandig studeren. In: Lodewijks, J.G.L.E. & P.R.J. Simons, (red.). *Zelfstandig leren*. Lisse: Swets & Zeitlinger.
- Elshout-Mohr, M. & Van Daalen-Kapteijns, M. (1986). Studievaardigheidstraining: Effect en waardering. *Onderzoek van Onderwijs*, 4, pp 64-67.
- Gerrits, G., Gielis, P., Hutjes, L., Rademakers, R. & Smits, J. (1988). *Studielessen: een handleiding*. Nijmegen: Berkhout.
- Hogeschool Groningen (1992). *Studie-ondersteuning propaedeuse HES-vt handleiding voor studenten cursusjaar 1991/1992*.
- Hogeschool Eindhoven (1989). *Vragenlijst propaedeuse zelfonderzoek*.
- Hogeschool Enschede (1992). *Leerwijzer*. Sector techniek.
- Ministerie van Onderwijs en Wetenschappen (1991). *Profiel van de tweede fase voortgezet onderwijs, voorstellen voor een betere toerusting van scholen en leerlingen in de tweede fase voortgezet onderwijs*. Zoetermeer.
- Ministerie van Onderwijs en Wetenschappen (1992). *Vervolgnota, 'Profiel van de tweede fase voortgezet onderwijs'*. Zoetermeer.
- Neuwahl, N.M.E., & Van den Bogaart, P.H.M. (1988). *VHVO, vragenlijst huiswerk voortgezet onderwijs*. Lisse: Zwets en Zeitlinger.
- Oosterhuis-Geers, J. (1992). *Diagnostische vragenlijst studiemethoden*. Onderwijskundig Centrum universiteit Twente.
- Pintrich et al. (1991). *MSLQ: A manual for the use of the motivated strategies for learning questionnaire*.
- Simons, R.J. (1987a). Leren zelfstandig te studeren. *Tijdschrift voor Hoger Onderwijs*, 5, 2, pp 6-11.

- Simons, R.J. (1987b). Bevordering van zelfstandig leervermogen. In: Halkes, R. & B. van Hoek (red.). *Basisvorming in onderwijskundige optiek*. Lisse: Swets & Zeitlinger.
- Simons, R.J. (1989). Learning to learn. In: Span, P. (red.). *Onderwijsleerprocessen*. Lisse: Swets & Zeitlinger.
- Simons, R.J. & Vermunt, J.D.H.M. (1986). Self regulation in knowledge acquisition: a selection of Dutch research. In: Beukhof, G. & R.J. Simons (red.). *German and Dutch research on learning and instruction*. 's Gravenhage: SVO.
- Sluijter, C. (red.) (in voorbereiding). *Verantwoording plaatsingstoetsen voor 3 havo en 3 vwo*. OPD-memorandum; Arnhem: Cito.
- Stuurgroep Profiel Tweede Fase Voortgezet Onderwijs (1994). *Tweede Fase: scharnier tussen basisvorming en hoger onderwijs: een uitwerking op hoofdlijnen van de nota's profiel van de tweede fase van het voortgezet onderwijs*. 's Gravenhage.
- TH Eindhoven (1987). *Studeergedragvragenlijst, rapport 27*. Bureau Studentenpsychologen.
- Verhelst, N.D. & Eggen, T.H.J.M. (1989). *Psychometrische en statistische aspecten van peilingsonderzoek*. PPON-rapport nr 4. Arnhem: Cito.
- Verhelst, N.D., Glas, C.A.W. & Verstralen, H.H.F.M. (1995). *OPLM: One-parameter logistic model: computer program and manual*. Arnhem: Cito.
- Verhelst, N.D. & Glas, C.A.W. (1995). The One Parameter Logistic Model. In: G.H. Fischer & I.W. Molenaar (Eds.). *Rasch Models: Foundations, Recent Developments, and Applications*. (pp. 215-239). New York: Springer.
- Willems, J.M.H.N. (1987). *Studietaken als instruktie-middel: een onderzoek naar de sturende werking van opdrachten bij studieteksten vanuit een informatietheoretisch kader*. Nijmegen: proefschrift KU.

## **Bijlage 1: Items in het prototype van de checklist studievaardigheid**

### **1 Huiswerk af**

U beoordeelt of de leerling het opgedragen huiswerk in de regel daadwerkelijk af heeft (stof geleerd, lesmateriaal doorgenomen, schriftelijk werk gemaakt). Is dat het geval, dan geeft u een positieve beoordeling (+). Laat de leerling hier in uw ogen te vaak steken vallen, dan levert dit een negatieve beoordeling (-).

### **2 Huiswerk netheid**

U beoordeelt of het schriftelijk huiswerk van de leerling er in de regel netjes verzorgd uitziet. Is dat het geval, dan geeft u een positieve beoordeling (+). Is er volgens u veelvuldig sprake van kliederen, kladderen, rommelig werken, enzovoort, dan leidt dat tot een negatieve beoordeling (-).

### **3 Motivatie**

U beoordeelt of de leerling naar uw indruk in het algemeen open staat voor wat in de lessen aan de orde komt, wat onder andere tot uitdrukking komt in actieve deelname aan de les. Is dat het geval, dan geeft u een positieve beoordeling (+). Is de leerling in de regel ongeïnteresseerd en passief, dan geeft u een negatieve beoordeling (-).

### **4 Opletten in de klas**

U beoordeelt of de leerling naar uw indruk in de lessen goed oplet, aandachtig is en niet snel is afgeleid (naar buiten kijken, kletsen, keten, etc.). Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Moet u een leerling te vaak 'bij de les halen', dan geeft u een negatieve beoordeling (-).

### **5 Gedrag in de klas**

U beoordeelt of de leerling zich in de lessen 'normaal' gedraagt. Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Sterk afwijkend (in de zin van storend) gedrag leidt tot een negatieve beoordeling (-).

### **6 Cijfers S.O.'s (hoogte)**

U beoordeelt of de leerling voor schriftelijke (en mondelinge) overhoringen voldoende cijfers haalt (cijfer 6 of beter). Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Regelmatig onvoldoende S.O. cijfers leidt tot een negatieve beoordeling (-).

### **7 Cijfers S.O.'s (constantheid)**

U beoordeelt of de leerling S.O. cijfers haalt die in hoogte vrij stabiel zijn (positief oordeel, +) of juist sterk fluctueren (negatief oordeel, -).

### **8 Cijfers Proefwerken (hoogte)**

U beoordeelt of de leerling voor proefwerken overwegend voldoende cijfers haalt (cijfer 6 of beter). Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Regelmatig onvoldoende proefwerk cijfers leidt tot een negatieve beoordeling (-).

### **9 Cijfers Proefwerken (constantheid)**

U beoordeelt of de leerling proefwerk cijfers haalt die in hoogte vrij stabiel (positief oordeel, +) zijn of sterk fluctueren (negatief oordeel, -).

### **10 Zelfvertrouwen**

U beoordeelt of de leerling naar uw indruk in de lessen en bij S.O.'s en proefwerken blijkt geeft van een voldoende mate van zelfvertrouwen. Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Een gebrek aan zelfvertrouwen (onzekerheid) of niet gerechtvaardigd overmatig zelfvertrouwen leidt tot een negatieve beoordeling (-).

### **11 Punctualiteit**

U beoordeelt of de leerling naar uw indruk in zijn / haar werk en lesvoorbereiding voldoende zorgvuldig is, netjes werkt, nauwkeurig is en zich aan afspraken houdt. Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Regelmatig voorkomende slordigheid, nonchalance en dergelijke, leidt tot een negatieve beoordeling (-).

## **12 Omgang met medeleerlingen**

U beoordeelt of de leerling naar uw indruk op een normale manier omgaat met medeleerlingen. Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Zich afsluiten of zich dominant manifesteren, ruzie of het optreden van fricties leidt tot een negatieve beoordeling (-).

## **13 Omgang met docenten**

U beoordeelt of de leerling naar uw indruk op een normale manier met docenten omgaat. Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Is er regelmatig sprake van gedrag dat afkeurende of corrigerende reacties van docenten oproept dan geeft u een negatieve beoordeling (-).

## **14 Spreekvaardigheid (m.b.t. nederlandse taal)**

U beoordeelt of de leerling naar uw indruk met spreektaken moeite heeft (ideeën naar voren brengen, uit zijn/haar woorden komen, argumenteren, vragen stellen). Zijn er geen problemen op dit vlak dan geeft u een positieve beoordeling (+). Heeft de leerling regelmatig problemen met spreektaken dan geeft u een negatieve beoordeling (-).

## **15 Leesvaardigheid (m.b.t. nederlandse taal)**

U beoordeelt of de leerling naar uw indruk in uw lessen met leestaken moeite heeft (technisch lezen, begrijpend lezen, en dergelijke). Zijn er geen problemen op dit vlak dan geeft u een positieve beoordeling (+). Heeft de leerling regelmatig problemen met leestaken dan geeft u een negatieve beoordeling (-).

## **16 Lesvoorbereiding**

U beoordeelt of de leerling naar uw indruk in de lessen voldoende voorbereid verschijnt (leerwerk geleerd en gekend, schriftelijk werk af, etc.). Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Zijn er regelmatig problemen op dit vlak dan geeft u een negatieve beoordeling (-).

### **17 Absentie**

U beoordeelt of de leerling naar uw indruk de lessen niet vaker dan normaal verzuimt. Is dat het geval dan geeft u een positieve beoordeling (+). Onder 'normaal' verstaan we incidenteel verzuim, bijvoorbeeld door een griepje. Frequent verzuim en ongeoorloofde afwezigheid (spijbelen) leidt in elk geval tot een negatieve beoordeling (-).

### **18 Zelfstandig werken**

U beoordeelt of de leerling naar uw indruk laat zien in de lessen in staat te zijn tot zelfstandig werken. Is dat het geval, dan geeft u een positieve beoordeling (+). Overmatige afhankelijkheid van initiatieven of sturing door anderen leidt tot een negatieve beoordeling (-).

### **19 Samenwerken**

U beoordeelt of de leerling naar uw indruk in de lessen laat zien in staat te zijn tot effectieve samenwerking met andere leerlingen, bijvoorbeeld bij groepswork. Is dat in de regel het geval, dan geeft u een positieve beoordeling (+). Aanpassingsmoeilijkheden, dominant gedrag, ruzie, fricties en disfunctioneren in samenwerkingsrelaties leiden tot een negatieve beoordeling (-).

### **20 Schrijfvaardigheid (m.b.t. nederlandse taal)**

U beoordeelt of de leerling naar uw indruk in de lessen laat zien in staat te zijn tot het correct uitvoeren van schrijftaken (een samenvatting maken, een opstel schrijven, een vertaling geven, het beschrijven van de uitkomsten van een opdracht, etc.). Is dat het geval, dan geeft u een positieve beoordeling (+). Zijn er regelmatig problemen op dit vlak (zinsbouw, tekststructuur, verhaal-lijn, logica, helderheid, etc.) dan geeft u een negatieve beoordeling (-).

### **21 Uitdrukkingsvaardigheid (m.b.t. nederlandse taal)**

U beoordeelt of de leerling naar uw indruk in de lessen laat zien zich mondeling helder te kunnen uitdrukken. Is dat het geval, dan geeft u een positieve beoordeling (+). Zijn er regelmatig problemen op dit vlak (moeite met het overbrengen van wat bedoeld wordt) dan geeft u een negatieve beoordeling (-).

## **22 Vragen (durven) stellen wanneer iets niet begrepen wordt**

U beoordeelt of de leerling naar uw indruk in de lessen **waar nodig** vragen stelt indien iets niet begrepen wordt. U geeft een positieve beoordeling (+) indien een leerling geen moeite heeft om met een vraag voor de dag te komen.

Alleen indien u duidelijke aanwijzingen heeft dat de leerling geen vragen stelt terwijl iets niet wordt begrepen geeft u een negatieve beoordeling (-).

## **23 Van buiten leren**

U beoordeelt of de leerling naar uw indruk in staat is tot adequaat memoriseren van lesstof-onderdelen (woordjes, begrippen, formules, regels, termen, relaties, ideeën).

U geeft een positieve beoordeling (+) indien uit het hoofd leren geen problemen geeft. U geeft een negatieve beoordeling (-) indien uit het hoofd leren **regelmatig** problemen geeft en/of het geleerde vaak weer snel vergeten blijkt.

## **24 Nauwkeurig werken**

U beoordeelt of de leerling naar uw indruk in de lessen te zien geeft geneigd te zijn tot nauwkeurig en precies werken. Is dat het geval, dan geeft u een positieve beoordeling (+). Slordigheid, knoeien en kliederwerk in die mate dat het fouten veroorzaakt levert een negatieve beoordeling op (-).

## **25 Begrijpen opgedragen taken**

U beoordeelt of de leerling naar uw indruk in de lessen opgedragen taken vlot begrijpt. Is dat het geval, dan geeft u een positieve beoordeling (+). Is uitvoerige of herhaalde uitleg ineermaals nodig dan leidt dit tot een negatieve beoordeling (-).

## **26 Aandacht in de les**

U beoordeelt of de leerling in de lessen normaal aandachtig is. Is dat het geval, dan geeft u een positieve beoordeling (+). Moet er regelmatig worden gewaarschuwd, of moeten er maatregelen worden genomen om de leerling 'bij de les te houden', dan levert dat een negatieve beoordeling op (-).

## **27 Zelfcontrole**

U beoordeelt of de leerling geneigd is tot een normale mate van controle op eigen werk. Is dat het geval, dan geeft u een positieve beoordeling (+). Snel tevreden zijn, 'het wel geloven', oppervlakkigheid en slordigheid, in die mate dat daardoor fouten ontstaan, leiden tot een negatieve beoordeling (-).

## **28 Verstandelijke capaciteit**

U beoordeelt het algemene intellectuele vermogen van de leerling. Het vermogen tot onthouden (feiten, begrippen, regels, formules), tot het leggen van relaties, het verwerven van inzicht, het vermogen tot transfer (geleerde kunnen toepassen en kunnen vertalen naar andere situaties). Zijn er, **gerelateerd aan het opleidingstype**, geen bijzonderheden op dit vlak, dan geeft u een positieve beoordeling (+), bij twijfel een middelmatige beoordeling ( $\square$ ). Beoordeelt u de verstandelijke capaciteit als onvoldoende (bijvoorbeeld op basis van u bekende serieuze leer- of studieproblemen) dan geeft u een negatieve beoordeling (-).

## **29 Theoretische interesse**

U beoordeelt of de leerling een duidelijke gevoel ('feeling') heeft voor de meer theoretische (abstracte, intellectuele) aspecten van de opleiding. Dit in contrast met een uitsluitend praktisch ingestelde instrumentele belangstelling. Een duidelijk theoretische interesse duidt u aan met een positieve beoordeling (+). Een leerling die duidelijk niet-theoretisch is aangelegd geeft u een negatieve beoordeling (-).

## **30 Incasseringsvermogen**

U beoordeelt of de leerling zich te vaak of bovenmatig druk maakt over tegenvallers of extra werkdruk (een laag cijfer, een terechtwijzing, een berisping, een fout antwoord, aankondiging van een proefwerk, etc.). U geeft een negatieve beoordeling (-) indien de leerling door dit soort situaties uit balans raakt. U geeft een positieve beoordeling (+) indien deze situaties zich niet voordoen of wanneer de leerling door dit soort situaties niet overmatig wordt belemmerd in het normale functioneren.

### **31 Doorzettingsvermogen**

U beoordeelt of de leerling geneigd is flink te willen aanpakken, niet snel het bijltje erbij neergooit, niet bij de pakken neer zit indien lastige of omvangrijke stof (evt. huiswerk) wordt aangeboden. Is er geen enkel probleem op dit vlak, dan geeft u een positieve beoordeling (+). Schiet de leerling in dit opzicht regelmatig tekort, dan geeft u een negatieve beoordeling (-).

### **32 Concentratievermogen**

U beoordeelt of de leerling zich niet bovenmatig laat afleiden door wat er om hem/haar heen gebeurt. Kan een leerling zich regelmatig slecht concentreren (snel afgeleid) dan geeft u een negatieve beoordeling (-). Is de leerling doorgaans onverstoorbaar en goed geconcentreerd, dan geeft u een positieve beoordeling (+).

### **33 Faalangst**

U beoordeelt of de leerling niet bovenmatig (negatief) faalangstig is, dichtklapt in bepaalde situaties, zenuwachtig wordt, of black outs heeft bij S.O.'s of proefwerken, etc. Zijn er duidelijke tekenen van negatieve faalangst dan geeft u een negatieve beoordeling (-). U geeft een positieve beoordeling (+) als de leerling geen tekenen van faalangst vertoont.

### **34 Stressgevoeligheid**

U beoordeelt of de leerling ook onder meer dan normale belasting (extra huiswerk, meer-dan-normaal lastige opgaven, tijdsdruk) goed blijft functioneren, niet in verwarring raakt, etc. Is een leerling bovenmatig gevoelig voor stress, dan geeft u een negatieve beoordeling (-). U geeft een positieve beoordeling (+) wanneer de leerling niet stressgevoelig is of onder druk goed blijft functioneren.

### **35 Werktempo, efficiënt werken**

U beoordeelt of de leerling aangeboden taken (opgaven, opdrachten) in een vlot tempo kan afwerken. Een leerling die doorgaans echt moeite heeft met het normale werktempo geeft u een negatieve beoordeling (-). Zijn er geen bijzondere opmerkingen op dit punt, dan geeft u een positieve beoordeling (+).



## Bijlage 2: De definitieve versie van de checklist studievaardigheid

| Leerling:<br>Klas (& Klastype):<br>Docent:<br>Vak:               |        |        |        |
|------------------------------------------------------------------|--------|--------|--------|
| Beoordelingsaspect                                               | 1<br>— | 2<br>□ | 3<br>+ |
| 1 Huiswerk af                                                    |        |        |        |
| 2 Netheid huiswerk                                               |        |        |        |
| 3 Motivatie                                                      |        |        |        |
| 4 Opletten in de klas                                            |        |        |        |
| 5 Gedrag in de klas                                              |        |        |        |
| 6 Cijfers S.O.'s (hoogte)                                        |        |        |        |
| 7 Cijfers proefwerken (hoogte)                                   |        |        |        |
| 8 Punctualiteit                                                  |        |        |        |
| 9 Omgang met docenten                                            |        |        |        |
| 10 Spreekvaardigheid                                             |        |        |        |
| 11 Leesvaardigheid                                               |        |        |        |
| 12 Lesvoorbereiding                                              |        |        |        |
| 13 Zelfstandig werken                                            |        |        |        |
| 14 Schrijfvaardigheid                                            |        |        |        |
| 15 Uitdrukkingsvaardigheid                                       |        |        |        |
| 16 Vragen wanneer iets niet begrepen wordt                       |        |        |        |
| 17 Van buiten leren                                              |        |        |        |
| 18 Nauwkeurig werken                                             |        |        |        |
| 19 Begrijpen opgedragen taken                                    |        |        |        |
| 20 Aandacht in de les                                            |        |        |        |
| 21 Zelfcontrole                                                  |        |        |        |
| 22 Theoretische interesse                                        |        |        |        |
| 23 Doorzettingsvermogen                                          |        |        |        |
| 24 Concentratievermogen                                          |        |        |        |
| 25 Werktempo, efficiënt werken                                   |        |        |        |
| Tel per kolom de kruisjes; de som der kolomtotalen moet 25 zijn. |        |        |        |
| Vermenigvuldigingsfactoren →                                     |        |        |        |
|                                                                  | x 1    | x 2    | x 3    |
|                                                                  | a      | b      | c      |

Let op: '-' = 'nee'; '+' = 'ja'

Totaalscore:   (a + b + c)





