

Wetenschappelijke verantwoording Leren leren sbo en so (voormalig) cluster 4

Hans Kuhlemeier, Nienke Lansink en Silvia Matijssen

cito.nl



**Leren leren
sbo en so (voormalig) cluster 4**

Wetenschappelijke verantwoording

Hans Kuhlemeier
Nienke Lansink
Silvia Matijssen

© Cito B.V. Arnhem (2020)

Niets uit dit werk mag zonder voorafgaande schriftelijke toestemming van Cito worden openbaar gemaakt en/of verveelvoudigd door middel van druk, fotokopie, scanning, computersoftware of andere elektronische verveelvoudiging of openbaarmaking, microfilm, geluidskopie, film- of videokopie of op welke wijze dan ook.

Inhoud

1	Inleiding	5
2	Uitgangspunten en achtergronden van de testconstructie	7
2.1	Meetpretentie, doelgroep en gebruiksdoel	7
2.1.1	Meetpretentie	7
2.1.2	Doelgroep	8
2.1.3	Gebruiksdoel en functie	9
2.2	Theoretische en onderwijskundige inkadering	11
2.2.1	Cognitieve vaardigheden van leren leren	11
2.2.2	Non-cognitieve vaardigheden van leren leren	13
2.2.3	Aansluiting bij onderwijsdoelen	16
2.2.4	Aansluiting bij bestaande methodieken voor leren leren	18
2.2.5	Dynamic assessment	19
3	Constructie en beschrijving van het pakket Leren leren	25
3.1	Uitgangspunten en overwegingen bij de constructie	25
3.2	De constructie van het pakket Leren leren	26
3.3	Opbouw, afname, beoordeling, rapportage en instructie	29
3.3.1	Opbouw	29
3.3.2	Afname	30
3.3.3	Observatie en beoordeling	34
3.3.4	Groeps- en leerlingrapportages	36
3.3.4.1	Rapportage van de resultaten van de afname van de checklist	36
3.3.4.2	Rapportage van de resultaten van de afname van de praktische opdrachten	38
3.3.5	Handreikingen voor het onderwijzen van leren leren	40
4	Uitvoerbaarheid, betrouwbaarheid en validiteit van de checklist	43
4.1	Doel, opzet en uitvoering van de eerste proefafname	43
4.2	Ervaringen van gebruikers	44
4.3	Betrouwbaarheid van de conceptversie van de checklist	44
4.4	Samenstelling van de definitieve checklist	46
4.5	De definitieve checklist	46
4.5.1	De schalen en uitspraken van de definitieve checklist	46
4.5.2	Beschrijvende gegevens per schaal van de definitieve checklist	47
4.5.3	Omzetting van somscores naar percentielen	48
4.6	Betrouwbaarheid van de definitieve checklist	48
4.6.1	Interne consistentie en item-restcorrelaties	48
4.6.2	Standaardmeetfout	49
4.7	Validiteit van de definitieve checklist	50
4.7.1	Unidimensionaliteit	50
4.7.2	Factorzuiverheid	51
4.7.2.1	Factorzuiverheid van de meting van zelfregulerende vaardigheden	51
4.7.2.2	Factorzuiverheid van de meting van werkhouding	53
4.7.3	Interne structuur	55
4.7.4	Samenhang met achtergrondkenmerken	56
4.7.4.1	Geslacht	57
4.7.4.2	Leeftijd en leerjaar	58
4.7.4.3	Uitstroomprofiel	60
4.7.4.4	Sbo versus so cluster 4	62
4.7.4.5	Doorverwijzing naar een praktische opdracht	62

5 Uitvoerbaarheid, betrouwbaarheid en validiteit van het beoordelingsinstrument 65

- 5.1 Doel, opzet en uitvoering van de tweede proefafname 65
 - 5.1.1 Doel 65
 - 5.1.2 Opzet en uitvoering 66
 - 5.1.3 Het onderzochte materiaal 66
 - 5.1.4 Statistische analyse 67
- 5.2 Uitvoerbaarheid 68
 - 5.2.1 Ervaringen van gebruikers 68
 - 5.2.2 Toewijzing van varianten aan leerlingen 70
 - 5.2.3 Hanteerbaarheid van de beoordelingsformulieren 71
- 5.3 Beschrijvende statistische gegevens 72
 - 5.3.1 Beschrijvende gegevens per beoordelingsaspect 72
 - 5.3.2 Beschrijvende gegevens per schaal 74
- 5.4 Betrouwbaarheid en meetnauwkeurigheid van het beoordelingsinstrument 74
 - 5.4.1 Interne consistentie en item-restcorrelaties 74
 - 5.4.2 Standaardmeetfout 75
- 5.5 Validiteit van het beoordelingsinstrument 75
 - 5.5.1 Factorzuiverheid 76
 - 5.5.2 Interne structuur 78
 - 5.5.2.1 Correlaties tussen de schalen 78
 - 5.5.2.2 Samenhang van leren leren met het geven van hints 79
 - 5.5.3 Verschillen tussen opdrachten en varianten qua moeilijkheidsgraad 83

6 Samenvatting en discussie 87

- 6.1 Samenvatting 87
- 6.2 Discussie 89

Literatuur 91

Bijlagen 97

- 1 Overzicht van doelen per hoofdvaardigheid van leren leren 98
- 2 De uiteindelijke versie van de checklist 100
- 3 De uiteindelijke versie van het beoordelingsinstrument 107
- 4 De omzetting van somscores op de checklist naar percentielen 112

1 Inleiding

Deze wetenschappelijke verantwoording speciaal basisonderwijs heeft betrekking op het pakket speciaal onderwijs Leren leren (Cito, 2018) voor leerlingen in het speciaal basisonderwijs (sbo) en het speciaal onderwijs (so) voormalig cluster 4 (hierna: so cluster 4).

Het pakket stelt leerkrachten in staat de ontwikkeling van vaardigheden in leren leren bij hun leerlingen vast te stellen, te volgen en verder te ontwikkelen. Bij deze zogeheten zelfregulerende vaardigheden gaat het om het leren nadenken over eigen kennis en vaardigheden en het kunnen toepassen ervan bij het uitvoeren van uiteenlopende taken. Daarnaast krijgt de leerkracht informatie over het leerpotentieel van de leerling en zijn of haar werkhouding. Scholen voor sbo en so cluster 4 kunnen het pakket gebruiken in het kader van de wettelijke verplichting tot opbrengstgericht werken, het vaststellen van het ontwikkelingsperspectief, het registreren van de voortgang en het onderwijzen van vaardigheden van leren leren.

Het materiaal omvat een checklist en tien praktische opdrachten met bijbehorende beoordelingsformulieren. Per opdracht is er een makkelijke, een wat moeilijkere en een nog moeilijkere variant beschikbaar gesteld. Hiermee kan de leerkracht de moeilijkheid van de opdracht afstemmen op het vaardigheidsniveau van de leerling.

De opdrachten en varianten zijn onder leiding van twee toetsdeskundigen ontwikkeld door vier ervaren leerkrachten die lesgeven in de groepen 5 tot en met 8 van het sbo en so cluster 4. De toetsdeskundigen werden daarbij ondersteund door een collega met ruime ervaring op het gebied van de constructie van praktische opdrachten voor de centraal schriftelijke praktijkexamens in het vmbo (CSPE).

Het materiaal sluit nauw aan bij:

- de (leergebiedoverstijgende) kerndoelen van het nationaal expertisecentrum leerplanontwikkeling (Ministerie van OCW, 1998; 2006; 2009);
- de '*PO Leerlijn Leren leren Vakoverstijgend*'; deze leerlijn is weliswaar bedoeld voor reguliere basisschoolleerlingen, maar is ook goed bruikbaar voor so- en sbo-leerlingen bij wie het leren niet vanzelf gaat (CED-groep, 2009).

In het constructieproces is de complexiteit van de opdrachten en varianten afgestemd op wat leerlingen in de groepen 5 tot en met 8 van het sbo en so cluster 4 aankunnen.

Het materiaal is ontwikkeld volgens de principes van *dynamic assessment*, een vorm van toetsend leren waarbij toetsen, onderwijzen en leren volledig geïntegreerd zijn (Lidz, 1987, 1991; Tzurriel, 2001). *Dynamic assessment* is een nog vrij onbekende toetsvorm die Cito één keer eerder heeft toegepast bij het Pakket Leren leren voor leerlingen in de uitstroombrofielen Dagbesteding en Arbeid van het Voorgezet Speciaal Onderwijs (Kuhlemeier & Chermin, 2015).

Samen met de uitgave zelf en de bijbehorende gebruikershandleiding (Cito, 2018) bevat deze verantwoording de informatie die nodig is voor een eerste beoordeling van de kwaliteit van het pakket Leren leren. In het vervolg van deze verantwoording beschrijven en verantwoorden we het pakket op de volgende aspecten:

- meetpretentie, doelgroep en gebruiksdoel;
- de keuze van de vaardigheden van leren leren (zowel theoretisch als onderwijskundig);
- de uitgangspunten en achtergronden van testconstructie (met speciale aandacht voor *dynamic assessment*);
- de constructieprocedures;
- de checklist, de praktische opdrachten, varianten, beoordelingsformulieren en handreikingen voor de leerkracht;
- de aanwijzingen voor de afname, beoordeling en rapportage;
- enkele aspecten van betrouwbaarheid en validiteit.

2 Uitgangspunten en achtergronden van testconstructie

In dit hoofdstuk geven we eerst een omschrijving van de meetpretenties, de doelgroep, en het gebruiksdoel en de functie van het pakket leren leren (paragraaf 2.1). Vervolgens beschrijven en verantwoorden we de keuze van de vaardigheden van leren leren (paragraaf 2.2.1 en 2.2.2), de aansluiting van het pakket bij de onderwijsdoelen en de methodieken van leren leren (paragraaf 2.2.3 en 2.2.4) en de gehanteerde toetsvorm: *dynamic assessment* (paragraaf 2.2.5).

2.1 Meetpretentie, doelgroep en gebruiksdoel

2.1.1 Meetpretentie

Het onderwijzen van vaardigheden van leren leren is een effectief middel om leerprestaties duurzaam te verhogen (De Boer, Donker-Bergstra & Kostons, 2012; Dignath & Büttner, 2008; Dignath, Büttner & Langfeldt, 2008; Haller, Child, & Walberg, 1988; Hattie, Biggs & Purdey, 1996; Zumbrunn, Tadlock & Roberts, 2011). Het belang van leren leren zien we ook terug in de officiële onderwijsdoelen en leerplannen voor het onderwijs. Naast vakspecifieke onderwijsdoelen heeft het speciaal onderwijs de opdracht vakoverstijgende doelen bij leerlingen na te streven. In de kerndoelen en leerlijnen voor het sbo en so cluster 4 is daarvoor een belangrijke plaats ingeruimd. Het pakket Leren leren sluit nauw aan bij de leergebied overstijgende kerndoelen van het nationaal expertisecentrum leerplanontwikkeling (Ministerie van OCW, 1998; 2006; 2009) en de vakoverstijgende leerlijnen voor leren leren (CED-groep, 2009).

Het pakket heeft tot doel meer inzicht te geven in de vaardigheden van leerlingen op het gebied van leren leren. Het materiaal is ontwikkeld op basis van de principes van *dynamic assessment*, een vorm van toetsend leren waarbij toetsen, onderwijzen en leren volledig geïntegreerd zijn (Feuerstein, 1979; Feuerstein, Feuerstein & Falik, 2010; Grigorenko & Sternberg, 1998; Lidz, 1987, 1991; Resing, 2006; Sternberg & Grigenko, 2002; Tzurriel, 2001). De meet- en gebruikspretentie kan als volgt worden omschreven:

- het in kaart brengen en volgen van de ontwikkeling van vaardigheden van leren leren;
- het verkennen van het leerpotentieel, dit wil zeggen a) het niveau dat de leerling met hulp van de leerkracht zou kunnen bereiken en b) de aard en omvang van de instructie en begeleiding die nodig is om dat niveau te bereiken;
- het verder ontwikkelen van de vaardigheden van leren leren tijdens de afname en na de afname tijdens instructie- en oefenmomenten (door het toepassen van de handelingsadviezen uit hoofdstuk 6 van de gebruikershandleiding).

Met het pakket kan de leerkracht vaststellen in hoeverre de leerling in staat is vaardigheden van leren leren toe te passen in nieuwe situaties. Bij deze zogeheten zelfregulerende vaardigheden gaat het om het kunnen nadenken over eigen kennis en vaardigheden en het kunnen toepassen ervan bij het uitvoeren van uiteenlopende taken. Het pakket stelt leerkrachten in staat de ontwikkeling van drie vaardigheden van leren leren vast te stellen, te volgen en verder te ontwikkelen:

- Oriënteren (begrijpen waar de opdracht over gaat) en Voorbereiden en Plannen (nadenken voor je begint en de werkzaamheden voorbereiden);
- Uitvoeren en Monitoren: nadenken terwijl je bezig bent en in de gaten houden of je bereikt wat je wilt bereiken;
- Evalueren en reflecteren: controleren of je hebt bereikt wat je wilde bereiken en terugkijken op en leren van de manier waarop je het hebt aangepakt.

Naast deze metacognitieve, zelfregulerende component omvat leren leren ook een non-cognitieve, affectief-motivationale component (o.a. Boekaerts, 1996, 1997; Corno, 1993). Behalve op de drie

bovengenoemde vaardigheden is het pakket gericht op het vaststellen en remediëren van eventuele problemen en blokkades met betrekking tot de werkhouding van de leerling. Belangrijke aspecten van zelfregulerend leren zijn daarbij aandachtig werken (Winne, 1995), gemotiveerd werken, zelfstandig werken en werken zonder gedragsproblematiek (Boekaerts, 1996, 1997).

Voor diagnostiek op het terrein van leren leren is het materiaal niet geschikt. Onderzoek op dit terrein dient te gebeuren met gespecialiseerde instrumenten en door of onder supervisie van een ter zake deskundige leerpsycholoog, orthopedagoog of gedragswetenschapper. Het pakket is slechts een hulpmiddel om de vaardigheden van leren leren vast te stellen en verder te ontwikkelen. Het instrument is uitdrukkelijk bestemd voor gebruik door de groepsleerkracht of iemand anders (bijvoorbeeld een intern begeleider) die de leerling goed kent.

De kern van het pakket Leren leren bestaat uit een checklist, een serie praktische opdrachten met bijbehorende beoordelingsformulieren en aanwijzingen voor de afname van de checklist en de praktische opdrachten. Het materiaal is echter niet in eerste instantie bedoeld om de beheersing van praktische vaardigheden vast te stellen, zoals het kunnen meten van de breedte van een tafel of het maken van een PowerPointpresentatie. De praktische opdrachten zijn een middel om algemene vaardigheden van leren leren te beoordelen en verder te ontwikkelen.

Het pakket is geen complete methode voor het beoordelen en onderwijzen van vaardigheden van leren leren. Wel bevat het pakket handreikingen en lessuggesties, onder meer voor het houden van leergesprekken tijdens de afname. Ook wordt de leerkracht in de gebruikershandleiding (Cito, 2018) verwezen naar gangbare methoden en methodieken voor het onderwijzen van leren leren, zoals de zogenoemde Beertjes-methodiek (Meichenbaum, 1981; Timmerman, 2002, 2017; Van Bon, 2009), en naar recente handreikingen voor het onderwijzen van zogeheten executieve functies (Cooper-Kahn & Dietzel, 2011; Dawson & Guare, 2009, 2010; Smidts, 2018; Smidts & Huizinga, 2011, 2017).

Bij de constructie van de checklist en de praktische opdrachten hebben traditionele psychometrische eigenschappen zoals moeilijkheidsgraad, discriminerend vermogen en betrouwbaarheid een belangrijke rol gespeeld. In overeenstemming met de principes van *dynamic assessment* heeft de keuze van de praktische opdrachten en beoordelingsaspecten echter vooral plaatsgevonden op basis van de mogelijkheden voor het onderwijzen van vaardigheden van leren leren (Tzurriel, 2001). De gerichtheid op het optimaliseren van onderwijzen en leren heeft echter een keerzijde die hier niet onvermeld mag blijven. Het pakket Leren leren is geschikt voor het vaststellen en volgen van de ontwikkeling van individuele leerlingen van eenzelfde leerkracht. Het is uitdrukkelijk niet geschikt voor het vergelijken van de vaardigheden van leerlingen van verschillende leerkrachten of van verschillende scholen.

2.1.2 Doelgroep

Het pakket Leren leren is bedoeld voor leerlingen van groep 5 tot en met 8 in het speciaal basisonderwijs en het speciaal onderwijs cluster 4. Leerlingen in het sbo en so cluster 4 hebben vaak last van externaliserende gedragsproblemen (40% van de leerlingen in het so en sbo), internaliserende gedragsproblemen (50%), een verstandelijke beperking (50%), een leerachterstand (60%), communicatieproblemen (40%) en/of werkhoudingsproblemen (50%) (Floor, De Zoete, Schaafsma & Van den Oetelaar, 2013).

Het speciaal basisonderwijs (sbo) is voor leerlingen die extra begeleiding nodig hebben om het regulier basisonderwijs te doorlopen. Als het regulier onderwijs de leerling niet goed genoeg kan begeleiden bij zijn leer-, opvoedings- en/of gedragsproblemen, gaat de leerling naar het sbo. Op de website van de Rijksoverheid staat dat de leerlingen in het sbo dezelfde kerndoelen moeten halen als leerlingen op reguliere basisscholen. Zij krijgen echter meer tijd om die te halen en mogen tot hun veertiende jaar naar een school voor speciaal basisonderwijs (Rijksoverheid, z.d.). Dat betekent in de praktijk echter niet dat leerlingen tot hun veertiende op het sbo zitten en dat zij alle kerndoelen hebben gehaald. Vaak stromen zij al eerder uit.

De doelgroep waarop het sbo zich voornamelijk richt, zijn moeilijk lerende kinderen, kinderen met opvoedingsmoeilijkheden en kinderen met gedragsproblemen (Rijksoverheid, z.d.).

Als een leerling extra ondersteuning nodig heeft die het regulier of het speciaal basisonderwijs niet kan bieden, gaat een leerling naar het speciaal onderwijs. Er zijn verschillende clusters in het speciaal onderwijs, er wordt altijd gekeken in welk cluster een leerling het beste past. Cluster 4 is voor leerlingen met stoornissen en gedragsproblemen (Rijksoverheid, 2017). Vaak is er sprake van een gecombineerde problematiek (denk aan ASS, ADHD, ODD en angststoornissen).

Met de invoering van passend onderwijs, zijn scholen sinds 1 augustus 2014 verantwoordelijk om alle leerlingen die extra ondersteuning nodig hebben een goede onderwijsplek te bieden. Reguliere en speciale scholen werken daarvoor samen in regionale samenwerkingsverbanden. Het samenwerkingsverband beslist of een leerling kan worden toegelaten tot het sbo of so (Passend Onderwijs, z.d.).

Hieronder beschrijven we twee problemen waar veel leerlingen in het sbo en so cluster 4 mee te kampen hebben en waar we ons bij de constructie van het pakket Leren leren op hebben gericht. Deze problemen kunnen de ontwikkeling van vaardigheden op het gebied van leren leren belemmeren.

Werkhoudingsproblemen

De werkhouding is het gedrag van een leerling dat een taakoplossing positief of negatief kan beïnvloeden. Als de taak negatief beïnvloed wordt, heeft de leerling een werkhoudingsprobleem.

Het gedrag kan zichtbaar zijn, maar er is ook gedrag dat niet zichtbaar is (denkgedrag).

Verklaringen voor werkhoudingsproblemen kunnen liggen op het gebied van: aandachts- en concentratiestoornissen, motivationele problemen en cognitieve factoren. Maar werkhoudingsproblemen kunnen ook komen door factoren in de omgeving, bijvoorbeeld de relatie met de leerkracht of de inhoud en vormgeving van de taak (Timmerman, 2002, 2017).

Gedragsproblemen

Er zijn vier typen gedragsproblemen: druk gedrag, opstandig en agressief gedrag, teruggetrokken gedrag, en sociaal opvallend gedrag (Bijl-van Gelder & van Elferen-Wiersma, 2016).

Leerlingen die *druk gedrag* vertonen, lijken een onuitputtelijke hoeveelheid energie te hebben, praten vaak veel en houden daarbij slecht rekening met anderen.

Leerlingen met *opstandig en agressief gedrag* vertonen geïrriteerd boos gedrag, ontevreden en prikkelbaar gedrag, maken vaak ruzie en hebben last van woedeaanvallen. Maar ze vertonen ook gedrag waarbij (bewust) schade aan anderen wordt aangebracht en waarbij gangbare normen en waarden worden overschreden.

Leerlingen met *teruggetrokken gedrag* zijn vaak stil en zoeken weinig contact met anderen. Het kan zijn dat ze dit niet durven of er weinig behoeften aan hebben, het kan ook komen doordat ze onhandig zijn in het leggen van contacten.

Kinderen die *sociaal opvallend gedrag* vertonen zijn vaak onvoldoende in staat om hun eigen gedrag te beoordelen en waar nodig aan te passen. Hierdoor komen zij vaak sociaal onhandig over. Ze hebben vaak moeite met het herkennen en begrijpen van gevoelens, gedachten en intenties van anderen.

2.1.3 Gebruiksdoel en functie

Het algemene doel van het pakket Leren leren is het vaststellen, volgen en verder ontwikkelen van vaardigheden van leren leren van sbo- en so cluster 4-leerlingen (in de groepen 5 tot en met 8).

Daartoe bevat het pakket een checklist, een serie praktische opdrachten met bijbehorende beoordelingsformulieren en handreikingen voor het al toetsend onderwijzen van vaardigheden van leren leren.

Sbo- en so cluster 4-scholen kunnen het pakket gebruiken in het kader van de wettelijke verplichting tot opbrengstgericht werken, het vaststellen van het ontwikkelingsperspectief, het registreren van de voortgang en het onderwijzen van vaardigheden van leren leren.

Doel van de checklist

De praktische opdrachten vormen de kern van het pakket Leren leren. In vergelijking met een klassikaal af te nemen toets is de individuele afname van een praktische opdracht tijdrovend. In het VSO waar de docent meestal samen met een onderwijsassistent voor de groep staat, is dat een minder groot probleem dan in het sbo- en so cluster 4-onderwijs waar de leerkracht er doorgaans alleen voorstaat. Niet onverwacht bleek tijdens verkennende gesprekken op de scholen en de proefafnames dat het in veel situaties erg lastig is om de praktische opdrachten aan alle leerlingen uit de groep voor te leggen. Wel gaven vrijwel alle bevraagde leerkrachten aan een praktische opdracht af te willen nemen bij een deel van hun leerlingen.

De leerkrachten gaven tegelijkertijd aan niet zo goed te weten welke leerlingen het meest bij afname van een praktische opdracht gebaat zouden zijn. Er was met andere woorden behoefte aan een efficiënt instrument waarmee de leerkracht een voorselectie kon maken. Voor dat doel is er een checklist aan het pakket toegevoegd.

De checklist heeft tot doel snel een indruk te verkrijgen van het niveau van leren leren in de groep als geheel. De groepsrapportage laat zien welke leerlingen meer en minder moeite hebben met leren en leren en welke vaardigheden in de groep als geheel meer en minder goed ontwikkeld zijn. De checklist is vooral een hulpmiddel om te bepalen welke vaardigheden extra (klassikale) aandacht vereisen en bij welke vaardigheden dat niet of minder nodig is. Daarnaast is de checklist een hulpmiddel om de leerlingen te selecteren die het meest voor afname van een arbeidsintensieve praktische opdracht in aanmerking komen. Daartoe kan de leerkracht ook te rade gaan bij de leerlingrapportage. Die laat zien hoe de leerling het op de onderscheiden vaardigheden doet in vergelijking met een landelijke referentiegroep van leerlingen uit sbo en so cluster 4.

De afnametijd van de checklist is ongeveer vijf à tien minuten per leerling.

Doel van de praktische opdrachten

De praktische opdrachten hebben tot doel de vaardigheden van leren leren vast te stellen, te volgen en verder te ontwikkelen. Deze meet- en gebruikspretentie komt het meest tot zijn recht als de praktische opdrachten individueel worden afgenomen.

Het vaststellen en onderwijzen van leren leren met de praktische opdrachten

Behalve de checklist bevat het pakket Leren leren tien praktische opdrachten met telkens drie varianten van een verschillende moeilijkheidsgraad (A = makkelijk; B = wat moeilijker; C = nog wat moeilijker). Na afloop van de afname van een variant kan de leerkracht een leerling- en/of een groepsrapportage aanmaken. De leerlingrapportage biedt informatie over de beheersing van de vaardigheden oriëntatie/voorbereiding, uitvoering en reflectie/evaluatie, de hoeveelheid benodigde hulp en de werkhouding. De globale betekenis van de scores voor de vaardigheden en de drie aspecten van werkhouding is 1 = zwakke vaardigheid; 2 = redelijke vaardigheid; 3 = goede vaardigheid; 4 = uitstekende vaardigheid. De scores voor de hoeveelheid benodigde hulp hebben de volgende betekenis: 1 = voortdurend hulp; 2 = regelmatig hulp; 3 = af en toe hulp; 4 = geen hulp. In de groepsrapportage zijn de toetsresultaten van meerdere leerlingen op een overzichtelijke manier naast elkaar gezet. Meer informatie over de rapportage is te vinden in paragraaf 3.3.4 en de gebruikershandleiding (Cito, 2018).

Het pakket Leren leren is niet bedoeld voor het diagnosticeren van storingen of problemen. Een lage score betekent niet dat de leerling – alleen op grond van deze lage score – als probleemleerling mag worden aangemerkt. Voor dat doel zijn aanvullende onderzoeksgegevens nodig waaruit blijkt dat leerlingen met lage scores ook daadwerkelijk problematisch zijn of later zijn geworden. Dit onderzoek is nog niet beschikbaar.

Het volgen van de ontwikkeling van leren leren

Met het pakket kan de leerkracht bepalen hoe vaardig een leerling is op het gebied van leren leren en hoeveel hulp hij of zij nodig heeft gehad. Door een tweede opdracht na de eerste opdracht af te nemen, kan worden vastgesteld in hoeverre de vaardigheid in leren leren vooruit is gegaan (en hoeveel de leerling van

de eerste afname en het onderwijs in leren leren in de tussenliggende periode geleerd heeft). Van elke opdracht is er een makkelijke, wat moeilijkere en nog moeilijkere variant. Welke variant de leerkracht het beste als tweede kan afnemen, is afhankelijk van de scores van de leerling op de eerste opdracht. Met de bijgeleverde digitale rapportagemodule (Excel) kan de leerkracht leerling- en groepsrapportages van verschillende afnames bewaren en met elkaar vergelijken. Hoe het volgen van leerlingen precies in zijn werk gaat, is uiteengezet in de gebruikershandleiding (Cito, 2018).

Het al toetsend onderwijzen van leren leren

Leren leren is een relatief vaag begrip dat zeer verschillend gedefinieerd wordt. Mede door het gebrek aan geschikt lesmateriaal is het onderwijzen van vaardigheden van leren leren nog geen gemeengoed in het onderwijs. Behalve testmateriaal bevat het pakket daarom ook uitgebreide handreikingen voor het verder ontwikkelen van vaardigheden van leren leren deels tijdens, maar vooral ook in de reguliere lessen na de testafname. Deze handreikingen zijn grotendeels gebaseerd op de bewezen effectieve BOL/EIM-methodiek (Aukes, Dekker, de Jong, Mulder & Timmer, 2000; Den Besten & Vulpen, 2013; Timmer, 1994; Timmer, Dekker & Voortman, 2003), de zogenoemde Beertjesmethodiek (Meichenbaum, 1981; Timmerman, 2002, 2017; Van Bon, 2009) en methodieken voor het verder ontwikkelen van zogeheten executieve functies (Cooper-Kahn & Dietzel, 2011; Dawson & Guare, 2009, 2010; Smidts, 2018; Smidts & Huizinga, 2011, 2017). Voor meer informatie over deze methodieken en de handreikingen verwijzen wij naar de gebruikershandleiding (Cito, 2018).

2.2 Theoretische en onderwijskundige inkadering

In deze paragraaf beschrijven en verantwoorden we eerst de keuze van de vaardigheden van leren leren (paragraaf 2.2.1 en 2.2.2). Daarna bespreken we de aansluiting bij de onderwijsdoelen (paragraaf 2.2.3) en de aansluiten bij bestaande methodieken (paragraaf 2.2.4). Tot slot gaan we in op *dynamic assessment*, de methode van toetsend leren waarmee de vaardigheden gemeten en onderwezen worden (paragraaf 2.2.5).

2.2.1 Cognitieve vaardigheden van leren leren

Theoretische inkadering

Om nieuwe kennis en vaardigheden te leren en deze in bekende en nieuwe situaties te kunnen toepassen zijn zogeheten zelfregulerende vaardigheden nodig (Timmer, Dekker & Voortman, 2009). Pintrich (2000) omschrijft zelfregulerend leren als: *“an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation and behavior, guided and constrained by their goals and the contextual features in the environment”* (p. 453). Zelfregulerende vaardigheden worden vaak ingedeeld op basis van de fase in het leerproces waarin de leerling ze kan toepassen. Men spreekt dan bijvoorbeeld over vaardigheden van oriëntatie en planning, uitvoering, en reflectie en evaluatie (Butler & Winne, 1995; Pintrich & Zusho, 2002; Zimmerman, 1990; 2002; Zumbunn, Tadlock & Roberts, 2011). Zelfregulerend leren wordt gezien als een recursieve cyclus van cognitieve en non-cognitieve activiteiten (Butler & Winne, 1995; Carver & Scheier, 1990; Zimmerman, 2000). Figuur 2.1 toont de drie fasen van dit recursieve proces zoals overgenomen uit Zumbunn, Tadlock en Roberts (2011).



Figuur 2.1 Het cyclische proces van zelfregulerend leren in drie fasen

Deze driedeling zien we terug in de volgende, door ons enigszins geparafraseerde beschrijving van de zelfregulerende leerling (Butler, 1996; Butler & Winne, 1995; Carver & Scheier, 1990; Pintrich & Zusho, 2002; Zimmerman, 2000):

Alvorens met een nieuwe taak te beginnen, analyseert en interpreteert de leerling de taak in relatie tot de reeds aanwezige kennis, vaardigheden en opvattingen (oriëntatie). Daarbij stelt de leerling taak-specifieke doelen en kiest hij of zij geschikte strategieën om die doelen te bereiken (voorbereiding en planning). Tijdens de uitvoering van de taak monitort de leerling de mate waarin het doel bereikt is (monitoring). In de fase na de uitvoering evalueert de leerling in hoeverre de inspanningen succesvol waren (evaluatie). De leerling past, als daar aanleiding toe is, de aanpak aan op basis van de perceptie van de vooruitgang (kalibratie, bijstelling, revisie) waarna de cyclus opnieuw begint.

Kortom: tijdens de voorbereiding, uitvoering en evaluatie van taken sturen zelfregulerende leerlingen hun leeractiviteiten waarbij zij hun aanpak zo nodig op een flexibele manier aanpassen. Het belang van zelfregulering zien we ook terug in recente literatuur over 21e-eeuwvaardigheden (o.a. Ledoux, Meijer, Van der Veen & Breetvelt, 2013; Marzano & Heflebower, 2013; Voogt & Pareja Roblin, 2010). Op basis van een uitgebreide literatuurstudie rekent de SLO (Thijs, Fisser & Van der Hoeven (2014) zelfregulering tot de 8 belangrijkste 21e-eeuwvaardigheden (samen met creativiteit, kritisch denken, probleemoplossingsvaardigheden, communiceren, samenwerken, digitale geletterdheid en sociale en culturele vaardigheden). Bij zelfregulering gaat het om "het kunnen realiseren van doelgericht en passend gedrag" (p. 37).

Meer specifiek gaat het bij zelfregulering om vaardigheden die grote gelijkenis vertonen met de denkvaardigheden en werkhoudingen die in het pakket Leren leren centraal staan:

- het stellen van realistische doelen en prioriteiten;
- doelgericht handelen (concentratie, zichzelf kunnen motiveren voor en richten op de uitvoering van een taak, zelfstandigheid) en monitoren van het proces (planning, timemanagement);
- reflectie op het handelen en de uitvoering van de taak, en feedback op het eigen gedrag en handelen benutten om adequate vervolgkeuzes te maken;
- inzicht hebben in de ontwikkeling van eigen competenties;
- verantwoording nemen voor eigen handelen en keuzes, en zicht hebben op consequenties van het eigen handelen voor de omgeving, ook op de lange termijn.

Vaardigheden van leren leren worden vooral van belang geacht voor minder vaardige leerlingen (Dinsmore & Parkinson, 2013; Zimmerman, 1990). Niet alleen omdat zij minder goed zijn in leren, maar ook omdat “they also tend to be less likely to *know* they are inaccurate or understand how close they are to the desired learning” (Dinsmore & Parkinson, p. 63). Moeilijk lerende kinderen zijn minder geneigd om effectieve zelfregulerende vaardigheden toe te passen dan reguliere leerlingen. Bij het uitvoeren van taken hebben zij vooral moeite met de vaardigheden plannen, monitoring, uittesten, checken, evalueren en reviseren. Onderzoek laat zien dat het mogelijk is om deze vaardigheden via gerichte instructie en begeleiding verder te ontwikkelen (De Boer, Donker-Bergstra & Kostons, 2012; Dignath & Büttner, 2008; Dignath, Büttner & Langfeldt, 2008; Haller, Child & Walberg, 1988; Hattie, Biggs & Purdey, 1996; Wang, Haertel & Walberg, 1990; Zumbunn, Tadlock & Roberts, 2011).

Zelfregulerende vaardigheden zijn vooral voor leerlingen in het speciaal onderwijs van groot belang. Velen van hen hebben moeite met het toepassen van praktische vaardigheden bij een nieuwe opdracht of in een nieuwe situatie. De leerling heeft bijvoorbeeld geleerd hoe hij tijdens de les begrijpend lezen vragen bij een tekst kan beantwoorden. Als hij echter tijdens de les geschiedenis een vergelijkbare opdracht krijgt, lukt het opeens niet meer. Om dat toch te kunnen heeft de leerling vaardigheden van leren leren nodig. Het gaat daarbij om oriënteren (begrijpen waar de opdracht over gaat), plannen (nadenken voor je begint en een plan maken), uitvoeren (nadenken terwijl je bezig bent en in de gaten houden of je bereikt wat je wilt bereiken) en evalueren (controleren of je hebt bereikt wat je wilde bereiken en terugkijken op en leren van de manier waarop je het hebt aangepakt). Als de leerling zich deze denkvaardigheden eigen heeft gemaakt, is het grote voordeel dat hij zijn praktische vaardigheden ook in andere situaties kan toepassen. De leerling kan dan zelf een oplossing bedenken en zo de doe-vaardigheden goed blijven uitvoeren (Timmer & Van Gemert, 1992).

2.2.2 Non-cognitieve vaardigheden van leren leren

Naast deze metacognitieve, zelfregulerende component omvat leren leren ook een non-cognitieve, affectief-motivationale component (o.a. Boekaerts, 1996, 1997; Corno, 1993). Tijdens het uitvoeren van nieuwe taken moeten leerlingen ook in staat zijn om non-cognitieve strategieën toe te passen als zij geconfronteerd worden met tegenslag, waarbij aandacht, motivatie, zelfstandigheid en impulsbeheersing een belangrijke rol spelen (Boekaerts, 1996, 1997; Corno, 1993; Pintrich, 2000, 2004; Pintrich & Zusho, 2002; Schunk & Zimmerman, 2008; Zimmerman, 2008). Behalve op de cognitieve, zelfregulerende vaardigheden is het pakket gericht op het vaststellen en remediëren van eventuele problemen en blokkades met betrekking tot de werkhouding van de leerling. Belangrijke aspecten van de werkhouding zijn daarbij aandachtig werken (Winne, 1995), gemotiveerd werken, zelfstandig werken en werken zonder gedragsproblematiek (Boekaerts, 1996, 1997).

Aandachtig werken is een krachtige voorspeller van leren in het cognitieve domein. Leerlingen die moeite hebben met het richten van hun aandacht blijken minder goede presteerders en zelfs soms onderpresteerders te zijn ten opzichte van leerlingen die in staat zijn om aandachtig en zorgvuldig te werken (DeShazo, Lyman & Grofer-Klinger, 2002). Vandaar dat er zowel in de checklist als in de beoordelingsformulieren bij de praktische opdrachten ruimschoots aandacht is besteed aan de mate waarin

de leerling in staat is tot aandachtig werken tijdens het voorbereiden, plannen, uitvoeren en evalueren van praktische opdrachten.

Motivatie voor leren behoort tot de persoonlijke opbrengsten van onderwijs en opvoeding. Het is een belangrijke indicator van het sociaal-emotionele functioneren op school en in de klas (Mönks & Knoers, 2009). Een aantal aspecten van motivatie draagt bij aan het behalen van hogere prestaties op school (Steinmayr & Spinath, 2009). Als leerlingen bijvoorbeeld plezier hebben in rekenen en taal, kan dit goede leerprestaties voor die vakgebieden bevorderen. Motivatie wordt over het algemeen gezien als een noodzakelijke voorwaarde voor leren (Boekaerts & Simons, 1995). Resultaten van onderzoek naar de samenhang tussen algemene motivatie voor school en schoolprestaties zijn niet eenduidig. Wellicht heeft dit te maken met de veelal indirecte verbanden. Leerlingen die geloven in hun eigen kunnen en intrinsieke interesse hebben in de leerstof, willen meer tijd en inzet investeren in het ontwikkelen van cognitieve strategieën en metacognitieve vaardigheden. Deze investering leidt uiteindelijk tot betere leerresultaten (Boekaerts & Simons, 1995). Intrinsiek gemotiveerde leerlingen voeren taken uit omdat zij daar plezier aan beleven, zonder dat zij zich daartoe gedwongen voelen en zonder dat daar een externe beloning of een straf tegenover staat (Deci & Ryan, 1985, 1991, 2000).

Onderzoek laat zien dat een aanzienlijk deel van de leerlingen in het regulier basisonderwijs te kampen heeft met een gebrekkige motivatie, bijvoorbeeld voor de vakgebieden lezen en rekenen (Kuhlemeier, Van Boxtel & Van Til, 2012). Leerkrachten besteden in de dagelijkse lespraktijk veel tijd en energie aan het bevorderen van de motivatie van hun leerlingen. Uit onderzoek blijkt dat intrinsiek gemotiveerde leerlingen zich meer voor school en leren inzetten en hogere leerprestaties behalen dan niet of extrinsiek gemotiveerde leerlingen. Waar algemene motivatie voor school niet eenduidig en altijd rechtstreeks samenhangt met schoolprestaties, hangt domeinspecifieke motivatie wel sterk samen met schools presteren (Boekaerts & Simons, 1995). Het is om deze redenen dat wij in de checklist en de beoordelingsformulieren een speciale plaats hebben ingeruimd voor het meten en bevorderen van het gemotiveerd werken aan praktische opdrachten. Een bijkomende reden is dat motivatie een belangrijk aandachtspunt is in het toezicht van de Inspectie op de kwaliteit van het onderwijs (Inspectie van het Onderwijs, 2010).

Bij ongeveer de helft van de leerlingen in het sbo en so is er sprake van een verstandelijke beperking. Bij hen gaat die beperking vaak samen met een beperkte zelfstandigheid en zelfredzaamheid. De gangbare methodieken voor het onderwijzen van zelfregulerende vaardigheden en werkhouding zijn gericht op het verhogen van deze zelfstandigheid en zelfredzaamheid (o.a. Aukes e.a., 2000; Cooper-Kahn & Dietzel, 2011; Dawson & Guare, 2009, 2010; Den Besten & Vulpen, 2013; Meichenbaum, 1981; Smidts, 2018; Smidts & Huizinga, 2011, 2017; Timmer, Dekker & Voortman, 2003; Timmerman, 2002; Van Bon, 2009). In de literatuur wordt daarbij vaak verwezen naar een 'kan-ik-niet-mentaliteit' ofwel 'doe-jij-het-maar-houding' die mede tot stand gekomen is door begeleiders die leerlingen met de beste bedoelingen zoveel mogelijk uit handen nemen (Maes, Penne & Petry, 2005; Timmer 1994; Timmer, Dekker & Voortman, 2009; Van Gennep, 1989, 1994). Alhoewel goed bedoeld, krijgen leerlingen in het speciaal onderwijs volgens deskundigen soms te veel hulp en begeleiding, wat met zich meebrengt dat ze zich passief opstellen en de leerkracht vaker om hulp vragen dan nodig en goed voor hen is. Tegenwoordig zijn de gangbare methodieken voor het ontwikkelen van vaardigheden van leren leren er mede op gericht om leerlingen te activeren en hen tevens te leren wanneer en hoe vaak zij hulp kunnen vragen en hoe zij dat het beste kunnen doen. Vandaar dat in de checklist ook een schaal opgenomen is over zelfstandig werken en hulp vragen. Een bijkomende reden is dat dit aspect van de werkhouding voor de leerkrachten die deelnamen aan de proefafnames een hoge prioriteit had.

Een deel van de leerlingen in het speciaal onderwijs heeft te lijden onder externaliserende problematiek die hen hindert hun cognitieve en sociale vaardigheden verder te ontwikkelen en in praktijk te brengen. Leerkrachten besteden een niet onaanzienlijk deel van hun tijd en energie aan de begeleiding van 'lastige' of anderszins problematische leerlingen (Nijntjes & Van Diepen, 2008). Van scholen wordt verwacht dat zij een beleid hebben waarin gedragsproblematiek vroegtijdig gesignaleerd en behandeld wordt (Inspectie van het Onderwijs, 2010). Het werken aan praktische opdrachten stelt leerlingen met een gebrekkige impulscontrole voor extra uitdagingen. Het is mede om deze redenen dat wij in de checklist en de

praktische opdrachten ook aandacht hebben geschonken aan het in kaart brengen van gedragsproblematiek tijdens het werken aan praktische opdrachten.

Meetbare ontwikkelingsdoelen

De vaardigheden van leren leren hebben wij vertaald in concrete en meetbare ontwikkelingsdoelen die de leerling in meer of minder mate kan bereiken. Onderstaand overzicht bevat de ontwikkelingsdoelen die passen bij wat sbo- en so cluster 4-leerlingen kunnen bereiken en die met de praktische opdrachten van het pakket Leren leren getoetst en geïnstrueerd kunnen worden.

Vaardigheid	Ontwikkelingsdoel van leren leren
1 Oriëntatie	De leerling geeft er blijk van te weten wat hij/zij straks gaat doen, welke spullen er nodig zijn, om welke werkzaamheden (stappen) het gaat, in welke volgorde deze worden uitgevoerd en wat er gedaan moet worden als de opdracht klaar is.
2 Voorbereiding en planning	De leerling pakt de benodigde spullen, pakt geen onnodige spullen en legt ze klaar op de werkplek (compleet en overzichtelijk).
3 Uitvoering	De leerling voert de werkzaamheden zichtbaar bewust uit (monitoring) en doet dat planmatig en zichzelf controlerend.
4 Reflectie en evaluatie	De leerling is zich zichtbaar bewust van/kan een beschrijving geven van: a) de werkzaamheden en gevoelens (reflectie); b) wat er goed en minder goed ging (terugkijken); c) wat hij/zij de volgende keer anders zou doen (vooruitkijken).
5 Werkhouding	De leerling voert de opdracht aandachtig, gemotiveerd en zonder gedragsproblematiek uit.

Leren leren en executieve functies

Sinds kort vormen de hiervoor onderscheiden vaardigheden van leren leren het object van neuropsychologisch onderzoek naar zogeheten executieve functies. In navolging van Gioia e.a. (2000; 2002) definiëren Cooper-Kahn en Dietzel (2011) executieve functies als “een verzameling van processen die te maken hebben met het beheren van jezelf en de bronnen die nodig zijn om een doel te bereiken” (p. 22) met de kanttekening dat het een containerbegrip is voor neurologische vaardigheden zoals mentale controle en zelfregulatie. Smidts en Huizing (2011) definiëren executieve functies als “... een verzamelterm voor denkprocessen (functies) die belangrijk zijn voor het uitvoeren (de executie) van sociaal en doelgericht gedrag, zoals op tijd komen”(p. 14). Volgens Chan, Shum, Touloupoulou en Chen (2008) betreft het “*an umbrella term comprising a wide range of cognitive processes and behavioral competencies which include verbal reasoning, problem-solving, planning, sequencing, the ability to sustain attention, resistance to interference, utilization of feedback, multitasking, cognitive flexibility, and the ability to deal with novelty*” (p. 201). In lijn met de eerder gepresenteerde indeling in cognitieve en non-cognitieve vaardigheden onderscheiden onderzoekers een aantal zogenaemde ‘koude’ en ‘hete’ executieve functies. Beperkingen op het gebied van cognitieve en non-cognitieve executieve functies kunnen een verwoestend effect hebben op allerlei activiteiten uit het leven van alledag, zoals het vermogen naar school te gaan of te werken, het zelfstandig functioneren in de thuissituatie en het ontwikkelen en onderhouden van sociale relaties (Chan e.a., 2008; Smidts & Huizinga, 2011).

De inhoudelijke overlap met ‘onze’ vaardigheden van leren leren blijkt onder meer uit het overzicht van Gioia e.a. (2000; 2002) waarin acht, onderling samenhangende executieve functies worden onderscheiden. Hierna presenteren we deze vaardigheden zoals geformuleerd door Cooper-Kahn & Dietzel (2011; p. 25 en 26):

- 1 Responsinhibitie. Het vermogen het eigen gedrag, handelingen en gedachten op tijd te stoppen.
- 2 Flexibiliteit. Het vermogen makkelijk van de ene situatie naar de andere te gaan en flexibel te denken om adequaat op de situatie te reageren.
- 3 Emotieregulering. Het vermogen emotionele reacties te reguleren met rationale gedachten.

- 4 Taakinitiatie. Het vermogen om zonder uitstel aan een taak of handeling te beginnen en zelfstandig ideeën, antwoorden of probleemoplossende strategieën te genereren.
- 5 Werkgeheugen. De vaardigheid om informatie in het geheugen te houden om een taak te voltooien.
- 6 Planning/organisatie. Het vermogen om datgene te beheersen wat nodig is voor huidige en toekomstige taken.
- 7 Organisatie van materiaal. Het vermogen om orde op zaken te stellen op het werk, tijdens het spel en in opslagruimten.
- 8 Zelfmonitoring. Het vermogen de eigen prestatie te controleren en te vergelijken met een soort standaard van wat nodig is of verwacht wordt.

De zelfregulerende vaardigheden uit het pakket Leren leren hebben eveneens veel gemeenschappelijk met de zes door Smidts en Huizinga (2011, p. 14) onderscheiden executieve functies:

- 1 Weerstand bieden aan impulsen
- 2 Efficiënt plannen maken
- 3 Omgaan met verandering
- 4 Actief gebruik van het geheugen
- 5 Inzicht hebben in je eigen gedrag en houding
- 6 Effectief omgaan met emoties

Tot slot verwijzen we nog naar de gelijkenis met de vaardigheden uit het overzicht van Dawson en Guare (2010) die twee dimensies van executieve functies onderscheiden: denken en doen.

Executieve Functies: denken	Executieve Functies: doen
werkgeheugen	responsinhibitie
planning/prioritering	emotieregulatie
organisatie	volgehouden aandacht
timemanagement	taakinitiatie
metacognitie	doelgericht gedrag
	flexibiliteit

2.2.3 Aansluiting bij onderwijsdoelen

De keuze van de relevante vaardigheden van leren leren is ook gebaseerd op de kerndoelen (Ministerie van OCW, 1998; 2006; 2009) en de vakoverstijgende leerlijnen voor leren leren (CED, 2009).

Ter ondersteuning van het onderwijs in leren leren heeft de overheid in 1998 in de tweede generatie van de kerndoelen basisonderwijs een aantal leergebiedoverstijgende kerndoelen geformuleerd (Ministerie van OCW, 1998). Inmiddels is de derde generatie kerndoelen van kracht. De leergebiedoverstijgende doelen uit de tweede generatie worden niet meer apart vermeld. Deze zijn zoveel mogelijk opgenomen in de kerndoelen per leergebied. (Ministerie van OCW, 2006). Toch geven we hier de voor leren leren relevante leergebiedoverstijgende doelen uit de tweede generatie weer, omdat deze helder weergeven welke vaardigheden door leerlingen ontwikkeld moeten worden. Het pakket Leren leren is gericht op vaardigheden van leren leren die gerelateerd zijn aan de volgende leergebiedoverstijgende kerndoelen (Ministerie van OCW, 1998):

Werkhouding

De leerlingen hebben belangstelling voor de wereld om hen heen en ze zijn gemotiveerd deze te onderzoeken:

- a) ze kunnen gerichte vragen stellen;
- b) ze kunnen relevante informatie zoeken en gebruiken;

- c) ze hebben plezier in het leren van nieuwe dingen;
- d) ze zetten door wanneer iets niet direct lukt.

Werken volgens plan

De leerlingen kunnen een plan opstellen en ernaar handelen:

- a) ze kunnen een doel formuleren;
- b) ze kunnen zich op een onderwerp oriënteren;
- c) ze begrijpen bij eenvoudige problemen wat oorzaken en gevolgen zijn;
- d) ze kunnen uit het voorgaande onder c. conclusies trekken;
- e) ze kunnen grotere activiteiten stap voor stap indelen en uitwerken;
- f) ze kunnen achteraf beoordelen of hun planning een goede was;
- g) ze kunnen de resultaten van hun werk presenteren in de vorm van een toelichting, een spreekbeurt, een werkstuk, een uitstalling.

Gebruik van uiteenlopende leerstrategieën

De leerlingen kunnen bij leeractiviteiten uiteenlopende strategieën en vaardigheden gebruiken:

- a) ze kunnen gerichte vragen stellen aan personen;
- b) ze kunnen feiten van meningen onderscheiden;
- c) ze kunnen relevante informatie zoeken en verwerken uit andere bronnen, waaronder een woordenboek, een atlas, een register;
- d) ze kunnen samenwerken en met elkaar overleggen om gezamenlijk tot oplossingen te komen;
- e) ze kunnen achteraf beoordelen of hun strategie een goede was.

Zelfbeeld

De leerlingen leren met hun eigen mogelijkheden en grenzen om te gaan:

- a) ze hebben zelfvertrouwen;
- b) ze kunnen gedragsimpulsen beheersen;
- c) ze kunnen en durven voor zichzelf en anderen op te komen.

Sociaal gedrag

De leerlingen leveren een positieve bijdrage in een groep:

- a) ze gaan respectvol met de anderen om;
- b) ze handelen naar algemeen geaccepteerde normen en waarden;
- c) ze respecteren verschillen in levensbeschouwing en cultuur;
- d) ze durven in de groep voor hun eigen standpunt uit te komen;
- e) ze houden rekening met gevoelens en wensen van anderen;
- f) ze durven in de groep steun te geven aan iemand met een afwijkend standpunt;
- g) ze nemen verantwoordelijkheid voor te verrichten taken.

Naast de kerndoelen voor het basisonderwijs zijn er ook kerndoelen voor het speciaal onderwijs opgesteld (Ministerie van OCW, 2009). In deze doelen zijn nog wel leergebiedoverstijgende doelen opgenomen.

Voor leren leren luidt het kerndoel als volgt:

“De leerlingen leren belangstelling hebben voor de wereld om hen heen, ze leren deze gemotiveerd onderzoeken en daarin taken uitvoeren, waarbij ze gebruikmaken van informatie, strategieën en vaardigheden en ze leren reflecteren op eigen handelen.”

De CED-groep heeft de leergebiedoverstijgende kerndoelen uitgewerkt in de leergebiedoverstijgende leerlijn voor leren leren die op veel sbo- en cluster 4-scholen gebruikt wordt (CED, 2009). In bijlage 1 is aangegeven hoe de schalen van de Checklist en de aandachtspunten uit de Beoordelingsformulieren bij de praktische opdrachten zijn gerelateerd aan de leergebiedoverstijgende leerlijn voor leren leren van de CED-groep (2009).

2.2.4 Aansluiting bij bestaande methodieken voor leren leren

Behalve op de hiervoor aangehaalde literatuur is de keuze van relevante vaardigheden gebaseerd op de BOL/EIM-methodiek voor het onderwijzen van vaardigheden van leren leren (Aukes e.a., 2000; Den Besten & Vulpen, 2013; Timmer, Dekker & Voortman, 2003). Genoemde auteurs maken een onderscheid in de denkvaardigheden oriënteren en plannen (nadenken voor je begint en een plan maken), uitvoeren (nadenken terwijl je bezig bent en in de gaten houden of je bereikt wat je wilt bereiken) en evalueren (controleren of je hebt bereikt wat je wilde bereiken en terugkijken op en leren van de manier waarop je het hebt aangepakt).

Daarnaast is de keuze van relevante vaardigheden gebaseerd op de zogeheten Beertjes-methodiek die op veel scholen voor sbo en so cluster 4 gebruikt wordt. Zowel de BOL/EIM- als de Beertjes-methodiek zijn erop gericht de zelfstandigheid en zelfredzaamheid van leerlingen te vergroten. Kern van de Beertjes-methodiek is de behandelingsgerichte observatie die tot doel heeft om problemen met aandacht en werkhouding zo objectief mogelijk vast te stellen en te begeleiden. Vergelijkbaar met de BOL/EIM-methodiek is er sprake van een integratie van meten en instructie: problemen met de cognitieve vaardigheden van leren leren en de werkhouding worden al tijdens het observeren aangepakt en niet alleen ervoor of erna (vgl. Timmerman, 2002, 2017).





De denkvaardigheden oriënteren, plannen, uitvoeren en evalueren uit de BOL/EIM-methodiek vertonen grote gelijkenis met de fasen van de Beertjes-methodiek (Meichenbaum, 1981; Timmerman, 2017):

- 1) De voorbereidingsfase, waarin de opdracht gegeven wordt. Het kind stelt zichzelf de vraag: "Wat moet ik doen?"
- 2) De planningsfase, waarin het kind nadenkt over de oplossingsstrategie. De vraag is: "Hoe ga ik het doen? (Ik maak een plan.)"
- 3) De uitvoeringsfase, waarin het kind de opdracht uitvoert. "Ik doe mijn werk."
- 4) De evaluatiefase, waarin de leerling terugkijkt op zijn werk. De vraag is: "Ik kijk na: wat vind ik ervan?"

Net als in de BOL/EIM-methode is de instructie erop gericht de leerling bewust te maken van de handelingen die nodig zijn om nieuwe taken tot een goed einde te brengen. De leerling leert het handelen bewust te sturen en minder afhankelijk te worden van derden. Door goed te observeren ziet de leerkracht wat de leerling doet en hoe deze te werk gaat (Timmerman, 2002, 2017). Tijdens het observeren verzamelt de leerkracht aanknopingspunten voor de leergesprekjes die tot doel hebben het leren leren verder te ontwikkelen en de gewenste transfer van het aangeleerde naar nieuwe situaties en taken te bewerkstelligen. Net als in de BOL/EIM-methodiek is het leergesprek een belangrijk didactisch middel waarbij leerlingen door middel van gerichte vragen gestimuleerd worden hun denkstappen te expliciteren. Overeenkomstig de denkvaardigheden van het vier-fasen model leren leerlingen zichzelf de volgende vragen te stellen (Meichenbaum, 1981; Timmerman, 2002, 2017; Van Bon, 2009):

- 1 Wat moet ik doen?
- 2 Hoe ga ik het doen? (Ik maak een plan.)
- 3 Ik doe mijn werk.
- 4 Ik kijk na: wat vind ik ervan?

In de Beertjes-methodiek wordt het zichzelf leren stellen van de vier vragen ondersteund door eenvoudige tekeningen, pictogrammen, kernwoorden en korte beschrijvingen. Figuur 1 laat zien hoe de vier stappen van de taakuitvoering en het oplossingsproces met behulp van beertjesplaatjes voor leerlingen inzichtelijk gemaakt worden. Ook zijn er kaarten voor activiteiten zoals stop, denk, praat en doen. Andere didactische hulpmiddelen zijn het stoplicht (rood, oranje, groen), de dobbelsteen ofwel het vragenblokje (rood, vraagteken, groen), het vraagkaartje (gevisualiseerd door een vraagteken), zogenoemde denkmantjeskaarten om de juiste denkhandelingen aan te leren bij het uitvoeren van een taak of het oplossen van een probleem, de time-timer die laat zien hoe lang de leerling nog aan de opdracht kan werken, et cetera.

Taaksituatie	Verbale begeleiding	Visuele ondersteuning	Bedoeling
1 De fase waarin de taak gegeven wordt.	Wat moet ik doen?		De opdracht komt auditief of visueel binnen en het kind denkt na over de informatie die het krijgt. Het analyseert wat het moet doen.
2 De fase waarin het kind nadenkt over een oplossingsstrategie.	Hoe ga ik dat doen?		Het kind denkt na over een oplossingsmethode om de opdracht uit te voeren.
3 De fase waarin het kind de opdracht uitvoert.	Ik doe mijn werk.		Met het wat en het hoe in zijn hoofd, voert het kind de opdracht uit. Het werk systematisch, doelgericht en in een gelijkmatig tempo.
4 De fase waarin het kind het werk evalueert.	Ik kijk na: wat vind ik ervan?		Het kind kijkt de oplossing na en stuurt bij als het nodig is.

Bron: Timmerman, K. (2017). *Aandachts- en werkhoudingsproblemen bij kinderen en jongeren*. Leuven: Acco, blz. 72.

Figuur 1 Visualisering van de vier fasen van het zelfinstructiemodel uit de Beertjesmethodiek (ontleend aan Timmerman, 2017, blz. 72)

Het pakket Leren leren sluit ook goed aan bij recente handreikingen voor het onderwijzen van executieve functies (o.a. Cooper-Kahn & Dietzel, 2011; Dawson & Guare, 2009, 2010; Smidts, 2018; Smidts & Huizinga, 2011, 2017). Deze publicaties bevatten een breed scala aan vuistregels en tips voor het onderwijzen van de zelfregulerende vaardigheden die in het pakket Leren leren centraal staan. In de gebruikershandleiding worden de geïnteresseerde leerkrachten dan ook veelvuldig naar deze literatuur verwezen.

2.2.5 Dynamic assessment

Onder toetsvorm verstaat Straetmans (2006) "... het concrete pakket van regels en procedures dat voorschrijft hoe gedrag wordt uitgelokt, gescoord en geëvalueerd" (p. 41). In de literatuur over moderne vormen van toetsing wordt vaak een onderscheid gemaakt tussen statische en dynamische vormen van toetsing, hierna aangeduid met de gebruikelijke termen *static assessment* (SA) en *dynamic assessment* (DA). DA is een nog vrij onbekende vorm van *assessment for learning* (Black & Williams, 1998a,b) die Cito één keer eerder heeft toegepast in het pakket Leren leren voor VSO-leerlingen Dagbesteding en Arbeid. In deze paragraaf verantwoorden we de keuze voor DA als een bij sbo- en so cluster 4-leerlingen passende manier om vaardigheden van leren leren te toetsen en verder te ontwikkelen.

Static assessments

Tot op heden zijn voor het speciaal onderwijs voornamelijk zogenoemde *static assessments* (SA) ontwikkeld. Met Tzurriel (2001) verstaan we onder een SA “a test where the examiner presents items to a child and records his or her response without an attempt to intervene in order to change, guide or improve the child’s performance” (p. 1). Enkele belangrijke kenmerken van een SA zijn:

- Een SA meet wat de leerling van het voorafgaande onderwijs heeft opgestoken; de toetsing geeft weinig informatie over deficiënt cognitief functioneren dat verantwoordelijk is voor leerproblemen en de onderwijsstrategieën om leren te faciliteren.
- Een SA is doorgaans landelijk genormeerd: de leerkracht vergelijkt de prestatie van een leerling met een landelijke steekproef van leeftijdsgenoten of met leerlingen in hetzelfde leerjaar.
- Een SA wordt volgens een strikt protocol afgenomen, waarbij de leerkracht de leerling tijdens de afname geen instructie mag geven.
- Een SA geeft geen informatie over niet-intellectuele factoren die van invloed zijn op de toetsprestaties.

Tzurriel (2001) noemt verschillende redenen waarom SA voor bepaalde groepen leerlingen minder geschikt kunnen zijn. Hierna noemen we enkele redenen die mogelijk gelden voor onze doelgroep:

- Er zijn aanwijzingen dat *static assessments* de prestaties van bepaalde groepen leerlingen systematisch onderschatten (Kaniel e.a., 1991; Resing, 2006; Scheltinga, Keuning & Kuhlemeier, 2014). Tot die groepen behoren leerlingen met een achterstand in de instructie- en toetstaal, leerlingen die door culturele factoren minder gelegenheid hebben gehad om zich de getoetste kennis en vaardigheden eigen te maken, leerlingen die minder vertrouwd zijn met de eisen die de afnamesituatie aan hen stelt en leerlingen die vanwege cognitieve, sociaal-emotionele en/of fysieke beperkingen extra hulp nodig hebben. Voor leerlingen in het speciaal onderwijs zijn tests die zij geheel zelfstandig moeten maken vaak lastig. Het kost hun vaak moeite om zich het taalgebruik en de procedures van gestandaardiseerde toetsen eigen te maken. Vooral als er een tijdslimiet op de toets zit, presteren zij vaak onder hun mogelijkheden wat hun zelfbeeld en motivatie nadelig kan beïnvloeden. De gebruikelijke toetsen zijn voor deze risicoleerlingen vaak te moeilijk waardoor ze de toekomstige schoolprestaties slecht voorspellen. Vermoed wordt dat zij daardoor nogal eens ten onrechte in speciale onderwijsprogramma’s terechtkomen, terwijl zij met passende hulp en begeleiding gewoon in het reguliere onderwijs hadden kunnen blijven (Scheltinga, Keuning & Kuhlemeier, 2014). Daardoor representeert een zwakke prestatie op een SA niet noodzakelijkerwijs ook een lage vaardigheid (Kester, Peña & Gillam, 2001; Peña, Quinn & Iglesias, 1992; Tzurriel, 2001).
- Een deel van de leerlingen in het speciaal onderwijs is tot op zekere hoogte hulpbehoevend. Sommigen zullen gedurende hun hele leven tenminste enige hulp en begeleiding nodig hebben (ook al zien we de laatste jaren een verschuiving van de begeleiding van ‘zorgen voor’ naar ‘zorgen dat’). De vraag is dan hoe relevant het is om te weten wat een leerling in een kunstmatige testsituatie geheel zelfstandig kan. Veel belangrijker lijkt het om te weten wat het leerpotentieel van de leerling is, dit wil zeggen wat een leerling met hulp van een betrokken en deskundige volwassene in een natuurlijke setting kan bereiken en hoeveel en welke hulp en begeleiding daarvoor in het vervolgtraject nodig zijn (Feuerstein, 1979; Tzurriel, 2001).
- Tussen leerlingen in het speciaal onderwijs bestaan grote verschillen in hun cognitieve, sociaal-emotionele en/of fysieke mogelijkheden en de benodigde hulp en begeleiding (zie paragraaf 2.1.2). Vergelijking met de gemiddelde leerling van een landelijke steekproef van leeftijds- of jaargenoten is daardoor minder zinvol dan voor bijvoorbeeld leerlingen in het regulier primair onderwijs (Van Boxtel, Hollenberg, Horsels & Kleintjes, 2012).
- Onderzoek heeft laten zien dat feedback een belangrijk middel is om de leerprestaties te bevorderen (Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Sluismans, Joosten-ten Brinke & Van der Vleuten, 2013). Gezien hun kortere spanningsboog hebben leerlingen in het speciaal onderwijs extra behoefte aan onmiddellijke, gerichte en gepersonaliseerde feedback op hun toetsprestaties. De afname van een SA biedt de leerkracht echter weinig informatie over wat er precies fout gaat in het leerproces en welke strategieën nodig zijn om het leren verder te ontwikkelen (Resing, 2006). Onmiddellijke feedback is in een SA niet mogelijk, omdat de leerkracht de leerling tijdens de afname niet mag instrueren of helpen.

Dit pleit ervoor om toetsing, onderwijzen en leren waar mogelijk en zinvol te integreren (Black & William, 1998a,b).

- Behalve op cognitief gebied hebben leerlingen in het speciaal onderwijs ook behoefte aan hulp en ondersteuning op sociaal-emotioneel en gedragsmatig gebied. SA biedt daarover geen informatie. Dit pleit voor een vorm van toetsing waarmee de leerkracht cognitieve en non-cognitieve vaardigheden in onderlinge samenhang kan vaststellen.

Om de hiervoor genoemde redenen hebben wij gekozen voor *dynamic assessment* (DA).

Dynamic assessment

Tzuriel (2001) geeft de volgende definitie van *dynamic assessment*: “*DA refers to an assessment of thinking, perception, learning, and problem solving by an active teaching process aimed at modifying cognitive functioning*”(p. 6). Op het gebied van de intelligentietoetsing is DA al veelvuldig met succes toegepast (Resing, 1993, 1997, 2006; Van der Aalsvoort, Resing & Ruijsenaars, 2002). In het domein van de vaardigheidstoetsing in het onderwijs is DA echter een nog relatief onbekend en weinig toegepast fenomeen. DA vertegenwoordigt een toekomstgerichte benadering van toetsing die gericht is op het exploreren van de zone van naaste ontwikkeling. Daaronder verstaat Vygotsky (1978) de kloof tussen “*the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers*” (p. 86). Een SA geeft informatie over de ondergrens van deze zone ofwel wat een leerling kan zonder de hulp van anderen. Een DA exploreert daarentegen ook de bovengrens en geeft hiermee een indicatie van het leerpotentieel, ook wel leerbaarheid, leervermogen of veranderbaarheid genoemd (Tzuriel, 2001). Hier gaat het om de mate waarin de leerling in staat is om te profiteren van de instructie, hulp en begeleiding van de leerkracht. Ofwel, wat de leerling in het vervolgtraject kan bereiken als er passende instructie, hulp en begeleiding gegeven zou worden. Hieronder gaan we nader in op enkele basale kenmerken van DA die in het pakket Leren leren geïmplementeerd zijn:

- Toetsen in dienst van het leren
- Verkennen van leervermogen of leerpotentieel
- Het train-within-test ontwerp
- De instructie en begeleiding tijdens de afname
- Integratie van (meta)cognitief en sociaal-emotioneel leren
- Observatie tijdens de afname

Toetsen in dienst van het leren

In een *dynamic assessment* staat de toetsing geheel in dienst van het leren. Tijdens de afname zoekt de leerkracht – via zorgvuldige en betrokken interacties met de leerling – naar oorzaken die ten grondslag liggen aan (meta)cognitieve en sociaal-emotionele problemen die het leren kunnen belemmeren. Een DA verschaft informatie over het leren van de leerling, zoals over de manier waarop de leerling de taak aanpakt, de fouten die hij of zij daarbij maakt en de vaardigheid om zichzelf te corrigeren. Daarnaast geeft de interactie met de leerling waardevolle informatie over wat er nodig is om het leren te verbeteren. Het belangrijkste verschil met SA is dat de leerkracht de leerling als dat nodig is tijdens de afname mag helpen, bijvoorbeeld door feedback te geven, oplossingsstrategieën aan te rijken of transfertaken aan te bieden. Het grote voordeel van DA is dat het procesgerichte informatie oplevert die direct gerelateerd is aan het voorafgaande onderwijs, via directe communicatie tussen leerkracht en leerling tot stand is gekomen en daardoor onmiddellijk bruikbaar is ter verbetering van het leerproces.

Verkennen van leervermogen of leerpotentieel

Een DA is ook gericht op het verkennen van het leervermogen of het leerpotentieel (Lidz, 1987; 1991; Sternberg & Grigorenko, 2002; Tzuriel, 2001). Daaronder verstaan we het niveau dat de leerling met hulp van een volwassene kan bereiken (Vygotsky, 1987). De toetsing kijkt niet alleen terug, maar vooral vooruit. De rapportage laat niet alleen zien wat de leerling al wel en nog niet kan, maar ook hoeveel en welke instructie, hulp en begeleiding de leerkracht heeft moeten geven om de ontwikkelingsdoelen te bereiken. De hoeveelheid hulp die de leerling daarbij nodig had, is een soort omgekeerde maat voor het

leerpotentieel (Resing, 2006). Hoe meer hulp de leerkracht de leerling tijdens de afname moet geven, hoe lager diens leerpotentieel. Leerlingen die weinig instructie nodig hebben maar toch sterk vooruitgaan, worden geacht over een groot leerpotentieel te beschikken. Bij leerpotentieel gaat het om de maximale prestatie die de leerling onder ideale omstandigheden kan behalen (Resing, 2006). Niet de gemiddelde prestatie, maar de piekprestatie die de leerling over een reeks van opdrachten behaalt, is de beste indicator voor het leerpotentieel (Tzuriel, 2001). Kortom, een DA geeft een indicatie van de mate waarin de leerling in staat is om te profiteren van de hulp van de leerkracht, het niveau dat de leerling in het vervolgtraject zou kunnen bereiken en hoeveel en welke instructie en begeleiding daarvoor in de toekomst nodig is (Lidz & Peña, 1996).

Het *train-within-test model*

Met betrekking tot het onderwijskundige ontwerp van een DA onderscheiden Sternberg en Grigorenko (2002) een sandwichmodel en een cakemodel. In het geval van het sandwichmodel wordt de toetsing uitgevoerd volgens een voormeting-instructie-nameting ontwerp. Met de voormeting wordt eerst de baseline score vastgesteld. Daarbij maakt de leerling de toets geheel zelfstandig, dus zonder hulp van de leerkracht (zodat er in feite sprake is van een SA). Daarna volgt een (korte) interventie van onderwijs en leren, vaak aangeduid als de fase van gemedieerde leerervaring, waarbij bijvoorbeeld de leerkracht de zogeheten mediator is. De mediatie bestaat uit instructie, met bijvoorbeeld hints en aanwijzingen, die wordt gegeven tot het ontwikkelingsdoel bereikt is en de leerling de taak zelfstandig kan uitvoeren (Resing, 2006). Na de fase van instructie volgt de natest, die de leerling opnieuw geheel zelfstandig moet maken.

In het zogeheten *cakemodel* is er geen sprake van een voor- en nameting. De hoeveelheid en aard van de instructie wordt tijdens de afname op adaptieve wijze gevarieerd al naar gelang de hulpbehoefte van de leerling. Het *cakemodel* wordt daarom ook wel het *train-within-test model* genoemd. Dit ontwerp is vooral geschikt voor leerlingen die moeite hebben met het zelfstandig maken van toetsopdrachten en bij wie een statische voor- of nameting tot een onderschatting van hun prestaties zou leiden. Vandaar dat wij voor het toetsen en onderwijzen van leren leren gekozen hebben voor het *train-within-test model* zonder voor- of nameting.

De instructie en begeleiding tijdens de afname

De kern van een DA is de instructie tijdens de afname van de test, veelal aangeduid als de mediatie of de gemedieerde leerervaring (Grigorenko & Sternberg, 1988). Naar de wijze waarop de mediatie het beste kan plaatsvinden, is al veel ontwerponderzoek verricht (Budoff, Meskin & Harrison, 1971; Campione, 1989; Carlson & Wiedl, 1978; Feuerstein, 1979; Lidz, 1987, 1991; Sternberg & Grigenko, 2002; Swanson & Lussier, 2001; Tzuriel, 2001). Een van de ontwerpprincipes is dat de leerkracht niet méér instructie geeft dan voor een zelfstandige uitvoering van de taak vereist is. Een kleine hint kan namelijk al voldoende zijn om het leerpotentieel te activeren. Daarnaast is het van belang dat de hoeveelheid hulp en de mate van expliciete instructie langzaam toeneemt al naar gelang de behoefte van de leerling. Idealiter wordt de mediatie gegeven tot het leer- of ontwikkelingsdoel bereikt is en de leerling de taak zelfstandig kan uitvoeren (Resing, 2006). Bij het pakket Leren leren is gekozen voor mediatie volgens de bewezen effectieve BOL/EIM-methodiek voor het onderwijzen van leren leren bij leerlingen in het speciaal onderwijs (Aukes e.a., 2000; Den Besten & Vulpen, 2013; Timmer, 2004; Timmer, Dekker & Voortman, 2003). Ook is aansluiting gezocht bij de Beertjes-methodiek (Meichenbaum, 1981; Timmerman, 2002; Van Bon, 2009) en recente literatuur over executieve functies (o.a. Cooper-Kahn & Dietzel, 2011; Dawson & Guare, 2009, 2010; Smidts, 2018; Smidts & Huizinga, 2011, 2017). Voor een nadere uitwerking verwijzen wij naar de gebruikershandleiding (Cito, 2018).

Aangezien de instructie optimaal wordt afgestemd op de specifieke leerproblemen van de leerling, biedt sterk geïndividualiseerde mediatie met onmiddellijke feedback en remediëring de grootste kans op verbetering van het leren. Een nadeel is dat de beperkte standaardisatie van de afnamecondities ten koste kan gaan van de objectiviteit en betrouwbaarheid van de meting. Een ander nadeel is dat de afname vaak bewerkelijker en tijdrovender is dan bij een SA het geval is.

Integratie van (meta)cognitief en sociaal-emotioneel leren

Leren leren heeft zowel een metacognitieve als non-cognitieve component. Een SA biedt geen informatie over niet-intellectuele factoren die het leren in de weg kunnen staan. DA is daarentegen een holistische benadering die mikt op het vaststellen en verbeteren van het cognitieve functioneren én het sociaal-emotionele en gedragsmatige functioneren in onderlinge samenhang.

Observatie tijdens de afname

Zelfregulerend lerende leerlingen zijn in staat om zelfregulerende en non-cognitieve vaardigheden toe te passen als zij tijdens het uitvoeren van taken geconfronteerd worden met tegenslag. In een DA is gedragsobservatie een belangrijke methode om hierover informatie in te winnen. De observatie wordt doorgaans uitgevoerd met observatielijsten. Deze meten doorgaans geen vakspecifieke vaardigheden, maar allerlei ondersteunende vaardigheden die van belang zijn voor het verkennen van het leerpotentieel, zoals de hoeveelheid hulp en begeleiding die de leerkracht heeft moeten geven tot de leerling de taak zelfstandig kan uitvoeren en de transfer naar andere dan de geoefende vaardigheden, taken en situaties (Lidz, 1987, 1991; Tzuriel, 2001). Voor de wijze waarop de observatie in het pakket Leren leren is uitgewerkt, verwijzen wij naar paragraaf 3.3.3 en de gebruikershandleiding (Cito, 2018).

3 Constructie en beschrijving van het pakket Leren leren

Zoals aangegeven in het vorige hoofdstuk bestaat de kern van het pakket Leren leren uit een checklist, een serie praktische opdrachten met bijbehorende beoordelingsformulieren en handreikingen. In het voorgaande hoofdstuk zijn we onder meer ingegaan op de keuze van de vaardigheden van leren leren en de toetsvorm (*dynamic assessment* volgens het *train-within-test model*). In dit hoofdstuk noemen we eerst enkele algemene overwegingen die ten grondslag lagen aan de constructie van de praktische opdrachten, beoordelingsformulieren en handreikingen (paragraaf 3.1). Vervolgens beschrijven we de manier waarop we het materiaal overeenkomstig deze uitgangspunten hebben ontwikkeld (paragraaf 3.2).

Daarna beschrijven we de opbouw, afname, beoordeling en rapportage van het uiteindelijke instrument en de handreikingen voor het onderwijzen van leren leren (paragraaf 3.3). De uitvoerbaarheid, betrouwbaarheid en validiteit staan centraal in het volgende hoofdstuk (hoofdstuk 4).

3.1 Uitgangspunten en overwegingen bij de constructie

Criteriumgerichte benadering

De standaard toetsen van het Cito Volgsysteem zijn landelijk genormeerd waardoor de vaardigheid van een leerling vergeleken kan worden met de vaardigheid van leeftijdgenoten. Door de grote variatie aan leerlingen binnen het speciaal onderwijs is dit voor deze doelgroep minder relevant. Daarom is er voor de meting van de praktische vaardigheden gekozen voor een *criteriumgerichte toets* in plaats van een *normgerichte toets*. Met de praktische opdrachten uit pakket Leren leren kan de leerkracht vaststellen op welk niveau de leerling de vaardigheden van leren leren beheerst en hoeveel instructie en hulp de leerling daarvoor nodig heeft gehad.

Praktische in plaats van schriftelijke opdrachten

Van Boxtel, Hollenberg, Horsels en Kleintjes (2012) onderzochten de voorkeur van leerkrachten in het speciaal onderwijs voor pen-en-papier toetsen dan wel praktische opdrachten. Schriftelijke toetsen geven volgens leerkrachten onvoldoende zicht op wat deze leerlingen daadwerkelijk kunnen (bereiken). Leerkrachten zijn van mening dat dergelijke toetsen hun prestaties onderschatten. Leerlingen uit de doelgroep komen beter tot hun recht in natuurlijke situaties waarin zij werken met concrete materialen dan in situaties waarin ze aan de hand van opgaven op papier moeten laten zien wat ze kunnen. Het is mede om deze reden dat wij voor toetsend leren aan de hand van praktische opdrachten hebben gekozen. Een andere reden is dat men zelfregulerende vaardigheden beter kan observeren in een natuurlijke omgeving aan de hand van praktische opdrachten dan met artificiële pen-en-papier toetsen.

Individuele in plaats van groepsgewijze of klassikale afname, observatie en instructie

De praktische opdrachten zijn bestemd voor individuele afname en niet geschikt voor groepsgewijze of klassikale afname. Tussen sbo- en so cluster 4-leerlingen bestaan grote verschillen in hun cognitieve, sociaal-emotionele en/of fysieke mogelijkheden die het leren leren kunnen faciliteren dan wel belemmeren. Een van de redenen voor de keuze voor individuele afname is dat men de leerling in een één-op-één situatie beter kan observeren, beoordelen en vooral verder helpen dan wanneer de opdracht groepsgewijs of klassikaal zou worden afgenomen. Een individuele benadering biedt meer mogelijkheid tot toetsing in dienst van het leren dan een groepsgewijze of klassikale benadering waarbij alle leerlingen over één kam geschoren worden en er niet op specifieke problemen van individuele leerlingen ingegaan kan worden. Een andere reden is dat groepsgewijze afname van een praktijkopdracht tot imitatiegedrag (afkijken, nadoen) kan leiden wat afbreuk doet aan de betrouwbaarheid en validiteit van de toetsing.

Afname, observatie en instructie door de 'eigen' leerkracht

Iedereen die specifieke kennis en vaardigheden op het gebied van het ontwikkelen van leren leren heeft en de leerling goed kent, kan de opdrachten afnemen (Den Besten & Vulpen, 2013).

De leerkracht of eventueel de intern begeleider is de meest geschikte persoon om vaardigheden van leren leren en het leerpotentieel vast te stellen. Zij werken immers dag in dag uit met de leerlingen samen en kennen hen door en door. Dit is vooral van belang als de communicatie gehinderd wordt door verstandelijke en andere beperkingen, zoals een slechte verstaanbaarheid. Als het de leerkracht niet lukt om de problemen voldoende te benoemen en te verhelpen, is het verstandig deskundige hulp van buitenaf in te roepen, zoals een orthopedagoog, leerpsycholoog of gedragswetenschapper.

3.2 De constructie van het pakket Leren leren

Fasering van de constructie van de checklist

De keuze van de inhoud en structuur van de checklist is gebaseerd op het hiervoor besproken theoretisch en onderwijskundig kader (zie paragraaf 2.2). De cognitieve vaardigheden van leren leren zijn daarbij ingedeeld al naar gelang de fase in het leerproces waarin de leerling ze kan toepassen: Oriënteren, voorbereiden en plannen, Uitvoeren en Monitoren, reflecteren en evalueren (Butler & Winne, 1995; Dawson & Guare, 2009, 2010; Den Besten & Vulpen, 2013; Meichenbaum, 1981; Pintrich & Zusho, 2002; Smidts, 2018; Smidts & Huizinga, 2011, 2017; Timmerman, 2002; Van Bon, 2009; Zimmermann, 1990; 2002; Zumbrunn, Tadlock & Roberts, 2011). Daarnaast is in de checklist een plaats ingeruimd voor non-cognitieve vaardigheden van leren leren (werkhouding): aandachtig werken, gemotiveerd werken, werken zonder gedragsproblematiek en zelfstandig werken en hulp vragen.

De constructie van de checklist verliep volgens de volgende fasen:

- Allereerst is een constructie-opdracht geschreven met onder meer informatie over het doel, de structuur en de inhoud van de checklist en aanwijzingen voor het formuleren van uitspraken.
- Vervolgens heeft de constructiegroep (hierna cg-groep) met 4 ervaren leerkrachten onder leiding van 2 toetsdeskundigen voor elk van de onderscheiden aspecten van leren leren een groot aantal uitspraken geformuleerd.
- De uitspraken zijn binnen de cg-groep besproken en op basis van de opmerkingen en suggesties bijgesteld en in sommige gevallen geschrapt.
- De overgebleven uitspraken zijn door de cg-leden tijdens een mini-pilot uitgeprobeerd. Op basis van de ervaringen van de cg-leden zijn de uitspraken verder aangescherpt.
- Uit de omvangrijke set uitspraken zijn 87 uitspraken gekozen die de toets der kritiek hadden doorstaan.
- In het voorjaar van 2017 is een proefafname gehouden waarbij de conceptversie van de checklist bij een grotere groep leerkrachten is uitgetest op bruikbaarheid en enkele aspecten van betrouwbaarheid en validiteit. 68 leerkrachten van 17 scholen vulden de checklist in voor in totaal 629 leerlingen uit 53 groepen. Daarvan zaten 226 leerlingen op een sbo-school en 403 leerlingen op een so cluster 4-school. De versie voor de proefafname bevatte 87 uitspraken die verdeeld waren over 7 a priori schalen, te weten
 - Oriënteren, voorbereiden en plannen van praktische taken;
 - Uitvoeren van praktische taken;
 - Monitoren, reflecteren en evalueren van praktische taken;
 - Aandachtig werken;
 - Gemotiveerd werken;
 - Werken zonder gedragsproblematiek;
 - Zelfstandig werken en hulp vragen.

In deze proefafname is nagegaan in hoeverre de uitspraken geschikt zijn voor afname bij sbo- en so cluster 4-leerlingen. Daarnaast is informatie verzameld om de checklist te verbeteren, onder andere door bestudering van psychometrische kenmerken met betrekking tot de verdeling van de scores, betrouwbaarheid en validiteit.

- Op basis van de resultaten van de proefafname is de definitieve versie van de checklist samengesteld die is beschreven in hoofdstuk 4. Daarbij is ook gebruikgemaakt van de waardevolle opmerkingen en suggesties die de proefafnameleerkrachten op het evaluatieformulier verstrekten.

De checklist is onverkort opgenomen in bijlage 2.

Fasering van de constructie van de praktische opdrachten

Het pakket Leren leren bestaat uit tien praktische opdrachten. Per opdracht is er een gemakkelijke A-variant, een wat moeilijkere B-variant en een nog moeilijkere C-variant. Onder leiding van de 2 toetsdeskundigen ontwikkelden 4 ervaren leerkrachten de opdrachten en varianten. De toetsdeskundigen werden ondersteund door een collega met ruime ervaring in het ontwikkelen van praktische opdrachten voor de centraal schriftelijke praktijkexamens in het vmbo (CSPE).

Bij de ontwikkeling van de praktische opdrachten zijn we niet over een nacht ijs gegaan. De constructie van een opdracht verliep volgens de volgende fasen:

- Allereerst is een overzicht opgesteld van de opdrachtkenmerken die bepalend zijn voor de complexiteit van de praktische opdrachten en hiermee voor de moeilijkheidsgraad voor de leerlingen. Dit overzicht is deels gebaseerd op ervaringen die zijn opgedaan tijdens de constructie van praktische opdrachten van de praktijkexamens voor de basisberoepsgerichte leerweg van het vmbo en het pakket Leren leren voor VSO Dagbesteding en Arbeid (Kuhlemeier & Chermin, 2015).
- Vervolgens is een constructieopdracht geschreven die met de leden van de constructiegroep (cg) werd doorgesproken en op grond van het overleg werd bijgesteld.
- Daarna maakten de cg-leden een eerste versie van de praktische opdracht met bijbehorende afnamekaart (dit wil zeggen: de instructie voor de afname van de praktische opdracht). Beide werden in een tweede cg-vergadering besproken en op grond van de bevindingen bijgesteld.
- Vervolgens werd de praktische opdracht met bijbehorende beoordelingsformulieren en afnamekaart door de cg-leden in een zogeheten mini-pilot bij 'eigen' leerlingen uitgetest (inclusief beoordeling), binnen de constructiegroep besproken en op basis van de ervaringen bijgesteld. Zo nodig volgde een tweede mini-pilot, waarbij de cg-leden elkaars praktische opdrachten uitprobeerden. Dit resulteerde in een versie die geschikt geacht werd voor gebruik door derden die niet bij het project betrokken waren.
- Daarna is een proefafname gehouden waarbij 2 praktische opdrachten met ieder 3 varianten bij een groep leerkrachten is uitgetest op bruikbaarheid en enkele aspecten van betrouwbaarheid en validiteit. Aan deze proefafname deden 21 leerkrachten van 12 verschillende scholen mee. In totaal is bij 115 leerlingen een praktische opdracht afgenomen. Daarvan zaten 62 leerlingen op een sbo-school en 53 leerlingen op een so cluster 4-school.
- Op basis van de uitkomsten van deze proefafname werden de opdrachten en beoordelingsformulieren nogmaals bijgesteld, inclusief de handleiding voor leerkrachten. De proefafname resulteerde in het definitieve instrumentarium dat we beschrijven in paragraaf 4.

De constructie van nominaal parallelle opdrachten en varianten

Bij de constructie is ernaar gestreefd om de overeenkomstige varianten van verschillende opdrachten ongeveer even moeilijk te maken. Zo is geprobeerd om de gemakkelijke A-variant van de opdracht *Meten* ongeveer even moeilijk te maken als de gemakkelijke A-variant van *Powerpoint*. Hetzelfde geldt ook voor de moeilijkere B- en C-varianten. De overeenkomstige varianten van verschillende opdrachten moeten met andere woorden 'nominaal parallel' zijn. De opdrachten en varianten verschillen onder meer qua vakgebied en onderwerp, maar moeten wel ongeveer even moeilijk zijn. Om dit laatste te realiseren hebben we eerst een overzicht opgesteld van de kenmerken die naar verwachting bepalend zijn voor de complexiteit van de opdracht. Vervolgens hebben we geprobeerd de overeenkomstige varianten van verschillende opdrachten ongeveer even moeilijk te maken door deze opdrachtkenmerken systematisch te variëren. Het schema op de volgende pagina toont de te variëren kenmerken zoals deze met de cg-leden zijn doorgesproken voordat zij met de constructie van de A, B- en C-varianten begonnen.

Opdrachtkenmerk	Complexiteit / moeilijkheidsgraad		
	A - Gemakkelijk	B - Moeilijker	C – Moeilijk
Introductie op de opdracht	De leerkracht introduceert de opdracht mondeling.	De leerkracht introduceert de opdracht mondeling.	De leerling leest de introductie geheel zelfstandig.
	De leerkracht neemt de opdracht mondeling met de leerling door. De leerling heeft een stappenplan, dat ondersteund wordt door beeldmateriaal (plaatjes + picto's)	De leerling leest de opdracht geheel zelfstandig (uiteraard met gelegenheid tot het stellen van vragen). De leerling heeft een stappenplan, dat ondersteund wordt door beeldmateriaal (plaatjes + picto's)	De leerling leest de opdracht geheel zelfstandig (uiteraard met gelegenheid tot het stellen van vragen). De opdracht is in de vorm van alleen schriftelijke instructies.
	De leerling krijgt waar nodig van het te maken eindproduct een afbeelding te zien.	De leerling krijgt waar nodig van het te maken eindproduct een afbeelding te zien.	De leerling krijgt waar nodig van het te maken eindproduct een afbeelding te zien.
	De leerling krijgt waar nodig van het te maken eindproduct een afbeelding te zien.	De leerling krijgt waar nodig van het te maken eindproduct een afbeelding te zien.	De leerling krijgt waar nodig van het te maken eindproduct een afbeelding te zien.
	Op de tafel staan geen spullen die de leerling niet nodig heeft.	Op de tafel staan ook (veel) spullen die de leerling voor de opdracht NIET nodig heeft.	Op de tafel staan ook (veel) spullen die de leerling voor de opdracht NIET nodig heeft.
Het materiaal	materiaal voorberekt (bijvoorbeeld: bij de opdracht Tekening perspectief zijn de tekenstreepjes op het voorblad voorgedrukt)	materiaal gedeeltelijk voorberekt	materiaal niet voorberekt
	vereenvoudigde beginsituatie	vereenvoudigde beginsituatie	beginsituatie niet vereenvoudigd (geen materiaal of handelingen vereenvoudigd of door leerkracht voorbereid; leerling voert alle handelingen zelf uit)
De uitvoering	klein aantal handelingen en stappen	groter aantal handelingen en stappen	groot aantal handelingen en stappen
	eenvoudige handelingen en stappen	gemiddelde handelingen en stappen	complexere handelingen en stappen
	alle noodzakelijke stappen en handelingen zijn aangegeven	alle noodzakelijke stappen en handelingen zijn aangegeven	slechts een deel van de handelingen en stappen is aangegeven
	stappen geïllustreerd met pictogrammen en tekeningen en eenvoudige tekst	stappen geïllustreerd met pictogrammen en tekeningen en tekst	stappen voornamelijk met tekst aangegeven (geen of weinig pictogrammen of tekeningen)
	vaste volgorde van stappen en handelingen	vaste volgorde van stappen en handelingen	volgorde van stappen en handelingen deels door de leerling zelf te bepalen
	geen schriftelijke verwerkingsopdrachten	ook schriftelijke verwerkingsopdrachten	ook schriftelijke verwerkingsopdrachten (bijv. zelf een lijstje met benodigde spullen opstellen)
Reflecteren en evalueren	voor alle varianten gelijk	Reflecteren en evalueren	voor alle varianten gelijk

In de praktijk bleek het lastig om de varianten volledig conform de constructie-opdracht te construeren. Soms moesten er concessies gedaan worden. In de kleinschalige try-outs bleek bijvoorbeeld dat de opdrachten bij het vakgebied Rekenen voor de leerlingen intrinsiek moeilijker waren dan die voor bijvoorbeeld Tekenen. Stringente doorvoering van de richtlijnen zou er dan bijvoorbeeld toe geleid hebben

dat de A-variant van Rekenen veel moeilijker zou worden dan de A-variant voor Tekenen. In dat geval is geprobeerd het verschil in intrinsieke moeilijkheid te compenseren door de varianten Rekenen wat gemakkelijker te maken. In paragraaf 5.5.3 gaan we aan de hand van twee opdrachten na in hoeverre het daadwerkelijk gelukt is om soortgelijke varianten ongeveer even moeilijk te maken (rekening houdend met de adaptieve toewijzing van varianten aan leerlingen).

3.3 Opbouw, afname, beoordeling, rapportage en instructie

3.3.1 Opbouw

Het pakket Leren leren bevat de volgende onderdelen:

- een checklist;
- tien praktische opdrachten met ieder drie varianten in een verschillende moeilijkheidsgraad;
- per variant één afnamekaart met aanwijzingen voor de voorbereiding en afname van de opdracht;
- vier generieke beoordelingsformulieren;
- twee digitale rapportagemodules (in Excel) voor de leerling- en groepsrapportage van de resultaten;
- een algemene gebruikershandleiding en twee gebruikershandleidingen bij de rapportagemodules.

De checklist

De checklist heeft tot doel snel een indruk te verkrijgen van de vaardigheid in leren leren. Het is vooral een hulpmiddel om te bepalen welke vaardigheden extra (klassikale) aandacht vereisen en bij welke vaardigheden dat niet of minder nodig of wenselijk is. Dit maakt de checklist geschikt voor leerkrachten die de vaardigheden van leren leren van hun leerlingen snel en globaal in kaart willen brengen of die vanwege tijdgebrek of andere redenen de praktische opdrachten niet individueel kunnen of willen afnemen.

Daarnaast is de checklist een hulpmiddel om de leerlingen te selecteren die het meest voor afname van een arbeidsintensieve praktische opdracht in aanmerking komen. Hiermee worden leerkrachten bediend die met de praktische opdrachten aan de slag willen, maar dat niet bij alle leerlingen uit hun groep willen of kunnen doen.

Het invullen van de checklist neemt per leerling vijf à tien minuten in beslag.

De praktische opdrachten met bijbehorende beoordelingsformulieren

De uiteindelijke uitgave kent tien praktische opdrachten. Er zijn vijf vakgebieden in het materiaal vertegenwoordigd: rekenen, tekenen, zaakvakken, schoolse vaardigheden (computer) en techniek. Voor elk vakgebied zijn er twee gelijkwaardige praktische opdrachten ontwikkeld. Zo is er voor het vakgebied Rekenen een opdracht Meten en een opdracht Plattegrond. De reden om twee opdrachten per vakgebied te ontwikkelen is dat de leerkracht hiermee kan nagaan in hoeverre de leerlingen de algemene vaardigheden die zijn aangeleerd met de ene opdracht ook kan toepassen bij een andere opdracht voor hetzelfde vakgebied (nabije transfer). Een van de redenen om opdrachten voor verschillende vakgebieden te ontwikkelen is dat de leerkracht hiermee kan nagaan in hoeverre de leerling een aangeleerde vaardigheid ook bij een geheel ander vakgebied kan toepassen (verre transfer). De tien opdrachten zijn weergegeven in onderstaand overzicht.

Vakgebied	Opdracht met elk een A-, B- en C-variant	Parallelopgdracht met elk een A-, B- en C-variant
Rekenen	Meten	Plattegrond
Tekenen	Tekening coördinaten	Tekening perspectief
Zaakvakken	Tijdslijn Geschiedenis	Topografie Aardrijkskunde
Schoolse vaardigheden (computer)	Powerpoint	Reisplan
Techniek	Stroomkring	Spijkerplank

De twee techniekopdrachten zijn niet in de fysieke uitgave opgenomen. Voor de uitvoering van deze twee opdrachten is relatief veel materiaal nodig dat niet per se op elke school aanwezig is. De leerkracht kan deze twee 'bonusopdrachten' desgewenst downloaden van Cito Portal (<http://portal.cito.nl>).

Drie varianten per opdracht

Het pakket bevat van elke opdracht een makkelijke A-variant, een wat moeilijkere B-variant en een moeilijke C-variant. De constructie van telkens drie varianten maakt het mogelijk de afname af te stemmen op het vaardigheidsniveau van de leerling. Tijdens de constructie is ernaar gestreefd om de parallele varianten van verschillende opdrachten even moeilijk te maken. Hoe we daarbij te werk zijn gegaan, is eerder uiteengezet in paragraaf 3.2.

3.3.2 Afname

Algemene aanwijzingen voor de afname van de checklist

De papieren gebruikershandleiding (Cito, 2018) bevat uitgebreide aanwijzingen voor de afname van de digitale of papieren versie van de checklist. Voor het opvragen en interpreteren van de groeps- en leerlingrapportage is er een afzonderlijke digitale handleiding (te downloaden van Cito Portal). Het kan voorkomen dat de leerkracht een of meer vragen wegens gebrek aan informatie niet onmiddellijk kan beantwoorden. In dat geval wordt hem of haar geadviseerd enkele dagen de tijd te nemen voor het observeren van het desbetreffende gedrag om de vragen daarna alsnog te beantwoorden. Ook bevat de afnamehandleiding aanwijzingen voor duo-leerkrachten die de vragen in onderling overleg willen beantwoorden. Tot slot zijn er adviezen voor de keuze van het juiste afnamemoment en het volgen van de ontwikkeling van leerlingen in de tijd.

Algemene aanwijzingen voor de afname van de praktische opdrachten

Ter ondersteuning van de afname van de praktische opdrachten zijn er twee typen documenten: de afnamekaarten voor de voorbereiding en de afname van de opdrachten en de gebruikershandleiding (Cito, 2018) met algemene aanwijzingen voor de afname, beoordeling, rapportage en instructie tijdens de afname.

Afname op ieder willekeurig moment

De afname van een praktische opdracht is niet aan een bepaald moment gebonden. De opdrachten en afnamekaarten zijn zo ontwikkeld dat geen bijzondere vakinhoudelijke kennis en vaardigheden vereist zijn. Zoals uiteengezet in hoofdstuk 3 kan iedereen die de leerling goed kent de opdrachten afnemen. In de meeste gevallen zal dat de 'eigen' leerkracht of de intern begeleider zijn.

Variabele afnametijd

In een DA is de duur van een afname niet altijd van te voren te voorspellen. Hoe minder problemen de leerling tijdens de afname tegenkomt, hoe korter de afname zal duren. De duur van de afname hangt ook af van hoe de leerkracht het pakket gebruikt. Wil de leerkracht alleen het niveau van leren leren bepalen, dan zal de afname weinig tijd in beslag nemen (maar zal de leerling weinig van de afname leren).

Combineert de leerkracht de afname en beoordeling met het onderwijzen van leren leren, dan vergt de afname natuurlijk meer tijd.

Gefaseerde opbouw van de afname

Een sterke voorstructurering van de afname kan een zelfstandige toepassing van vaardigheden van zelfregulerend leren in de weg staan. Vanwege de kortere spanningsboog van sbo- en so-cluster-4-leerlingen is er toch voor gekozen om de afname in verschillende, onderling samenhangende onderdelen op te knippen. In overeenstemming met het onderscheid in de vier hoofdvaardigheden die in het pakket centraal staan, kent elke afname vier fasen:

- oriëntatie op de opdracht;
- voorbereiding en planning van de opdracht;
- uitvoering van de opdracht;
- reflectie en evaluatie van de opdracht.

Het onderscheid in deze vier fasen is van cruciaal belang voor een soepele organisatie van de afname. Zo bevat de algemene gebruikershandleiding voor elke fase afzonderlijke instructies waarin de werkzaamheden van leerkracht en leerling beknopt omschreven zijn. Bij elke fase hoort een beoordelingsformulier waarmee de leerkracht de leerling beoordeelt.

De toewijzing van varianten aan leerlingen

Tussen leerlingen met een beperking bestaan grote verschillen qua cognitieve en praktische vaardigheden, zelfstandigheid en ondersteuningsbehoefte. Daarom bevat het pakket voor elke opdracht een makkelijke, een moeilijkere en een moeilijke variant. Het is de bedoeling dat de leerkracht de variant kiest die naar zijn of haar inschatting het beste past bij wat de leerling aankan. De praktische opdracht moet uitdagend zijn en de leerling met problemen confronteren, maar mag niet te makkelijk of te moeilijk zijn. De moeilijkheid van de opdracht moet vallen binnen de zone van naaste ontwikkeling van de leerling. De opdracht mag niet zo makkelijk zijn dat de leerling deze geheel zelfstandig kan maken. Tegelijkertijd mag de opdracht niet zo moeilijk zijn dat de leerling deze ondanks de hulp van de leerkracht niet tot een goed einde kan brengen. In de gebruikershandleiding (Cito, 2018) zijn de leerkrachten er met nadruk op gewezen dat de scores alleen betekenis hebben als de variant op de juiste wijze aan de leerling wordt toegewezen. Leerkrachten wordt dus geadviseerd de opdracht af te stemmen op het niveau van de leerling. Ter ondersteuning bevat de handleiding het volgende toewijzingschema:

Variante	Groep 5	Groep 6	Groep 7	Groep 8
A - makkelijk	X	X		
B - moeilijker		X	X	
C - moeilijk			X	X

In de gebruikershandleiding zijn de leerkrachten er uitdrukkelijk op gewezen dat het schema slechts een vrijblijvend advies bevat en geen bindende richtlijn is. De leerkrachten zijn verzocht om vooral op basis van hun eigen inschatting te beoordelen welke variant het meest geschikt is voor hun leerling. Zij zijn er tevens op gewezen dat het daarvoor belangrijk is om eerst alle drie de varianten goed te bestuderen.

Door de adaptieve toewijzing van opdrachten en varianten aan leerlingen, is het niet mogelijk om de scores van leerlingen die verschillende varianten gemaakt hebben met elkaar te vergelijken. De leerkracht geeft namelijk geen absoluut oordeel over de vaardigheid van de leerling, maar een relatief oordeel gegeven de moeilijkheid van de gekozen variant. Daardoor is het niet zinvol om de vaardigheid van leerlingen die verschillende varianten gemaakt hebben met elkaar te vergelijken. Hetzelfde geldt voor het vergelijken van de vaardigheid van bijvoorbeeld leerlingen in sbo en so cluster 4. Wel is het pakket een bruikbaar middel voor het vaststellen, oefenen en verder ontwikkelen van de algemene vaardigheden van individuele leerlingen van eenzelfde groep. Ook kan de leerkracht onder bepaalde voorwaarden de resultaten

vergelijken met die van een collega die aan dezelfde leerlingen lesgeeft. De gebruikershandleiding (Cito, 2018) bevat aanwijzingen die ervoor moeten zorgen dat verschillende leerkrachten de opdrachten zoveel mogelijk op dezelfde manier afnemen. Het pakket bevat echter geen uitgewerkte protocollen of scripts waarin precies is vastgelegd welke hulp de leerkracht in welke situatie moet geven. Daardoor kan de ene leerkracht in eenzelfde situatie meer of andere hulp bieden dan de andere leerkracht. Ook kan de ene leerkracht net iets strenger of soepeler beoordelen dan de andere leerkracht. Dit beperkt de mogelijkheden om de scores van verschillende leerkrachten met elkaar te vergelijken. Wel kan de leerkracht de resultaten relatief objectief vergelijken als de collega's nagaan in hoeverre zij de afname, beoordeling en instructie inderdaad op een vergelijkbare manier uitvoeren.

Het volgen van de ontwikkeling van leren leren

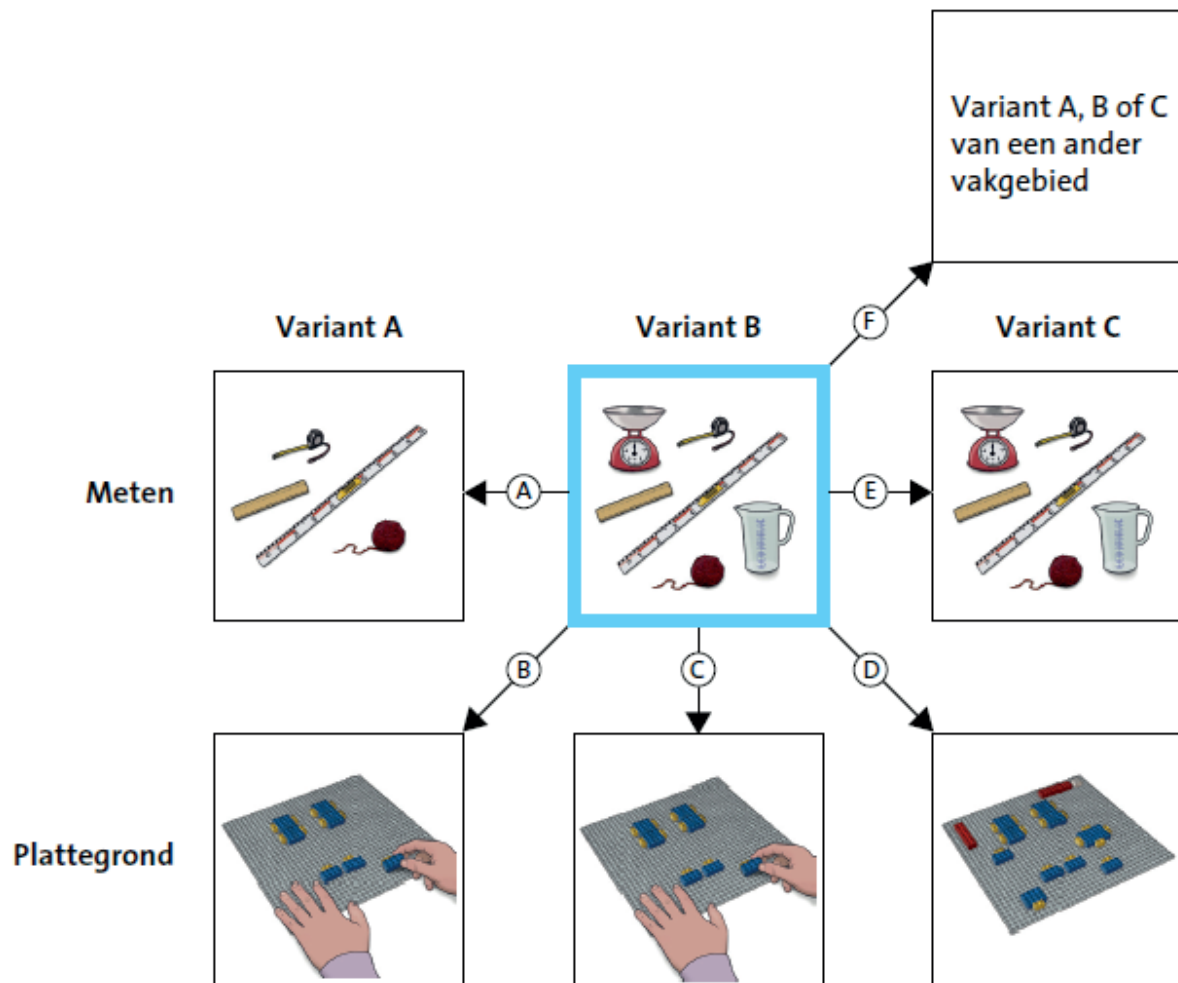
Met het pakket kan de leerkracht vaststellen hoe vaardig een leerling is op het gebied van leren leren en inschatten hoeveel hulp hij of zij in het vervolgtraject nodig zal hebben. Door na enige tijd een tweede opdracht na de eerste opdracht af te nemen, kan worden vastgesteld of en zo ja in hoeverre de vaardigheid in leren leren vooruit is gegaan (en hoeveel de leerling van de eerste afname en het onderwijs in leren leren geleerd heeft). Welke variant de leerkracht het beste als tweede kan afnemen, is afhankelijk van de scores van de leerling op de eerste opdracht. Was de eerste opdracht bijvoorbeeld zo moeilijk dat het de toepassing van leren leren in de weg stond, dan ligt de keuze van een makkelijkere variant voor de hand. Gaf de eerste afname een positief resultaat te zien, dan zijn er afhankelijk van de mate van succes ten minste drie mogelijkheden (Tzuril, 2001):

- afname van dezelfde variant of de parallelle variant van hetzelfde vakgebied (om *maintenance* vast te stellen);
- afname van een moeilijkere variant van hetzelfde vakgebied (om nabije transfer vast te stellen);
- afname van een bij benadering even moeilijke of moeilijkere variant van een ander vakgebied (om verre transfer vast te stellen).

De algemene gebruikershandleiding (Cito, 2018) geeft een eenvoudige vuistregel voor de keuze van de vervolgoopdracht die is weergegeven in het volgende overzicht.

Vaardigheid	Scores	Advies voor de keuze van de vervolgoopdracht
Lage vaardigheid en regelmatig tot voortdurend hulp nodig	Overwegend scores 1 en 2	Een makkelijkere variant van dezelfde of een andere opdracht (eventueel van een ander vakgebied)
Middelmatige vaardigheid en af en toe tot regelmatig hulp nodig	Overwegend scores 2 en 3	Een even moeilijke variant van een andere opdracht (eventueel van een ander vakgebied)
Hoge vaardigheid en geen of slechts af en toe hulp nodig	Overwegend scores 3 en 4	Een moeilijkere variant van dezelfde of een andere opdracht (eventueel van een ander vakgebied)

Figuur 3.1 geeft weer hoe de leerkracht deze vuistregel in de praktijk kan toepassen. Voor een nadere toelichting op de figuur verwijzen we naar paragraaf 5.2 van de gebruikershandleiding (Cito, 2018).



Figuur 3.1 De keuze van een vervolgoopdracht om de voortgang van leren vast te stellen.

De tijd tussen twee afnames

Door de resultaten van twee afnames met elkaar te vergelijken, kan de leerkracht nagaan in hoeverre zijn of haar onderwijs in leren leren effectief was en welke ondersteuning de leerling eventueel nog meer nodig heeft. De tijd tussen de eerste en tweede afname is afhankelijk van het doel dat de leerkracht heeft. Als de leerkracht wil vaststellen wat de leerling geleerd heeft van de eerste afname en hoeveel ondersteuning nog nodig is, dan kunnen de twee afnames het beste kort na elkaar gepland worden (direct of hooguit enkele dagen na elkaar). Hoewel wij hier nog geen onderzoek naar gedaan hebben, verwachten wij dat de vooruitgang doorgaans bescheiden zal zijn. Eén afname van hooguit en half uur zal voor veel leerlingen eenvoudigweg te weinig opleveren. Aanzienlijk meer vooruitgang kan de leerkracht verwachten als hij of zij de vaardigheden van leren leren gedurende een langere periode systematisch onderwijst, bijvoorbeeld volgens de Beertjes-methodiek (Meichenbaum, 1981; Timmerman, 2002; Van Bon, 2009). Als de leerkracht daarna de voortgang wil vaststellen, kan hij of zij de tweede opdracht het beste na een periode van drie à zes maanden afnemen.

3.3.3 Observatie en beoordeling

In een *dynamic assessment* is gedragsobservatie een belangrijk instrument ter verkrijging van informatie over het leerproces van de individuele leerling (Lidz, 1987, 1991; Zimmerman, 2008). De observatie wordt doorgaans uitgevoerd met observatielijsten. Deze meten meestal geen vakspecifieke vaardigheden, maar wel allerlei ondersteunende vaardigheden die van belang zijn voor het vaststellen van het leerpotentieel (zoals aandacht, plannen, zelfregulatie, impulscontrole), de gevoeligheid voor de instructie die de leerkracht heeft moeten geven tot de leerling de taak zelfstandig kon uitvoeren en de transfer naar andere dan de geoefende vaardigheden, taken en situaties (Lidz, 1987, 1991; Tzuriel, 2001).

Drie korte beoordelingsformulieren

Het pakket Leren leren bevat vier korte observatielijsten die in de gebruikershandleiding beoordelingsformulieren worden genoemd. Bij praktische opdrachten reageert elke leerling op zijn of haar eigen manier. Bij het beoordelen van algemene vaardigheden is er geen sprake van helemaal goed of helemaal fout, maar van een glijdende schaal met allerlei gradaties tussen zwakke en goede prestaties. De beoordeling is dan lastig. Verschillende leerkrachten kennen aan een zelfde prestatie of product verschillende scores toe. Het werken met descriptieve beoordelingsformulieren – vaak rubrics genoemd – maakt de beoordeling objectiever en consistent (Arter & McTighe, 2001; Jonsson & Svingby, 2007). Bovendien bieden descriptieve beoordelingsformulieren een inhoudelijk kader om de voortgang van het leren vast te stellen. De leerkracht beoordeelt de leerling door zijn of haar algemene vaardigheid te vergelijken met de niveauomschrijvingen op het formulier. Hiermee plaatst de leerkracht de leerling als het ware op de schaal. Het is de kunst om precies die niveauomschrijving te kiezen die het beste overeenkomt met het vaardigheidsniveau van de leerling.

Het pakket bevat vier formulieren voor het beoordelen van:

- vaardigheden tijdens de oriëntatie en voorbereiding op de opdracht (fase 1 en 2);
- vaardigheden tijdens de uitvoering van de opdracht (fase 3);
- vaardigheden van reflectie en evaluatie nadat de opdracht is uitgevoerd (fase 4);
- de hoeveelheid benodigde hulp tijdens de fasen van Oriëntatie en voorbereiding, Uitvoeren en monitoren, Reflecteren en evalueren (als omgekeerde maat voor leerpotentieel);
- werkhouding: aandachtig werken, gemotiveerd werken en werken zonder gedragsproblematiek.

Formulier 1 vult de leerkracht deels in aan het einde van afnamefase 1 en deels aan het einde van afnamefase 2, formulier 2 aan het einde van fase 3, formulier 3 en 4 aan het einde van fase 4 als de leerling de opdracht helemaal voltooid heeft. De vier formulieren zijn integraal opgenomen in bijlage 3.

Elk formulier bevat een aantal aspecten waar de leerkracht bij de beoordeling op moet letten.

De beoordelingsaspecten zijn vaak geformuleerd als vragen die de leerkracht tijdens de leergesprekken aan de leerling stelt. De beoordelingsformulieren hebben dus twee functies:

- het beoordelen van de vaardigheid van leren leren;
- het geven van richting aan de leergesprekken. De leerkracht voert de leergesprekken namelijk aan de hand van de vragen die zijn vermeld in de eerste kolom van het beoordelingsformulier.

De beoordelingsaspecten

De uiteindelijke versie van het beoordelingsinstrument bevat in totaal zestien beoordelingsaspecten. Zoals weergegeven in tabel 3.1 zijn deze aspecten onderverdeeld in zelfregulerende vaardigheden van leren leren, benodigde hoeveelheid hulp (als omgekeerde maat voor het leerpotentieel) en werkhouding (aandachtig en gemotiveerd werken zonder gedragsproblematiek).

Tabel 3.1 *Verdeling van de beoordelingsaspecten naar vaardigheid en type beoordelingsaspect in het uiteindelijke beoordelingsinstrument*

Vaardigheid en beoordelingsaspect	Type aspect
Oriënteren en voorbereiden	
Wat ga je doen?	Zelfregulering
Wat ga je eerst doen en wat daarna?	Zelfregulering
Heb je alle spullen klaargezet? en Wat moet je nog pakken?	Zelfregulering
Benodigde hulp tijdens oriënteren en voorbereiden	Leerpotentieel
Uitvoeren	
Werken volgens (eigen) plan	Zelfregulering
Zelfmonitoring	Zelfregulering
Controle eigen werk	Zelfregulering
Benodigde hulp tijdens de uitvoering	Leerpotentieel
Reflecteren en evalueren	
Lijkt wat je gemaakt hebt op het voorbeeld? Of: Past wat je hebt gedaan/gemaakt bij de opdracht?	Zelfregulering
Wat ging er goed? (terugkijkvraag)	Zelfregulering
Wat ging er niet zo goed of fout? (terugkijkvraag)	Zelfregulering
Wat kan er beter? (vooruitkijkvraag)	Zelfregulering
Benodigde hulp tijdens de reflectie en evaluatie	Leerpotentieel
Werkhouding	
Aandachtig werken	Aandacht
Gemotiveerd werken	Motivatie
Werken zonder gedragsproblematiek	Impulsbeheersing

Generieke beoordelingsschalen

Voor de beoordeling van leren leren is gekozen voor generieke in plaats van opdracht-specifieke beoordelingsschalen (voor een uiteenzetting over de wederzijdse voor- en nadelen zie Kuhlemeier, 2011). In een opdracht-specifieke schaal zijn de ontwikkelings- of prestatieniveaus omschreven in termen van opdracht-specifieke kenmerken of producteigenschappen. Deze schalen zijn geschikt voor eenmalig gebruik en vooral van belang als het gaat om duidelijk afgebakende kenniselementen of om procedures die in een vaste volgorde moeten worden doorlopen. Het gebruik ervan is zinvol als een hoge betrouwbaarheid van de praktijktoets gewenst is, omdat er bijvoorbeeld belangrijke beslissingen in het geding zijn (wat bij leren leren niet het geval is).

In een generieke beoordelingsschaal zijn de beoordelingsaspecten en -criteria daarentegen voor alle opdrachten gelijk. Een generieke beoordelingsschaal past goed bij de algemene vaardigheden waarvan leren leren een voorbeeld is. Dit type schalen is bedoeld voor herhaalde toepassing gedurende een langere periode. Leerkracht en leerlingen passen dezelfde beoordelingscriteria toe op verschillende opdrachten, problemen en contexten. Het voordeel is dat de kwaliteitseisen daardoor op den duur geïnternaliseerd worden. De leerling leert zo zijn of haar eigen werk te beoordelen. De generieke beoordelingsschalen in het pakket Leren leren kunnen leerlingen helpen bij de transfer van kennis en vaardigheden naar nieuwe situaties met nieuwe problemen. In ons geval vergemakkelijken generieke schalen de transfer van vaardigheden die met de ene opdracht geoefend zijn naar andere, nog niet geoefende opdrachten.

Vier schaalpunten

Na de keuze van een geschikt schaaltype moet men een beslissing nemen over het aantal schaalpunten. Het wenselijke aantal schaalpunten hangt onder meer af van vier factoren (Kuhlemeier, 2011):

- De complexiteit van het beoordelingsobject: hoe complexer, hoe meer schaalpunten.
- De grootte van de prestatieverschillen tussen de leerlingen: hoe groter de verschillen tussen leerlingen, hoe meer schaalpunten voor een gedifferentieerde beschrijving van deze verschillen nodig zijn.

- Het doel van de toetsing: toetsen van leren versus toetsen voor leren. Indien de praktische opdracht in het kader van een *dynamic assessment* gebruikt wordt, moet de beoordelingsschaal de leerkracht handzame informatie geven over waar de leerling goed in is en waar verbetering nodig is. Het aantal schaalpunten mag dan niet te groot zijn.
- De tijd die de leerkracht in de afnamesituatie aan de beoordeling kan besteden: hoe minder tijd, hoe minder schaalpunten.

De complexiteit van de vaardigheid leren leren en de grootte van de verschillen tussen sbo- en so cluster 4-leerlingen rechtvaardigen een groter aantal schaalpunten, terwijl het doel van de toetsing en de complexiteit van de afname- en beoordelingssituatie aanleiding geven tot een kleiner aantal schaalpunten. De ervaring leert dat meer dan zeven schaalpunten in de praktijk niet goed werkt. Het lukt de beoordelaars dan niet meer om de prestaties van de leerlingen betrouwbaar te onderscheiden. Drie of minder schaalpunten is vaak weer te weinig om de vooruitgang van de leerlingen goed te kunnen rapporteren en het testresultaat te gebruiken in dienst van het leren. Na deze deels tegenstrijdige argumenten tegen elkaar afgewogen te hebben, hebben wij uiteindelijk gekozen voor schalen met vier antwoordmogelijkheden.

Scoring en rapportage in vier vaardigheidsniveaus

De leerkracht kan in de beoordelingsformulieren steeds kiezen uit vier antwoordmogelijkheden (zie bijlage 3). De globale betekenis van de scores voor de vaardigheden van leren leren is als volgt:

- Score 1: Zwakke vaardigheid in leren leren;
- Score 2: Redelijke vaardigheid in leren leren;
- Score 3: Goede vaardigheid in leren leren;
- Score 4: Uitstekende vaardigheid in leren leren.

Bij de beoordelingsaspecten voor de hoeveelheid benodigde hulp hebben de scores de volgende betekenis:

- Score 1: Voortdurend hulp nodig;
- Score 2: Regelmatig hulp nodig;
- Score 3: Af en toe hulp nodig;
- Score 4: Geen hulp nodig.

3.3.4 Groeps- en leerlingrapportages

De gebruikershandleiding (Cito, 2018) bevat aanwijzingen voor het toekennen van scores en het interpreteren van de uitslag van individuele leerlingen en van groepen leerlingen. Het gebruik van de leerling- en groepsrapportages wordt geïllustreerd aan de hand van concrete voorbeelden. Voor het invullen van de checklist en de beoordelingsformulieren en het opvragen van rapportages bevat de uitgave twee digitale rapportagemodules die de leerkracht kan downloaden van Cito Portal (<http://portal.cito.nl>).

3.3.4.1 Rapportage van de resultaten van de afname van de checklist

De groepsrapportage van de checklist

De groepsrapportage bij de checklist heeft tot doel snel een indruk te geven van het niveau van leren leren in de groep als geheel. De rapportage laat zien welke leerlingen meer en minder moeite hebben met leren en leren en welke vaardigheden in de groep als geheel meer en minder goed ontwikkeld zijn. De leerkracht kan hiermee bepalen welke vaardigheden of leerlingen extra aandacht vereisen en waar dat niet of minder nodig is. Een voorbeeld van een groepsrapportage is weergegeven in figuur 3.2.

School:	De klarinet	De klarinet	De klarinet	De klarinet	De klarinet	De klarinet	De klarinet
Groep:	6	6	6	6	6	6	6
Leerkracht:	José	José	José	José	José	José	José
Leering:	Thomas	Tamara	Demi	Julian	Youri	Esra	Dave
Afnamedatum:	11-07-18	11-07-18	11-07-18	11-07-18	11-07-18	11-07-18	11-07-18
Percentielscore	52	60	57	15	56	35	13
VAARDIGHEDEN LEREN LEREN							
Oriënteren, voorbereiden en plannen van praktische taken	● 72	● 8	● 94	● 72	● 66	● 66	● 77
Uitvoeren van praktische taken	● 45	● 65	● 31	● 5	● 45	● 5	● 16
Monitoren, reflecteren en evalueren van praktische taken	● 60	● 63	● 13	● 13	● 22	● 13	● 4
Zelfstandig werken en hulp vragen	● 53	● 53	● 19	● 15	● 33	● 9	● 15
WERKHOUDING							
Aandachtig werken	● 2	● 94	● 94	● 16	● 91	● 62	● 16
Gemotiveerd werken	● 94	● 86	● 79	● 16	● 91	● 56	● 23
Werken zonder gedragsproblematiek	● 89	● 73	● 98	● 39	● 85	● 89	● 9
Kopiëren als afbeelding		leerling-rapportage	leerling-rapportage	leerling-rapportage	leerling-rapportage	leerling-rapportage	leerling-rapportage

Figuur 3.2 Voorbeeld van een groepsrapportage bij de checklist

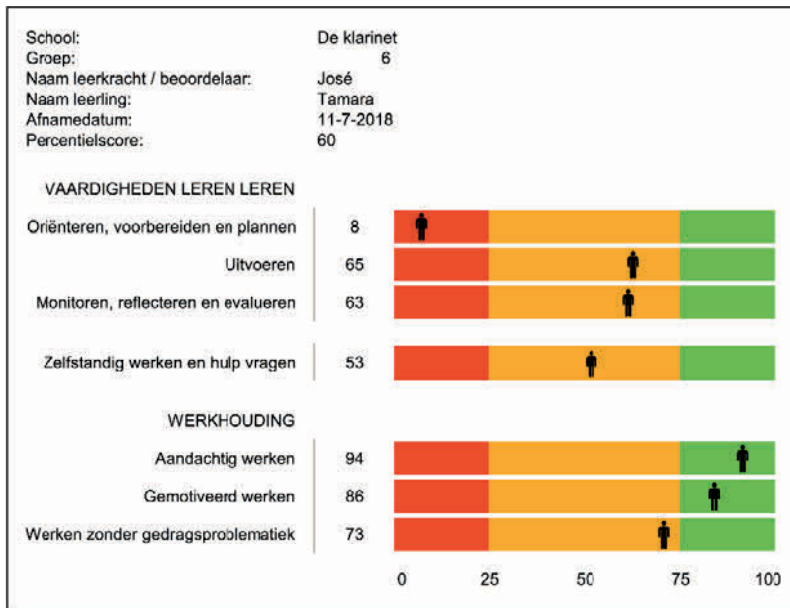
De vaardigheid van de leerling in vergelijking met de landelijke referentiegroep is weergegeven met een gekleurd bolletje en een getal, de zogenaamde percentielscore. De percentielscore geeft aan welk percentage van de leerlingen in de referentiegroep de leerling achter zich laat (dit wil zeggen: welk deel van de referentiegroep eenzelfde of lagere somscore op de checklist heeft behaald dan de leerling). Dus, hoe hoger de percentielscore, hoe gunstiger de vaardigheden in leren leren. De figuur toont eerst de percentielscore voor de volledige checklist en vervolgens de percentielscores voor de zeven schalen (ofwel vaardigheden van leren leren).

De betekenis van de kleuren is als volgt:

- groen: de leerling behoort tot de bovenste 25% leerlingen met de meest gunstige scores (percentiel 75-100);
- oranje: de leerling behoort tot de middelste 50% van de leerlingen (percentiel 25-75);
- rood: de leerling behoort tot de onderste 25% leerlingen met de minst gunstige scores (percentiel 0 - 25).

De leerlingrapportage van de checklist

Figuur 3.3 toont een voorbeeld van een leerlingrapportage. De percentielscores zijn voor elk van de zeven schalen weergegeven in de vorm van het poppetjesdiagram dat bekend is van de Centrale Eindtoets Basisonderwijs. De kleuren hebben dezelfde betekenis als in het groepsoverzicht: rood voor het onderste kwart, oranje voor de middelste helft en groep voor het bovenste kwart.



Figuur 3.3 Voorbeeld van een leerlingrapportage bij de checklist

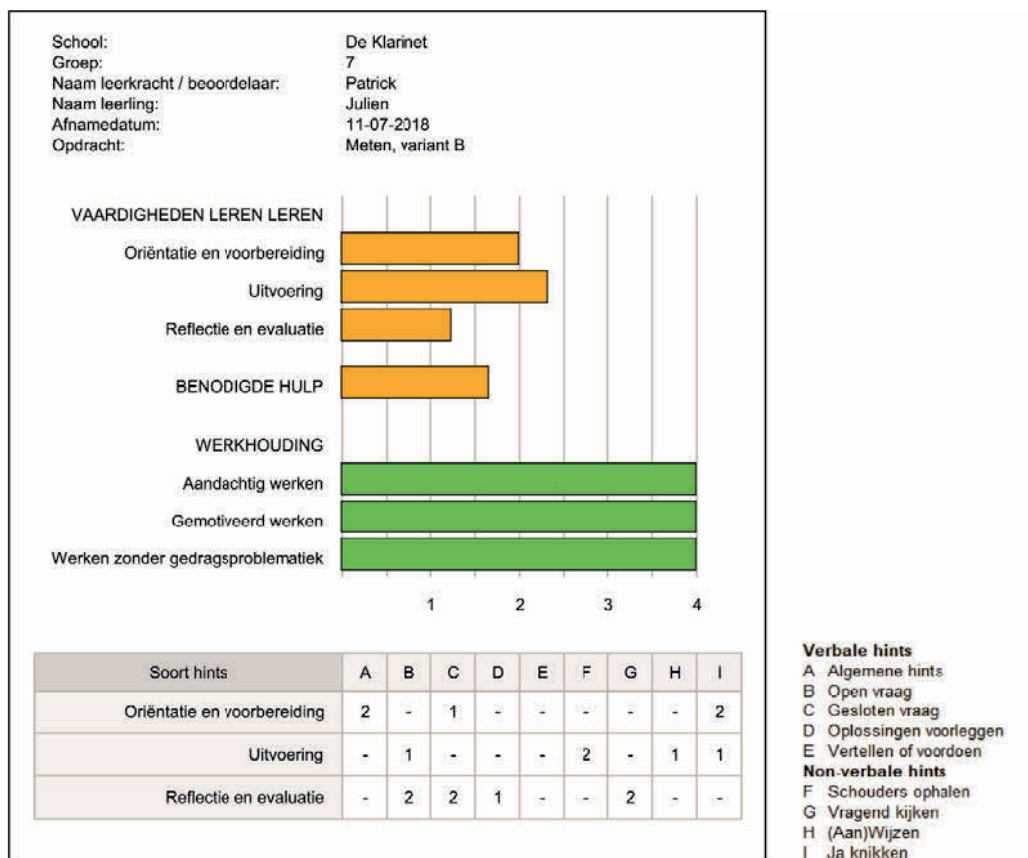
3.3.4.2 Rapportage van de resultaten van de afname van de praktische opdrachten

Leerkrachten kunnen de digitale rapportagemodule bij de praktische opdrachten op drie manieren gebruiken:

- De leerkracht heeft voor een of meer leerlingen het papieren beoordelingsformulier ingevuld en wil een digitale leerling- en groepsrapportage opvragen. Dat kan door de scores over te brengen naar de applicatie.
- De leerkracht wil de scores tijdens de afname digitaal invoeren zonder gebruik te maken van het papieren beoordelingsformulier (en eventueel op een later moment een leerling- en groepsrapportage opvragen).
- De leerkracht wil de digitale leerling- en groepsrapportage bewaren om deze op een later moment te gebruiken voor het volgen van de ontwikkeling van de leerlingen.

De leerlingrapportage bij de praktische opdrachten

Figuur 3.4 toont een voorbeeld van een leerlingrapportage zoals aangemaakt met de digitale rapportagemodule. De figuur bevat de gemiddelde scores voor de drie hoofdvaardigheden (Oriënteren en voorbereiden, Uitvoering en Reflecteren en evalueren) en de hoeveelheid benodigde hulp. Daarnaast zijn de door de leerkracht toegekende scores voor werkhouding een-op-een weergegeven.



Figuur 3.4 Voorbeeld van een leerlingrapportage bij de praktische opdrachten

De uitslag wordt niet als een getal getoond, maar als een gekleurde staaf, ook wel een signaal genoemd. Hoe langer de staaf, hoe gunstiger de uitslag. Voor de drie hoofdvaardigheden oriënteren/voorbereiden, uitvoeren, reflecteren/evalueren en de hoeveelheid benodigde hulp zijn er drie signalen:

- groen voor een hoge vaardigheid of geen of hooguit af te toe hulp nodig (gemiddelde scores in het bereik van 3 tot en met 4);
- oranje voor een middelmatige vaardigheid of hooguit af en toe tot regelmatig hulp nodig (gemiddelde scores in het bereik van 2 tot 3);
- rood voor een lage vaardigheid of regelmatig tot voortdurende hulp nodig (gemiddelde scores in het bereik van 1 tot 2).

Voor de onderscheiden aspecten van werkhouding is de betekenis van de signalen als volgt:

- groen voor score 4;
- oranje voor score 3;
- rood voor score 1 of 2.

De ruimte onder het staafdiagram bevat:

- een getalsmatig overzicht van het aantal hints per type hint dat de leerkracht heeft moeten geven (als omgekeerde maat voor het leerpotentieel);
- ruimte voor het maken van opmerkingen bij de uitvoering van de vaardigheden van leren leren en de werkhouding (in figuur 3.4 is dit niet ingevuld)

De groepsrapportage bij de praktische opdrachten

Figuur 3.5 toont een groepsrapportage met de signalen voor alle leerlingen in de door de leerkracht gedefinieerde groep.

School	De Klarinet	De Klarinet	De Klarinet	De Klarinet	De Klarinet
Groep:	7	7	7	7	7
Naam leerkracht / beoordelaar:	Patrick	Patrick	Patrick	Patrick	Patrick
Naam leerling:	Julien	Kim	Jan	Corien	Ilyas
Afnamedatum:	11-07-2018	11-07-2018	11-07-2018	11-07-2018	11-07-2018
Opdracht:	MET	REIS	POW	TIJD	TEK_P
Variant:	B	B	C	B	A

VAARDIGHEDEN LEREN LEREN					
Oriëntatie en voorbereiding	●	●	●	●	●
Uitvoering	●	●	●	●	●
Reflectie en evaluatie	●	●	●	●	●

BENODIGDE HULP					
	●	●	●	●	●

WERKHOUDING					
Aandachtig werken	●	●	●	●	●
Gemotiveerd werken	●	●	●	●	●
Werken zonder gedragsproblematiek	●	●	●	●	●

Kopiëren als afbeelding

leerling-rapportage

leerling-rapportage

leerling-rapportage

leerling-rapportage

leerling-rapportage

Figuur 3.5 Voorbeeld van een groepsrapportage bij de praktische opdrachten

3.3.5 Handreikingen voor het onderwijzen van leren leren

Omdat vaardigheden van leren leren als moeilijk onderwijsbaar gelden, is in de gebruikershandleiding (Cito, 2018) veel aandacht besteed aan suggesties voor het onderwijzen van vaardigheden van leren leren. Voor het onderwijzen van leren leren is een scala aan effectieve didactische werkwijzen beschikbaar (Boekaerts & Corno, 2005; Dignath & Büttner, 2008; Dignath, Büttner & Langfeldt, 2008; Zumbunn, Tadlock & Roberts, 2011). Daartoe behoren directe instructie en modellering (voordoen), begeleid ontdekkend leren en sociale ondersteuning en feedback. De handreikingen in de gebruikershandleiding (Cito, 2018) zijn voornamelijk ontleend aan de BOL/EIM-methodiek, de Beertjes-methodiek en literatuur over executieve functies en metacognitie. De BOL/EIM-methodiek is een bewezen effectieve methode voor het onderwijzen van vaardigheden van leren leren (Aukes e.a., 2000; Den Besten & Vulpen, 2013; Timmer, 1994; Timmer, Dekker & Voortman, 2003). BOL staat voor begeleid ontdekkend leren en EIM staat voor eigen initiatief model. Over de effectiviteit van de Beertjes-methodiek (Meichenbaum, 1981; Timmerman, 2002, 2017; Van Bon, 2009) zijn voor zover ons bekend nog geen 'harde' onderzoeksgegevens beschikbaar. Wel wordt de Beertjes-methodiek naar verluid in het sbo en so cluster 4 veel gebruikt. Beide methodieken zijn gericht op het vergroten van de zelfstandigheid van leerlingen met een of meer beperkingen. Leerkrachten die ermee werken, leren leerlingen vaardigheden van leren leren die hen in staat stellen zelfstandiger te functioneren en minder afhankelijk te worden van anderen. Fouten zijn er om van te leren, zolang de leerling ze maar zelf ontdekt en dan nadenkt over een manier om ze te herstellen. Leerlingen leren gaandeweg zelfstandiger te werken, meer verantwoordelijkheid te nemen, initiatieven te nemen en flexibeler om te gaan met nieuwe situaties.

De praktische opdrachten zijn zo gemaakt dat de leerling ze meestal niet volledig zelfstandig kan uitvoeren. Tijdens de afname observeert de leerkracht welke werkzaamheden de leerling al wel zelfstandig kan uitvoeren en waar hij of zij tegen problemen aanloopt of zelfs vastloopt.

De BOL/EIM-methodiek (Den Besten & Vulpen, 2013) en de Beertjes-methodiek (Meichenbaum, 1981; Timmerman, 2002; Van Bon, 2009) geven praktische aanwijzingen over hoe de leerkracht de leerling in zo'n situatie kan helpen. Zo wordt in de gebruikershandleiding ruimschoots aandacht geschonken aan het principe van de geleidelijke toename van algemene naar specifieke hulp (Den Besten & Vulpen, 2013), het geven van hints, het stellen van verschillende soorten vragen, het reageren op hulpvragen, het voeren van leergesprekken tijdens de afname en het gebruik van de resultaten voor de planning van het onderwijs in leren leren. Voor meer informatie verwijzen wij naar de hoofdstukken 6 en 7 van de gebruikershandleiding (Cito, 2018).

4 Uitvoerbaarheid, betrouwbaarheid en validiteit van de checklist

In het vorige hoofdstuk deden we onder meer verslag van de constructie, opbouw, afname, beoordeling en rapportage van het pakket leren leren. In dit hoofdstuk bespreken we eerst het doel, de opzet en de uitvoering van de proefafname waarin een conceptversie van de checklist is uitgetest op uitvoerbaarheid, betrouwbaarheid en validiteit (paragraaf 4.1). Vervolgens rapporteren we de uitkomsten van een gebruikersonderzoek waarin de proefafnameleerkrachten onder meer is gevraagd naar hun ervaringen met de checklist en waarbij zij verbeteringsuggesties konden aangeven (paragraaf 4.2). Vervolgens bespreken we de wijze waarop we de definitieve versie van de checklist op basis van de proefafname hebben samengesteld (paragraaf 4.3, 4.4 en 4.5). Tot slot doen we verslag van de betrouwbaarheid van de meting van vaardigheden van leren leren en gaan we nader in op enkele aspecten van de validiteit (paragraaf 4.6 en 4.7). Vanwege budgettaire beperkingen zijn slechts enkele aspecten van betrouwbaarheid en validiteit onderzocht. Zo is geen onderzoek gedaan naar de interbeoordelaarsbetrouwbaarheid en de test-hertestbetrouwbaarheid. Evenmin is er onderzoek gedaan naar eventuele misclassificaties (waarbij de leerling onterecht naar de praktische opdrachten wordt verwezen terwijl hij het niveau in feite wel beheerst of waarbij de leerling ten onterechte niet naar de praktische opdrachten verwezen wordt terwijl hij in feite niet over voldoende vaardigheid beschikt).

Omdat het pakket Leren leren vooralsnog enig in zijn soort is, is onderzoek naar de convergente en discriminante validiteit nog niet mogelijk.

4.1 Doel, opzet en uitvoering van de eerste proefafname

Doelen van de proefafname

In het vroege voorjaar van 2017 heeft een onderzoek plaatsgevonden naar de uitvoerbaarheid, betrouwbaarheid en validiteit van de conceptversie van de checklist. De proefafname kende drie doelen:

- het inventariseren van gebruikerservaringen en verbeteringsuggesties bij de leerkrachten die de checklist invulden;
- het uittesten van de concept-checklist met als doel het aantal uitspraken en hiermee de afnametijd zoveel mogelijk te beperken (zonder afbreuk te doen aan de betrouwbaarheid en validiteit);
- het onderzoeken van de betrouwbaarheid en de validiteit van de meting van vaardigheden van leren leren.

Inventariseren gebruikerservaringen

De gebruikerservaringen zijn vastgesteld door de deelnemende leerkrachten een vragenlijst naar hun ervaringen en eventuele verbeteringsuggesties voor te leggen.

Uittesten checklist

In de proefafname is een eerste versie van de checklist uitgetest. De checklist is een hulpmiddel om te bepalen welke vaardigheden extra klassikale aandacht vereisen en bij welke vaardigheden dat niet of minder nodig is. De checklist stelt leerkrachten in staat om op efficiënte wijze de leerlingen te selecteren die het meest voor afname van een arbeidsintensieve praktische opdracht in aanmerking komen.

De conceptversie van de checklist bevatte meer uitspraken over de vaardigheid in leren leren dan voor een betrouwbare meting naar verwachting nodig zouden zijn. De reden is dat wij meer wilden weten over de interne structuur van de vaardigheid in leren leren en wilden bepalen hoeveel uitspraken er nodig zijn om deze vaardigheid betrouwbaar te kunnen meten. Het doel was informatie te verzamelen op basis waarvan het aantal uitspraken met zo weinig mogelijk verlies aan informatie zo ver mogelijk kon worden teruggebracht.

Vaststellen betrouwbaarheid en validiteit

De betrouwbaarheid van de eerste versie van de checklist is vastgesteld aan de hand van de interne consistentie, de item-restcorrelaties en de standaardmeetfout. Een van de validiteitsvragen is in hoeverre het a priori onderscheid in vaardigheden van leren leren (zie paragraaf 2.2) is terug te vinden in de data.

Opzet en uitvoering

Aan de proefafname van de eerste versie van de checklist is deelgenomen door 68 docenten van 17 scholen (53 groepen). In totaal is voor 629 leerlingen een checklist ingevuld. Daarvan zaten 226 leerlingen op een sbo-school en 403 leerlingen op een so cluster 4-school.

4.2 Ervaringen van gebruikers

Hoe functioneerde de checklist?

- de gemiddelde invultijd van de checklist was 12,5 minuten;
- 61% vond dat de checklist goed in te vullen was en 39% vond dat dit ten dele het geval was;
- 73% vond de aandachtspunten duidelijk, 24% deels duidelijk en 3% vond ze niet duidelijk;
- 82% miste geen aandachtspunten;
- 75% gaf aan dat er geen overbodige vragen in de checklist stonden;
- 72% vond de checklist bruikbaar voor het selecteren van de leerlingen die het meest in aanmerking komen voor afname van een praktische opdracht.

Hoe functioneerde het gebruikte schaaltype?

- 88% vond de frequentieschaal (*(bijna) nooit - soms - vaak - (bijna) altijd*) duidelijk en werkbaar;
- 32% vond de kwalitatieve beoordelingsschaal (*zeer zwak ontwikkeld - zwak ontwikkeld - matig ontwikkeld - sterk ontwikkeld*) duidelijk en 68% vond dat dat niet het geval was. De meeste leerkrachten gaven aan dat ze een optie misten tussen 'matig' en 'sterk'. Velen hadden een voorkeur voor een schaal met 5 antwoordmogelijkheden, met bijvoorbeeld de optie 'voldoende' tussen de opties 'matig' en de optie 'sterk'.

Op basis van de feedback van de leerkrachten zijn onder meer de volgende verbeteringen doorgevoerd:

- De checklist is ingekort van 87 naar 58 vragen. Hiermee is de afnametijd teruggebracht van ruim 12 tot gemiddeld 8 minuten.
- In de handleiding is uitgelegd waarom er voor een kwalitatieve beoordelingsschaal met de alternatieven zeer zwak, zwak, matig en sterk gekozen is en zijn aanwijzingen gegeven over hoe de leerkracht het beste kan omgaan met het verschil tussen 'matig' en 'sterk'.

4.3 Betrouwbaarheid van de conceptversie van de checklist

Zoals aangegeven in tabel 4.1 bevat de conceptversie van de checklist 87 vragen verdeeld over 7 a priori schalen.

Tabel 4.1 Aantal uitspraken, Cronbach's alfa en laagste, hoogste en gemiddelde item-restcorrelatie per a priori schaal

A priori schaal	Aantal uitspraken	Cronbach's alfa	Item-restcorrelaties		
			Min	Max	Gem
Oriënteren, voorbereiden en plannen van praktische taken	12	0,92	0,47	0,80	0,67
Uitvoeren van praktische taken	12	0,92	0,50	0,78	0,68
Monitoren, reflecteren en evalueren van praktische taken	15	0,96	0,60	0,82	0,75
Aandachtig werken	12	0,94	0,17	0,87	0,72
Gemotiveerd werken	12	0,94	0,52	0,83	0,71
Werken zonder gedragsproblematiek	12	0,91	0,43	0,77	0,63
Zelfstandig werken en hulp vragen	12	0,92	0,53	0,74	0,66
Totaal	87	0,99	0,21	0,80	0,65

Per uitspraak kon de leerkracht kiezen uit 4 antwoordmogelijkheden: 1 = (bijna) nooit; 2 = soms; 3 = vaak; 4 = (bijna) altijd of 1 = zeer zwak ontwikkeld; 2 = zwak ontwikkeld; 3 = matig ontwikkeld; 4 = sterk ontwikkeld. De respondenten is gevraagd de antwoordmogelijkheid 'n.t.b' en/of 'n.v.t.' te omcirkelen als de vraag niet te beantwoorden was of niet van toepassing was. De percentages NTB of NVT zijn belangrijke indicatoren van mogelijk disfunctionerende uitspraken. Bij geen van de uitspraken was het percentage NTB en NVT groter dan 5%. Er was dus geen aanleiding om op basis hiervan uitspraken uit de vragenlijst te verwijderen. In de verdere analyses zijn NTB- en NVT-scores daarom als ontbrekende waarnemingen beschouwd (en dus niet vervangen door bijvoorbeeld het gemiddelde over de als 1, 2, 3 en 4 gecodeerde antwoordmogelijkheden).

Een inspectie van de frequentieverdelingen per uitspraak bracht aan het licht dat de antwoorden relatief gelijkmatig verdeeld waren over de vier antwoordmogelijkheden. Bij geen enkele uitspraak was de verdeling zo scheef dat er antwoorden voor een antwoordmogelijkheid ontbraken. Ook nu was er dus geen aanleiding om uitspraken op voorhand te verwijderen.

De betrouwbaarheid van de conceptversie is bepaald aan de hand van de interne consistentie ofwel homogeniteit (Cronbach's alfa) en de item-restcorrelaties. Tabel 4.1 laat zien dat de interne consistentie van alle zeven schalen zeer hoog is (alfa > 0,90). De betrouwbaarheidscoëfficiënten voldoen alle aan het COTAN-criterium voor de betrouwbaarheid van meetinstrumenten die bedoeld zijn voor het nemen van belangrijke beslissingen (Evers, Lucassen, Meijer & Sijtsma, 2010), zoals de verwijzing naar een ander schooltype. Maar zoals eerder betoogd, is de checklist daar niet voor bedoeld.

Volgens de COTAN-normen geldt een item-totaalcorrelatie (Rit) van 0,19 of lager als onvoldoende, tussen 0,20 en 0,29 als voldoende en 0,30 of hoger als goed (Evers e.a., 2010). Passen we de COTAN-normen voor de Rit toe op onze (lagere) item-restcorrelaties (Rir) dan heeft slechts een uitspraak een Rir die hoger is dan de ondergrenswaarde die de COTAN geeft om een uitspraak als goed te beoordelen (waarbij we opmerken dat het oordeel over de kwaliteit van die ene dissonante uitspraak nog steeds 'voldoende' is). De zeer hoge item-restcorrelaties zijn deels te danken aan het gegeven dat het hier niet gaat om dichotome goed-foutscores (zoals bij toetsitems gebruikelijk is), maar om vier-puntschalen. Desalniettemin kunnen we concluderen dat er sprake is van in psychometrisch opzicht goede uitspraken.

4.4 Samenstelling van de definitieve checklist

Zoals eerder betoogd, hebben we meer uitspraken in de conceptversie opgenomen dan voor een betrouwbare meting naar verwachting nodig zouden zijn. De reden is dat wij meer wilden weten over de interne structuur van de vaardigheid in leren leren en wilden bepalen hoeveel uitspraken er nodig zijn om de vaardigheden betrouwbaar en valide te kunnen meten. Het doel was informatie te verzamelen op basis waarvan het aantal uitspraken met zo weinig mogelijk verlies aan informatie zo ver mogelijk kon worden teruggebracht.

Uitgevoerde analyses

Voor de samenstelling van de ingekorte schalen voor de uiteindelijke checklist zijn onder meer de volgende analyses uitgevoerd:

- analyse van verdelingskenmerken van de schalen (o.a. gemiddelden, spreiding, scheefheid, kurtosis);
- analyse van het discriminerend vermogen van uitspraken (item-restcorrelaties) en de interne consistentie van schalen (Cronbach's alfa);
- analyse van de correlaties tussen de uitspraken van een schaal met de uitspraken van de overige schalen (teneinde 'verkeerd' geplaatste uitspraken te detecteren);
- analyse van de uni-dimensionaliteit van de schalen door middel van exploratieve principale componentenanalyse over de uitspraken van de individuele schalen;
- analyse van de factorzuiverheid van de schalen door middel van exploratieve principale componentenanalyse over de uitspraken van meerdere schalen (met het doel schalen samen te stellen die de onderscheiden vaardigheden zo zuiver mogelijk meten);
- analyse van de correlaties tussen de schalen ter bestudering van de dimensionaliteit van de meting van de vaardigheid in leren leren;
- analyses van de relaties met achtergrondvariabelen zoals geslacht, leeftijd, leerjaar, schooltype en uitstroomprofiel.

Op basis van de analyseresultaten en inhoudelijke overwegingen zijn uiteindelijk zeven schalen samengesteld: drie 'cognitieve' schalen voor 'Oriënteren, voorbereiden en plannen', 'Uitvoeren' en 'Monitoren, reflecteren en evalueren' en vier schalen voor Werkhouding te weten 'Aandachtig werken', 'Gemotiveerd werken', 'Werken zonder gedragsproblematiek' en 'Zelfstandig werken en hulp vragen'. De precieze samenstelling van deze schalen is beschreven in paragraaf 4.5.1, de psychometrische kenmerken in paragraaf 4.5.2 en de omzetting van somscores naar percentielen in paragraaf 4.5.3 en bijlage 4.

4.5 De definitieve checklist

4.5.1 De schalen en uitspraken van de definitieve checklist

De uiteindelijke checklist bestaat uit zes schalen met acht uitspraken en één schaal met tien uitspraken. De eerste kolom van tabel 4.2 toont voor elke schaal de uiteindelijk geselecteerde uitspraken. Een inhoudelijke inspectie van de uitspraken per schaal laat zien dat er geen sprake meer is van storende overlap (oftewel van uitspraken die in inhoudelijk opzicht te sterk op elkaar lijken). In het rechter gedeelte van tabel 4.2 is ook de procentuele verdeling van de antwoorden over de vier antwoordmogelijkheden gepresenteerd. Voor alle uitspraken geldt dat elk van de vier antwoordmogelijkheden door een deel van de respondenten gekozen is.

Tabel 4.2 De uitspraken en enkele beschrijvende gegevens per schaal per uitspraak van de uiteindelijke checklist

	N	Min	Max	Gem	SD					
						1	2	3	4	
Oriënteren, voorbereiden en plannen van praktische taken										
Is betrokken bij de instructie in een groepssituatie.	622	1	4	2,80	0,87	6%	32%	38%	24%	
Leest de opdracht vóórdat hij start met de opdracht.	621	1	4	2,48	0,85	12%	39%	37%	12%	
Legt het benodigde materiaal voor de opdracht klaar.	622	1	4	2,51	0,91	14%	35%	37%	14%	
Kan uitleggen wat hij gaat doen.	617	1	4	2,94	0,87	6%	22%	43%	29%	
Kan uitleggen hoe hij de opdracht gaat aanpakken.	616	1	4	2,73	0,89	9%	29%	41%	21%	
Maakt een realistische planning voor een opdracht.	573	1	4	2,36	0,91	20%	35%	36%	10%	
Is in staat prioriteiten te stellen.	593	1	4	2,56	0,96	15%	32%	35%	18%	
Kan vooraf benoemen met welk eindresultaat hij tevreden is.	573	1	4	2,59	0,91	13%	31%	40%	16%	
Uitvoeren van praktische taken										
Raadpleegt de opdracht opnieuw wanneer hij vastloopt.	621	1	4	2,32	0,91	20%	39%	30%	11%	
Voert de stappen van een opdracht in de aangegeven volgorde uit.	622	1	4	3,07	0,80	3%	19%	45%	33%	
Probeer een gemaakte fout te herstellen.	618	1	4	2,88	0,92	9%	23%	40%	28%	
Past na aanwijzing van de leerkracht zijn werkhouding aan.	612	1	4	3,12	0,81	3%	18%	43%	36%	
Lost een taakinhoudelijk probleem zelf op.	621	1	4	2,53	0,90	13%	35%	37%	14%	
Gebruikt de beschikbare tijd op een efficiënte manier.	616	1	4	2,65	0,96	14%	28%	37%	21%	
Kijkt hoe de leerkracht een taak aanpakt en neemt deze aanpak over.	615	1	4	2,92	0,82	4%	25%	46%	25%	
Vult zijn werk aan als hij iets vergeten is.	620	1	4	2,80	0,95	10%	26%	37%	26%	
Monitoren, reflecteren en evalueren van praktische taken										
Houdt in de gaten of de opdracht volgens plan verloopt.	608	1	4	2,59	0,91	13%	31%	40%	16%	
Vraagt zich bij een fout af: "Hoe komt het dat ik dit fout doe"?	613	1	4	2,45	0,95	19%	32%	35%	15%	
Beoordeelt of zijn voorbereiding goed is geweest.	561	1	4	2,14	0,91	29%	33%	32%	6%	
Vergelijkt zijn eigen eindproduct met de opdracht (of het voorbeeld).	579	1	4	2,63	0,94	13%	31%	37%	19%	
Beoordeelt het eindresultaat.	590	1	4	2,67	0,94	12%	29%	37%	21%	
Onderbouwt zijn beoordeling van het eindresultaat.	567	1	4	2,38	0,94	20%	33%	34%	12%	
Benoemt wat goed is gegaan.	619	1	4	2,85	0,86	7%	26%	43%	24%	
Benoemt wat niet goed is gegaan.	615	1	4	2,84	0,86	6%	29%	41%	25%	
Geeft aan wat hij heeft geleerd van de opdracht.	610	1	4	2,60	0,90	13%	29%	43%	15%	
Benoemt wat hij de volgende keer anders zou doen.	598	1	4	2,53	0,88	13%	35%	39%	13%	
Aandachtig werken										
Luistert geconcentreerd naar een langere instructie.	620	1	4	2,53	0,97	16%	34%	31%	19%	
Richt zijn aandacht - na het starten - op het uitvoeren van de opdracht.	622	1	4	2,90	0,84	5%	26%	43%	26%	
Richt na afleiding zijn aandacht opnieuw op de opdracht.	621	1	4	2,70	0,82	6%	36%	40%	18%	
Kan zijn aandacht bij het werk houden tot de opdracht af is. (spanningsboog)	621	1	4	2,77	0,98	13%	24%	37%	26%	
Werkt aandachtig in een rustige groepssituatie.	619	1	4	3,16	0,89	5%	18%	33%	44%	
Kan geconcentreerd aan een korte opdracht werken.	621	1	4	3,29	0,82	4%	13%	35%	49%	
Werkt geconcentreerd aan opdrachten die hij niet leuk vindt.	619	1	4	2,44	0,95	19%	32%	36%	14%	
Kan geconcentreerd aan een langere opdracht werken.	620	1	4	2,74	0,94	11%	27%	38%	24%	
Gemotiveerd werken										
Werkt met enthousiasme aan een opdracht die aansluit bij zijn interesse.	621	1	4	3,35	0,69	1%	9%	43%	46%	
Werkt met betrokkenheid aan een opdracht die niet aansluit bij zijn interesse.	617	1	4	2,25	0,85	18%	46%	27%	8%	
Reageert op vragen die de leerkracht stelt tijdens de instructie.	620	1	4	3,30	0,77	2%	13%	38%	47%	
Werkt zelfstandig zonder aansporing van de leerkracht.	618	1	4	2,91	0,98	11%	20%	36%	33%	
Zet door ook als de opdracht moeilijk is.	619	1	4	2,57	0,96	16%	28%	39%	17%	
Zet door op het moment dat hij tegenslag ervaart.	622	1	4	2,43	0,96	20%	31%	35%	14%	
Maakt af waar hij aan begonnen is.	619	1	4	2,96	0,91	7%	23%	37%	33%	
Kan zichzelf motiveren voor de opdracht.	616	1	4	2,73	0,88	9%	29%	42%	20%	
Werken zonder gedragsproblematiek										
Stoort anderen tijdens het werken.	620	1	4	3,04	0,82	5%	16%	47%	31%	
Reageert impulsief.	621	1	4	2,64	0,95	13%	30%	37%	20%	
Vraagt continu om bevestiging van de leerkracht.	620	1	4	3,03	0,85	5%	21%	41%	33%	
Vraagt op ongepaste wijze hulp wanneer hij vastloopt.	621	1	4	3,09	0,93	7%	17%	35%	41%	
Vertoont storend gedrag als hij tegen een werkprobleem aanloopt.	622	1	4	2,98	0,97	11%	15%	40%	35%	
Kan omgaan met kritiek op gemaakt werk en/of de werkhouding.	617	1	4	2,51	0,91	16%	32%	39%	13%	
Zet door op het moment dat hij gefrustreerd raakt.	612	1	4	2,23	0,92	25%	36%	30%	9%	
Kan emoties in bedwang houden, zodat ze de uitvoering van de opdracht niet beïnvloeden.	586	1	4	2,36	0,97	22%	34%	30%	14%	
Zelfstandig werken en hulp vragen										
Past de afgesproken regels van het zelfstandig werken toe.	623	1	4	2,93	0,82	3%	28%	42%	27%	
Probeer eerst zelf het probleem op te lossen voor hij hulp vraagt.	623	1	4	2,51	0,91	15%	34%	37%	14%	
Vraagt hulp op een passend moment.	622	1	4	2,64	0,85	9%	35%	40%	16%	
Maakt bij de start van een opdracht een inschatting: Kan ik aan de slag of heb ik eerst hulp nodig?	613	1	4	2,66	0,91	12%	28%	42%	18%	
Werkt zelfstandig.	619	1	4	3,24	0,83	4%	14%	37%	45%	
Kan een hulpvraag goed verwoorden.	619	1	4	2,81	0,90	9%	25%	42%	24%	
Kan omgaan met uitgestelde aandacht.	622	1	4	2,87	0,95	10%	23%	38%	30%	

4.5.2 Beschrijvende gegevens per schaal van de definitieve checklist

Voor elk van de zeven schalen van de checklist is een totaalscore geconstrueerd door de scores over de desbetreffende uitspraken te sommeren. Enkele verdelingskenmerken van de schalen zijn weergegeven in tabel 4.3. Voor geen van de schalen geldt dat de scores uitzonderlijk scheef verdeeld zijn.

Tabel 4.3 Enkele verdelingskenmerken voor de definitieve checklist per schaal

Schaal	N	Min	Max	Gem	Std-dev	Scheefheid	P5	P10	P20	P50	P80	P90	P95
Oriënteren, voorbereiden en plannen	546	8	32	20,92	5,55	-0,10	11	13	16	21	26	28	30
Uitvoeren	598	8	32	22,27	5,51	-0,19	13	15	18	22	28	30	31
Monitoren, reflecteren en evalueren	529	10	40	25,70	7,35	-0,03	14	15	19	26	33	35	38
Aandachtig werken	603	8	32	22,50	6,15	-0,29	12	14	17	23	29	30	32
Gemotiveerd werken	605	8	32	22,45	5,45	-0,31	13	15	18	23	28	29	31
Werken zonder gedragsproblematiek	569	8	32	21,81	5,39	-0,25	12	14	17	22	27	29	30
Zelfstandig werken en hulp vragen	602	8	32	22,37	5,27	-0,16	13	15	18	23	27	29	31

4.5.3 De omzetting van somscores naar percentielen

Uit de leerling- en groepsrapportages (zie paragraaf 3.3.4) kan de leerkracht opmaken hoe de leerling respectievelijk de groep als geheel het doet in vergelijking met een landelijke referentiegroep van leerlingen uit sbo en so cluster 4. Hiertoe zijn de scores omgezet in percentielen die aangeven welk percentage van de leerlingen uit de referentiegroep dezelfde of een lagere score heeft behaald dan de leerling in kwestie. In bijlage 4 is aangegeven hoe de somscores zijn omgezet naar percentielen. Tabel 1 uit deze bijlage toont de omzetting voor de somscore op de checklist als geheel, tabel 2 voor de drie cognitieve schalen en tabel 3 voor de vier schalen van werkhouding.

4.6 Betrouwbaarheid van de definitieve checklist

De checklist is niet bedoeld om er belangrijke beslissingen mee te onderbouwen. Het doel reikt niet verder dan verkrijgen van een indruk van het niveau van leren leren in de groep als geheel en het selecteren van de leerlingen die het meest voor afname van een praktische opdracht in aanmerking komen. In dit hoofdstuk gaan we na in hoeverre de vastgestelde betrouwbaarheid voldoet aan de eisen die het COTAN Beoordelingssysteem hieraan stelt (Evers, Lucassen, Meijer & Sijtsma, 2010). Daarbij bespreken we twee aspecten: interne consistentie en meetnauwkeurigheid.

4.6.1 Interne consistentie en item-restcorrelaties

Op basis van de hiervoor besproken analyses zijn zeven schalen samengesteld door de toegekende scores over de desbetreffende uitspraken te sommeren. De betrouwbaarheid van de conceptversie is bepaald aan de hand van de interne consistentie ofwel homogeniteit (Cronbach's alfa) en de item-restcorrelaties. De betrouwbaarheid in termen van interne consistentie kan als zeer hoog worden aangemerkt ($0,88 < \text{alfa} < 0,94$). Volgens de COTAN-norm voor de betrouwbaarheid voor het nemen van minder zwaarwegende beslissingen op individueel niveau, zoals voortgangscontrole, hebben alle schalen een goede

betrouwbaarheid (Evers, Lucassen, Meijer & Sijtsma, 2010). De betrouwbaarheidscoëfficiënten van vijf van de zeven schalen voldoen zelfs aan het COTAN-criterium voor de betrouwbaarheid van meetinstrumenten die bedoeld zijn voor het nemen van belangrijke beslissingen (Evers e.a., 2010), zoals de verwijzing naar een ander schooltype. Maar zoals eerder betoogd, is de checklist daar niet voor bedoeld. De betrouwbaarheid van de totale schaal met alle 58 uitspraken bedraagt overigens 0,98.

Tabel 4.4 toont ook de item-restcorrelaties (Rit) per schaal per uitspraak. Volgens de COTAN-normen geldt een item-totaalcorrelatie (Rit) van 0,19 of lager als onvoldoende, tussen 0,20 en 0,29 als voldoende en 0,30 of hoger als goed (Evers e.a., 2010). Passen we deze norm toe op onze (lagere) item-restcorrelaties (Rir) dan heeft geen enkele uitspraak een Rir die lager is dan de ondergrenswaarde die de COTAN geeft om een uitspraak als goed te beoordelen. We kunnen concluderen dat er sprake is van in psychometrisch opzicht zeer goede uitspraken.

Tabel 4.4 Aantal uitspraken, Cronbach's alfa en laagste, hoogste en gemiddelde item-restcorrelatie per schaal van de checklist

Schaal	Aantal uitspraken	Cronbach's alfa	Item-restcorrelaties		
			Gem	Min	Max
Oriënteren, voorbereiden en plannen van praktische taken	8	0,90	0,69	0,54	0,78
Uitvoeren van praktische taken	8	0,91	0,71	0,65	0,75
Monitoren, reflecteren en evalueren van praktische taken	10	0,94	0,75	0,69	0,84
Aandachtig werken	8	0,94	0,80	0,74	0,87
Gemotiveerd werken	8	0,90	0,69	0,48	0,82
Werken zonder gedragsproblematiek	8	0,88	0,64	0,44	0,77
Zelfstandig werken en hulp vragen	8	0,88	0,65	0,49	0,71
Totaal	58	0,98	0,69	0,54	0,78

4.6.2 Standaardmeetfout

Tabel 4.5 toont de standaardmeetfout met het bijbehorend 90%- en 95%-betrouwbaarheidsinterval. Het bereik van plusminus (bijna) twee standaardmeetfouten rond de gemiddelde score geeft aan waar de ware score zich met 90%- of 95%-zekerheid zal bevinden. De gepresenteerde standaardmeetfouten zijn klein. Zo ligt de ware somscore voor de schaal Oriënteren, voorbereiden en plannen van praktische taken met 95% waarschijnlijkheid tussen $\pm 3,44$. De conclusie is dat de betrouwbaarheid van alle zeven schalen en van de checklist als geheel als zeer goed beoordeeld kan worden.

Tabel 4.5 Standaardmeetfout met bijbehorend 90%- en 95%-betrouwbaarheidsinterval

Schaal	Standaardmeetfout	Standaardmeetfout met 90%- en 95%- betrouwbaarheidsinterval	
		90%	95%
Oriënteren, voorbereiden en plannen van praktische taken	1,75	±2,88	±3,44
Uitvoeren van praktische taken	1,65	±2,72	±3,24
Monitoren, reflecteren en evalueren van praktische taken	1,80	±2,96	±3,53
Aandachtig werken	1,51	±2,48	±2,95
Gemotiveerd werken	1,72	±2,84	±3,38
Werken zonder gedragsproblematiek	1,87	±3,07	±3,66
Zelfstandig werken en hulp vragen	1,82	±3,00	±3,58
Totaalscore over alle 58 uitspraken	4,99	±8,20	±9,78

4.7 Validiteit van de definitieve checklist

In deze paragraaf doen we verslag van de uitkomsten van analyses die zijn uitgevoerd met het oog op het vaststellen van de begripsvaliditeit van de uiteindelijke checklist. Het doel reikt niet verder dan verkrijgen van een indruk van het niveau van leren leren in de groep als geheel en het selecteren van de leerlingen die het meest voor afname van een praktische opdracht in aanmerking komen. Omdat de checklist niet bedoeld is voor 'voorspellend gebruik' blijven voorspellende kwaliteiten en criteriumvaliditeit in deze paragraaf verder onbesproken.

Er is op verschillende manieren onderzoek gedaan naar de begripsvaliditeit. Eerst doen we verslag van een analyse van de interne structuur van de meting van leren leren. Daarbij gaan we in op de vraag in hoeverre het gelukt is om voor de onderscheiden aspecten van de vaardigheid in leren leren uni-dimensionele schalen te construeren. Vervolgens kijken we met behulp van exploratieve factoranalyse naar de factor-zuiverheid ofwel de empirische onderscheidbaarheid van de veronderstelde aspecten van leren leren. Daarna bespreken we de correlaties tussen de schalen zoals deze door de factoranalyses in grote lijnen bevestigd werden. Tot slot gaan we in op verschillen tussen groepen leerlingen.

4.7.1 Unidimensionaliteit

Met het oog op een zo eenduidig mogelijke interpretatie van de scores door de gebruiker is tijdens de constructiefase veel aandacht besteed aan het ontwikkelen van betrouwbare en uni-dimensionele schalen. In paragraaf 4.6 hebben we laten zien dat de betrouwbaarheid van alle zeven schalen op basis van de interne consistentie (Cronbachs alfa) en de item-restcorrelaties zeer goed te noemen is. Als check op de nagestreefde uni-dimensionaaliteit zijn wij met behulp van exploratieve factoranalyse – principale componentenanalyse met varimaxrotatie – nagegaan in hoeverre de uitspraken inderdaad slechts één factor meten dan wel een beroep doen op meerdere factoren. Tabel 4.6 toont voor elke schaal het aantal uitspraken en het aantal gevonden factoren met het bijbehorende percentage verklaarde variantie.

Tabel 4.6 Resultaten van de exploratieve factoranalyse per schaal

Schaal	Aantal uitspraken	Aantal factoren	Percentage verklaarde variantie	
			Factor 1	Factor 2
Oriënteren, voorbereiden en plannen van praktische taken	8	1	59	
Uitvoeren van praktische taken	8	1	61	
Monitoren, reflecteren en evalueren van praktische taken	10	1	64	
Aandachtig werken	8	1	71	
Gemotiveerd werken	8	2	60	13
Werken zonder gedragsproblematiek	8	2	55	15
Zelfstandig werken en hulp vragen	8	1	65	

Van vijf van de zeven schalen blijken de uitspraken slechts één factor te meten (zie tabel 4.6). De beide uitzonderingen zijn Gemotiveerd werken en Werken zonder gedragsproblematiek die een beroep doen op twee factoren (waarbij de tweede factor duidelijk minder variantie naar zich toetrekt dan de eerste factor).

4.7.2 Factorzuiverheid

Hiervoor hebben we vastgesteld dat vijf van de zeven schalen voor de meting van de vaardigheden in leren leren uni-dimensioneel zijn. Naast homogeniteit en uni-dimensionaliteit is tijdens de constructie ook aandacht besteed aan het verkrijgen van factor-zuivere schalen, dit wil zeggen schalen waarvan de inhoud zo weinig mogelijk overlapt met die van andere schalen. De meting van de vaardigheid in leren leren geldt als factor-zuiver als de veronderstelde aspecten (schalen) in de gegevens terug te vinden zijn. In deze paragraaf onderzoeken we de factor-zuiverheid ofwel de empirische onderscheidbaarheid van de definitieve schalen. Dat doen we met behulp van exploratieve factoranalyse: principale componenten-analyse met varimax rotatie, uitgevoerd op de gegevens van de proefafname. We beschouwen een schaal als factor-zuiver als de ladingen op de factor die de schaal pretendeert te meten hoog zijn (d.w.z. $\geq 0,40$) en de ladingen op de 'eigen' factor hoger zijn dan die op andere factoren die de schaal niet beoogt te meten.

Omdat het aantal uitspraken relatief groot is in verhouding tot het aantal respondenten, hebben we de factoranalyses afzonderlijk uitgevoerd voor de uitspraken over zelfregulerende vaardigheden en over werkhouding.

4.7.2.1 Factorzuiverheid van de meting van zelfregulerende vaardigheden

Op basis van de theoretische overwegingen die aan de constructie ten grondslag lagen (zie hoofdstuk 2) zijn drie a priori schalen ontwikkeld: Oriënteren, voorbereiden en plannen, Uitvoeren en Monitoren, reflecteren en evalueren. In de definitieve checklist zijn deze 3 vaardigheden met respectievelijk 8, 8 en 10 uitspraken vertegenwoordigd. Tabel 4.7 toont de uitkomsten van een exploratieve factoranalyse op alle 26 uitspraken. Er worden 3 factoren gevonden die tezamen 65% van de variantie in leren leren naar zich toetrekken. De verklaarde variantie voor factor 1, 2 en 3 is achtereenvolgens 55%, 6% en 4%. Alle uitspraken hebben ten minste één factorlading die hoger is dan de ondergrens van 0,40.

Tabel 4.7 *Uitkomsten van een exploratieve factoranalyse van de uitspraken van de definitieve checklist*

	F1	F2	F3
Oriënteren, voorbereiden en plannen			
Is betrokken bij de instructie in een groepssituatie.		0,61	
Leest de opdracht vóórdat hij start met de opdracht.		0,64	
Legt het benodigde materiaal voor de opdracht klaar.		0,49	0,50
Kan uitleggen wat hij gaat doen.			0,73
Kan uitleggen hoe hij de opdracht gaat aanpakken.			0,74
Maakt een realistische planning voor een opdracht.			0,70
Is in staat prioriteiten te stellen.		0,47	0,63
Kan vooraf benoemen met welk eindresultaat hij tevreden is.	0,55		0,51
Uitvoeren			
Raadpleegt de opdracht opnieuw wanneer hij vastloopt.		0,59	
Voert de stappen van een opdracht in de aangegeven volgorde uit.		0,67	
Probeert een gemaakte fout te herstellen.	0,45	0,68	
Past na aanwijzing van de leerkracht zijn werkhouding aan.		0,71	
Lost een taakinhoudelijk probleem zelf op.		0,50	0,49
Gebruikt de beschikbare tijd op een efficiënte manier.		0,67	0,42
Kijkt hoe de leerkracht een taak aanpakt en neemt deze aanpak over.	0,42	0,63	
Vult zijn werk aan als hij iets vergeten is.		0,66	
Monitoren, reflecteren en evalueren			
Houdt in de gaten of de opdracht volgens plan verloopt.	0,53	0,47	0,43
Vraagt zich bij een fout af: "Hoe komt het dat ik dit fout doe"?	0,64	0,44	
Beoordeelt of zijn voorbereiding goed is geweest.	0,61		0,41
Vergelijkt zijn eigen eindproduct met de opdracht (of het voorbeeld).	0,73		
Beoordeelt het eindresultaat.	0,75		
Onderbouwt zijn beoordeling van het eindresultaat.	0,79		
Benoemt wat goed is gegaan.	0,59		0,48
Benoemt wat niet goed is gegaan.	0,67		
Geeft aan wat hij heeft geleerd van de opdracht.	0,64		0,43
Benoemt wat hij de volgende keer anders zou doen.	0,66		

Het patroon van factorladingen lijkt tot op zekere hoogte overeen te komen met de beoogde structuur. De eerste factor, die de meeste variantie aan zich bindt, wordt gedomineerd door de 10 uitspraken van de vaardigheid 'Monitoren, reflecteren en evalueren'. Op de tweede factor hebben alle 8 uitspraken over de vaardigheid 'Uitvoeren' hoge ladingen. De derde factor wordt grotendeels bepaald door 6 van de 8 uitspraken van 'Oriënteren, voorbereiden en plannen'. De factorzuiverheid wordt echter enigszins verstoord door de volgende afwijkingen van de beoogde structuur:

- 12 van de 26 uitspraken laden behalve op de 'eigen' factor ook op 1 of 2 van de beide overige factoren;
- wel zijn de bij-ladingen behoudens 3 uitzonderingen meestal aanzienlijk lager dan de ladingen op de 'eigen' factor;
- de eerste uitzondering doet zich voor bij de uitspraak 'Kan vooraf benoemen met welk eindresultaat hij tevreden is'; deze uitspraak beoogt vooral de vaardigheid Oriënteren, voorbereiden en plannen te meten, maar laadt iets hoger op Monitoren, reflecteren en evalueren dan op de 'eigen' factor (0,55 versus 0,51);
- de tweede uitzondering betreft de uitspraak 'Legt het benodigde materiaal voor de opdracht klaar'; deze uitspraak was ingedeeld bij Oriënteren, voorbereiden en plannen maar laadt ongeveer even hoog op de factor voor Uitvoeren (0,50 versus 0,49);
- de derde uitzondering is 'Lost een taakinhoudelijk probleem zelf op'; deze uitspraak valt onder 'Uitvoeren' maar laadt ook bijna even hoog op Oriënteren, voorbereiden en plannen (0,50 tegen 0,49);

- de tweede factor die we eerder als Uitvoeren benoemden, is gecontamineerd met 2 uitspraken over Oriënteren, voorbereiden en plannen, te weten 'Is betrokken bij de instructie in een groepssituatie' (0,61) en 'Leest de opdracht vóórdat hij start met de opdracht' (0,64).

Al met al wijkt het gevonden patroon van factorladingen enigszins af van het veronderstelde patroon. De 3 schalen vertonen meer inhoudelijke overlap dan op voorhand was ingeschat. De zuiverheid van de 3 veronderstelde factoren biedt ruimte voor verbetering. De afwijkingen van de veronderstelde structuur zijn echter niet extreem groot, terwijl de bij-ladingen overwegend weinig hoger zijn dan de ondergrens van 0,40. Vandaar dat ondanks de geconstateerde onvolkomenheden besloten is om de verdeling van de uitspraken over de 3 cognitieve schalen van leren leren intact te laten.

4.7.2.2 Factorzuiverheid van de meting van werkhouding

Tabel 4.8 toont de uitkomsten van een exploratieve factoranalyse op alle 32 uitspraken die werkhouding betreffen. Omdat we op basis van de bestudeerde literatuur 4 factoren veronderstellen, is gekozen voor een gedwongen vier-factoroplossing. De 4 factoren trekken tezamen 67% van de variantie naar zich toe. De verklaarde variantie voor factor 1, 2, 3 en 4 is achtereenvolgens 49%, 9%, 5% en 4%. Alle uitspraken hebben ten minste 1 factorlading die hoger is dan de ondergrens van 0,40.

Tabel 4.8 *Uitkomsten van een exploratieve factoranalyse van de uitspraken in de definitieve checklist voor Werkhouding (gedwongen vier-factoroplossing)*

	1	2	3	4
Aandachtig werken				
Luistert geconcentreerd naar een langere instructie.	0,55			0,47
Richt zijn aandacht - na het starten - op het uitvoeren van de opdracht.	0,72			
Richt na afleiding zijn aandacht opnieuw op de opdracht.	0,71			
Kan zijn aandacht bij het werk houden tot de opdracht af is (spanningsboog).	0,82			
Werkt aandachtig in een rustige groepssituatie.	0,76			
Kan geconcentreerd aan een korte opdracht werken.	0,77			
Werkt geconcentreerd aan opdrachten die hij niet leuk vindt.	0,53		0,41	
Kan geconcentreerd aan een langere opdracht werken.	0,80			
Gemotiveerd werken				
Werkt met enthousiasme aan een opdracht die aansluit bij zijn interesse.				0,74
Werkt met betrokkenheid aan een opdracht die niet aansluit bij zijn interesse.			0,40	0,47
Reageert op vragen die de leerkracht stelt tijdens de instructie.	0,42			0,65
Werkt zelfstandig zonder aansporing van de leerkracht.	0,83			
Zet door ook als de opdracht moeilijk is.	0,49	0,69		
Zet door op het moment dat hij tegenslag ervaart.	0,43	0,76		
Maakt af waar hij aan begonnen is.	0,68	0,43		
Kan zichzelf motiveren voor de opdracht.	0,52	0,51		
Werken zonder gedragsproblematiek				
Stoort anderen tijdens het werken.			0,79	
Reageert impulsief.			0,79	
Vraagt continu om bevestiging van de leerkracht.		0,48		
Vraagt op ongepaste wijze hulp wanneer hij vastloopt.			0,79	
Vertoont storend gedrag als hij tegen een werkprobleem aanloopt.			0,79	
Kan omgaan met kritiek op gemaakt werk en/of de werkhouding.		0,64		
Zet door op het moment dat hij gefrustreerd raakt.		0,74		
Kan emoties in bedwang houden, zodat ze de uitvoering van de opdracht niet beïnvloeden.		0,71		
Zelfstandig werken en hulp vragen				
Past de afgesproken regels van het zelfstandig werken toe.	0,57		0,41	
Probeert eerst zelf het probleem op te lossen voor hij hulp vraagt.	0,52	0,48		
Vraagt hulp op een passend moment.			0,47	
Maakt bij de start van een opdracht een inschatting: Kan ik aan de slag of heb ik eerst hulp nodig?	0,48			
Werkt zelfstandig.	0,76			
Kan een hulpvraag goed verwoorden.				0,54
Kan omgaan met uitgestelde aandacht.		0,50	0,48	
Werkt aan een alternatieve opdracht als hij niet verder kan.	0,55			

Het patroon van factorladingen kan als volgt worden beschreven:

- Op de eerste factor, die bijna de helft van de variantie naar zich toetrekt, hebben 6 van de 8 uitspraken over Aandachtig werken zeer hoge factorladingen. Zeer hoge ladingen zien we ook voor de uitspraak 'Werkt zelfstandig zonder aansporing van de leerkracht.' uit de schaal Gemotiveerd werken en de uitspraak 'Werkt zelfstandig.' uit de schaal Zelfstandig werken en hulp vragen. Deze eerste factor zouden we dan ook met enige goede wil kunnen benoemen als Aandachtig werken met als

kanttekening dat er onbedoeld ook enkele aspecten van Gemotiveerd werken en Zelfstandig werken zonder gedragsproblematiek worden 'meegemeten'.

- De tweede factor wordt gedomineerd door twee uitspraken over doorzettingsvermogen uit de schaal Gemotiveerd werken ('Zet door ook als de opdracht moeilijk is.' en 'Zet door op het moment dat hij tegenslag ervaart.') en drie positief geformuleerde uitspraken uit de schaal Werken zonder gedragsproblematiek ('Zet door op het moment dat hij gefrustreerd raakt.', 'Kan emoties in bedwang houden, zodat ze de uitvoering van de opdracht niet beïnvloeden.' en 'Kan omgaan met kritiek op gemaakt werk en/of werkhouding.'). De drie uitspraken over doorzettingsvermogen lijken de gemeenschappelijke noemer van deze schaal te vormen. De kern ervan zou men met de nodige slagen om de arm kunnen benoemen als gemotiveerd werken in de zin van doorzettingsvermogen (in de wetenschap dat daarnaast onbedoelde aspecten zoals het kunnen omgaan met kritiek en impulsbeheersing worden 'meegemeten');
- Op de derde factor hebben vier van de vijf negatief geformuleerde uitspraken over Werken zonder gedragsproblematiek zeer hoge factorladingen ('Stoort anderen tijdens het werken.', 'Reageert impulsief.', 'Vraagt op ongepaste wijze hulp wanneer hij vastloopt.' en 'Vertoont storend gedrag als hij tegen een werkprobleem aanloopt.'). Deze factor zou men dan ook met enige terughoudendheid kunnen benoemen als Afwezigheid van storende gedragsproblematiek.
- De uitspraken die op de vierde factor laden, lijken vanuit inhoudelijk perspectief weinig gemeenschappelijk te hebben waarmee deze factor als oninterpreteerbaar moet worden beschouwd;
- 11 van de 32 uitspraken laden op twee factoren tegelijkertijd waarbij de bij-lading soms hoger is dan die op de 'eigen' factor.

Al met al wijkt het gevonden patroon van factorladingen af van de veronderstelde patroon. Met name de vierde factor lijkt lastig te interpreteren. Voor het overige lijkt er grosso modo sprake van de volgende drie factoren:

- een eerste factor voor Aandachtig werken;
- een tweede factor voor Gemotiveerd werken in de zin van doorzettingsvermogen;
- een derde factor voor Afwezigheid van storende gedragsproblematiek.

Al met al moeten we concluderen dat het gevonden patroon van factorladingen enigszins afwijkt van het veronderstelde patroon waarbij elke uitspraak uniek op de 'eigen' factor zou moeten laden. Het streven naar inhoudelijk relevante en gevarieerde schalen is hier ten koste gegaan van de empirische onderscheidbaar. Een te sterk accent op inhoudelijke overwegingen kan ertoe leiden dat de samenhang tussen de schalen zo hoog wordt dat het voor de rangordening van de leerlingen niet meer uitmaakt met welke schaal de vaardigheid van leren leren gemeten wordt. De schalen zijn in dat geval volledig uitwisselbaar. In de volgende paragraaf gaan we na in hoeverre de geconstateerde inhoudelijke overlap tot uiting komt in een te hoge samenhang tussen de schalen.

4.7.3 Interne structuur

De begripsvaliditeit van de meting van de vaardigheid in leren leren moet onder meer blijken uit de interne structuur van de meting. Als de veronderstelde of beoogde structuur hout snijdt, zou die ook in de gegevens terug te vinden moeten zijn. In de toetsingsfase van het valideringsproces (Cronbach, 1970) zijn er tenminste drie situaties die aanleiding geven om de begripsvaliditeit van de meting van de vaardigheid in leren leren in twijfel te trekken:

- Ten eerste kunnen de schalen die de onderscheiden aspecten van de vaardigheid in leren leren pretenderen te meten perfect samenhangen. Het onderscheid tussen de aspecten is dan empirisch betekenisloos. Dit kan aanleiding zijn om het instrument (en de eraan ten grondslag liggende theorie) te vereenvoudigen door het onderscheid tussen de desbetreffende aspecten te laten vallen.
- Ten tweede kan er sprake zijn van een niet perfecte samenhang van de aspecten, terwijl de richting van de samenhang anders is dan op grond van theorie te verwachten valt. Ook dan is er aanleiding om de validiteit van de meting en/of de onderliggende theorie in twijfel te trekken.

- Ten derde kunnen aspecten geen enkele of slechts een geringe samenhang vertonen. Ze kunnen dan niet meer worden beschouwd als behorend tot hetzelfde overkoepelende begrip. Het ware dan beter om te spreken van afzonderlijke, op zichzelf staande begrippen.

Hiervóór hebben we beargumenteerd dat er voor vijf van de zeven schalen geen enkele reden is om aan de uni-dimensionaliteit van de meting van de onderscheiden aspecten van leren leren te twijfelen.

De resultaten van de factoranalyses bleken grosso modo overeen te komen met de op basis van theoretische argumenten gebaseerde verdeling van de 56 uitspraken over de zeven schalen. Wel was er vooral bij de vier schalen van werkhouding sprake van enige conceptuele overlap. In deze sectie gaan we aan de hand van de correlaties tussen de schalen na in hoeverre de geconstateerde inhoudelijke overlap ten koste is gegaan van de onderscheidbaarheid van de onderscheiden aspecten van leren leren.

De correlaties tussen de zeven schalen van de definitieve checklist zijn weergegeven in tabel 4.9. Over het algemeen is sprake van een middelmatig tot hoge samenhang tussen de schalen. Een buitenbeentje is de schaal Werken zonder gedragsproblematiek die relatief laag correleert met vooral Oriënteren, voorbereiden en plannen ($r = 0,43$) en Monitoren, reflecteren en evalueren ($r = 0,41$). Mogelijk weerhoudt gedragsproblematiek leerlingen ervan hun vaardigheden in leren leren toe te passen.

De samenhang is nergens zo hoog dat aan de onderscheidbaarheid van de onderscheiden aspecten van leren leren getwijfeld moet worden. Het hoogst is de samenhang tussen de drie cognitieve vaardigheden van leren leren ($0,80 < r < 0,83$), tussen Gemotiveerd werken en Uitvoeren ($r = 0,81$), tussen Gemotiveerd werken en Aandachtig werken ($r = 0,85$), tussen Zelfstandig werken en hulp vragen en Uitvoeren ($r = 0,81$), tussen Zelfstandig werken en hulp vragen en Aandachtig werken ($r = 0,81$) en tussen Zelfstandig werken en hulp vragen en Gemotiveerd werken ($r = 0,84$). De deels hoge correlaties tussen de aspecten van Werkhouding zijn in overeenstemming met de beperkte onderscheidbaarheid van deze factoren zoals dit in de factoranalyses is vastgesteld (zie paragraaf 4.7).

Tabel 4.9 Samenhang tussen de oorspronkelijke en ingekorte schalen

	1	2	3	4	5	6	7
1 Oriënteren, voorbereiden en plannen van praktische taken	1,00						
2 Uitvoeren van praktische taken	0,83	1,00					
3 Monitoren, reflecteren en evalueren van praktische taken	0,82	0,80	1,00				
4 Aandachtig werken	0,73	0,77	0,59	1,00			
5 Gemotiveerd werken	0,76	0,81	0,69	0,85	1,00		
6 Werken zonder gedragsproblematiek	0,43	0,56	0,41	0,60	0,63	1,00	
7 Zelfstandig werken en hulp vragen	0,77	0,81	0,70	0,81	0,84	0,67	1,00

4.7.4 Samenhang met achtergrondkenmerken

De vaardigheid in leren leren van leerlingen in het speciaal onderwijs is een nieuw en nog nauwelijks ontgonnen onderzoeksgebied. Over de samenhang van vaardigheden in leren leren met (achtergrond)kenmerken van scholen, leerkrachten en leerlingen is nog weinig bekend. We hebben dan ook geen uitgesproken verwachtingen over eventuele verschillen. Toch vinden we het belangrijk om hier de uitkomsten van een analyse van de samenhang met achtergrondkenmerken te presenteren. De gebruiker van het pakket Leren leren sbo en so cluster 4 kan deze informatie immers gebruiken bij de interpretatie van de resultaten. Bij de evaluatie van de grootte van de verschillen tussen groepen baseren we ons op de vuistregel van Cohen (1977) waarbij 0,20, 0,50 en 0,80 respectievelijk een klein, middelgroot en groot verschil vertegenwoordigt. Om van een duidelijke dalende of stijgende trend tussen opeenvolgende antwoordmogelijkheden te kunnen spreken moeten de opeenvolgende effectgroottes groter zijn dan +0,20 of -0,20. In de effectgrootte is het verschil tussen twee gemiddelden uitgedrukt als een proportie van de standaarddeviatie.

4.7.4.1 Geslacht

Naar sekseverschillen op het gebied van leren leren in het speciaal onderwijs heeft voor zover ons bekend geen theorievorming of onderzoek plaatsgevonden. Vandaar dat wij het niet hebben aangedurfd om van te voren toetsbare hypothesen te formuleren met betrekking tot de richting en de sterkte van eventuele verschillen tussen jongens en meisjes waar het de vaardigheid in leren leren betreft. De verdeling van de deelnemende leerlingen naar geslacht in de responsgroep is 452 jongens, 139 meisjes en van 38 leerlingen is het geslacht onbekend. Tabel 4.10 toont voor elk van de definitieve schalen de gemiddelden en standaarddeviaties voor jongens en meisjes. De gemiddelden en standaarddeviaties zijn lastig rechtstreeks met elkaar te vergelijken. Daarom rapporteren we ook de effectgroottes. Daarnaast presenteren we de p-waarde en *Eta squared*. De p-waarde geeft de kans dat het verschil tussen de beide gemiddelden op toeval berust. *Eta squared* is een maat voor de proportie variantie in de vaardigheid leren leren die door geslacht verklaard wordt.

Tabel 4.10 Verschillen tussen jongens en meisjes per schaal

	jongen		meisje		Effect-grootte	p	Eta ²
	GEM	SD	GEM	SD			
Oriënteren, voorbereiden en plannen	20,75	5,67	21,45	5,18	-0,13	0,234	0,00
Uitvoeren	21,91	5,53	23,63	5,38	-0,31	0,002	0,02
Monitoren, reflecteren en evalueren	25,47	7,50	26,37	7,10	-0,12	0,258	0,00
Aandachtig werken	22,06	6,13	24,09	5,86	-0,34	0,001	0,02
Gemotiveerd werken	22,07	5,41	23,77	5,35	-0,32	0,002	0,02
Werken zonder gedragsproblematiek	21,31	5,42	23,52	5,13	-0,42	< 0,001	0,03
Zelfstandig werken en hulp vragen	21,85	5,37	24,05	4,85	-0,43	< 0,001	0,03

Eerst kijken we naar de cognitieve aspecten van leren leren. Meisjes scoren gemiddeld hoger op de vaardigheid Uitvoeren dan jongens. De grootte van het verschil houdt ongeveer het midden tussen een klein en middelgroot verschil. Geen sekseverschil zien we daarentegen bij Oriënteren, voorbereiden en plannen en Monitoren, reflecteren en evalueren.

Op alle vier aspecten van de werkhouding tijdens het werken aan praktische opdrachten zijn meisjes gemiddeld vaardiger dan jongens: Aandachtig werken, Gemotiveerd werken, Werken zonder gedragsproblematiek en Zelfstandig werken en hulp vragen. De effectgroottes bevinden zich in het bereik van een klein en een middelgroot verschil. Deze uitkomsten zijn in overeenstemming met hetgeen op basis van onderzoek in het regulier basisonderwijs verwacht mag worden en vormen een ondersteuning voor de validiteit van de meting (o.a. Kuhlemeier, Knoop, Van Boxtel, Papenburg & Hollenberg, 2016).

Een verklaring voor het ontbreken van vaardigheidsverschillen in leren leren tussen jongens en meisjes bij de schalen Oriënteren, voorbereiden en plannen en Monitoren, reflecteren en evalueren is niet eenvoudig te geven. Naar sekseverschillen op het gebied van leren leren in het speciaal onderwijs heeft voor zover ons bekend geen theorievorming of onderzoek plaatsgevonden. Wel is uit onderzoek in het regulier basis- en voortgezet onderwijs bekend dat sekseverschillen sterk afhankelijk zijn van het specifieke vakgebied en de specifieke opdracht (voor rekenen bijvoorbeeld groter dan voor sommige taalvaardigheden). Het ontbreken van een sekseverschil bij een volledig nieuw onderzoeksgebied als leren leren hoeft dan ook niet per se te wijzen op een gebrekkige validiteit. Daar komt bij dat er bij leerlingen in het speciaal basisonderwijs vaak sprake is van aanzienlijke beperkingen. Die wegen vaak zwaarder dan de vaardigheidsverschillen tussen jongens en meisjes zoals we deze bijvoorbeeld in het regulier basisonderwijs aantreffen.

4.7.4.2 Leeftijd en leerjaar

Leeftijd

Normaliter neemt de vaardigheid toe naarmate de leerling ouder wordt. Tabel 4.11 laat zien hoe de leerlingen in de responsgroep verdeeld zijn naar leeftijd. De leeftijden 10, 11 en 12 komen het vaakst voor namelijk 22%, 27% en 27%.

Tabel 4.11 Verdeling van de leerlingen naar leeftijd

Leeftijd	N	%	Valide %
8	27	4	4
9	82	13	13
10	137	22	22
11	166	26	27
12	170	27	27
13	38	6	6
Totaal	620	99	100
Missing	9	1	
Totaal	629	100	

Tabel 4.12 toont de gemiddelde vaardigheid in leren leren per leeftijdscategorie.

Tabel 4.12 Gemiddelde score per leeftijdscategorie

	Leeftijd						Totaal
	8	9	10	11	12	13	
Oriënteren, voorbereiden en plannen	18,80	20,04	20,74	20,95	21,78	21,45	20,96
Uitvoeren	20,23	21,37	22,32	21,97	23,11	23,71	22,30
Monitoren, reflecteren en evalueren	22,71	24,18	26,02	25,28	26,86	26,85	25,74
Aandachtig werken	19,64	22,34	22,67	22,63	22,79	22,63	22,52
Gemotiveerd werken	21,81	22,00	22,47	22,28	22,86	22,91	22,46
Werken zonder gedragsproblematiek	21,56	21,65	22,31	20,92	21,95	23,49	21,79
Zelfstandig werken en hulp vragen	20,62	21,23	22,59	22,21	22,98	23,66	22,38

De effectgroottes voor de verschillen tussen opeenvolgende leeftijdscategorieën zijn weergegeven in tabel 4.13. De effectgroottes geven geen duidelijke, gestage en gelijkmatige toename van de vaardigheid in leren leren met het stijgen van de leeftijd te zien. Volgens de verwachting is het algebraïsche teken van de effectgroottes overwegend negatief, maar de grootte van de effecten is doorgaans klein. Bovendien wordt de trend regelmatig verstoord door zeer lage of zelfs positieve waarden. Het patroon kan daarom als grillig gekenschetst worden. Dit is met name het geval bij de schaal Aandachtig werken en Werken zonder gedragsproblematiek. Voor een deel is dit toe te schrijven aan soms zeer kleine aantal waarnemingen per leeftijdscategorie waardoor toevallige steekproeffluctuaties het verwachte patroon kunnen verstoren. Verder valt op dat de verschillen tussen de leeftijdscategorieën 8 en 9 en tussen 9 en 10 groter lijken dan tussen met name 10 en 11 en tussen 12 en 13.

Tabel 4.13 Effectgroottes van de gemiddelde scoreverschillen tussen opeenvolgende leeftijdscategorieën

	8 min 9	9 min 10	10 min 11	11 min 12	12 min 13
Oriënteren, voorbereiden en plannen van praktische taken	-0,26	-0,13	-0,04	-0,15	0,06
Uitvoeren van praktische taken	-0,22	-0,18	0,06	-0,20	-0,12
Monitoren, reflecteren en evalueren van praktische taken	-0,21	-0,25	0,10	-0,22	0,00
Aandachtig werken	-0,45	-0,05	0,01	-0,03	0,03
Gemotiveerd werken	-0,04	-0,09	0,03	-0,10	-0,01
Werken zonder gedragsproblematiek	-0,02	-0,14	0,25	-0,18	-0,30
Zelfstandig werken en hulp vragen	-0,12	-0,27	0,07	-0,14	-0,14

Leerjaar

De verwachting is dat leerlingen in de hogere leerjaren gemiddeld vaardiger zijn dan die in de lagere leerjaren (als gevolg van onderwijs en van rijping met het stijgen van de leeftijd). Tabel 4.14 laat zien dat de leerlingen in de responsgroep relatief gelijkmatig over de leerjaren 5, 6, 7 en 8 verdeeld zijn.

Tabel 4.14 Verdeling van de leerlingen naar leerjaar

Leerjaar	N	%	Valide %
5	89	14	14
6	181	29	29
7	147	23	24
8	203	32	33
Totaal	620	99	100
Missing	9	1	
Totaal	629	100	

De effectgroottes voor de verschillen tussen opeenvolgende leerjaarcategorieën zijn weergegeven in tabel 4.15.

Tabel 4.15 Gemiddelden en standaarddeviaties per schaal en effectgroottes van de gemiddelde scoreverschillen tussen opeenvolgende leerjaarcategorieën

	5		6		7		8		Effectgrootte		
	GEM	SD	GEM	SD	GEM	SD	GEM	SD	5	6	7
										minus	minus
									6	7	8
Oriënteren, voorbereiden en plannen	18,45	5,22	20,80	5,50	21,18	5,57	21,79	5,46	-0,44	-0,07	-0,11
Uitvoeren	20,35	4,99	22,17	5,61	22,48	5,72	23,01	5,33	-0,34	-0,06	-0,10
Monitoren, reflecteren en evalueren	21,81	6,59	25,83	7,27	26,38	7,76	26,63	7,02	-0,58	-0,07	-0,03
Aandachtig werken	20,88	6,17	23,25	6,01	22,10	6,64	22,70	5,77	-0,39	0,18	-0,10
Gemotiveerd werken	21,18	5,01	22,88	5,35	21,94	5,55	22,85	5,60	-0,33	0,17	-0,16
Werken zonder gedragsproblematiek	21,48	5,07	21,96	5,41	21,44	5,76	21,99	5,27	-0,09	0,09	-0,10
Zelfstandig werken en hulp vragen	20,47	4,87	22,58	4,96	22,12	5,97	23,12	4,99	-0,43	0,09	-0,18

De effectgroottes geven geen duidelijke, gestage en gelijkmatige toename van de vaardigheid in leren leren met het stijgen van het leerjaar te zien. Behoudens één uitzondering wijzen de effectgroottes voor de contrasten tussen leerjaar 5 en 6 op kleine tot middelmatige verschillen tussen beide groepen in gemiddelde vaardigheid in leren leren. De uitzondering is de schaal Werken zonder gedragsproblematiek die geen enkele samenhang met leerjaar vertoont. De verschillen tussen leerjaar 5 en 6 zijn veel groter dan die tussen leerjaar 6 en 7 en tussen leerjaar 7 en 8. Vooral bij Aandachtig werken, Gemotiveerd werken en Zelfstandig werken is er geen sprake van een duidelijke, consistente en gelijkmatig stijgende trend van leerjaar 5 naar leerjaar 8.

Interpretatie

Men zou logischerwijs verwachten dat de vaardigheid in leren leren toeneemt met de leeftijd van de leerling en de groep waarin de leerling zich bevindt. Het nagenoeg ontbreken van het veronderstelde effect van leeftijd en groep op de vaardigheid in leren leren vraagt om een verklaring. Een voor de hand liggende verklaring is het relatief kleine aantal leerlingen in de onderscheiden groepen dat aan de proefafname heeft meegedaan. Een tweede mogelijke verklaring veronderstelt dat leeftijd en groep in sbo en so cluster 4 minder goede indicatoren zijn van effectieve leertijd en daarmee van het bereikte vaardigheidsniveau dan in het regulier basis- en voortgezet onderwijs. Een derde mogelijke verklaring verwijst naar de grote invloed van de handicap van de leerling op de vaardigheid in leren leren. Dat sterke effect zou met andere woorden het effect van leeftijd en groep kunnen overschaduwen.

4.7.4.3 Uitstroomprofiel

De verdeling van de leerlingen uit het zevende en achtste leerjaar naar uitstroomprofiel is weergegeven in tabel 4.16. De leerlingen zijn relatief gelijkmatig verdeeld over de vijf uitstroomprofielen, al is havo/vwo wat minder sterk vertegenwoordigd dan de overige uitstroomprofielen. De verwachting is dat leerlingen vaardiger zijn in leren leren naar mate het uitstroomprofiel 'hoger' is.

Tabel 4.16 Verdeling van de leerlingen naar uitstroomprofiel in leerjaar 7 en 8

	N	%	Valide %
Praktijkonderwijs	70	20	23
Vmbo bb	62	18	20
Vmbo kb	77	22	25
Vmbo gt	64	18	21
Havo/Vwo	34	10	11
Missing	43	12	
Totaal	350	100	100

De tabellen 4.17 en 4.18 geven inzicht in de richting en grootte van de vaardigheidsverschillen tussen de groepen leerlingen in verschillende uitstroomprofielen.

Tabel 4.17 Gemiddelden en standaarddeviaties per schaal per uitstroomprofiel

	Pro		Vmbo bb		Vmbo kb		Vmbo-gt		Havo//Vwo	
	Gem	SD	Gem	SD	Gem	SD	Gem	SD	Gem	SD
Oriënteren, voorbereiden en plannen van praktische taken	19,58	5,47	20,31	4,49	20,52	5,55	23,11	4,67	24,95	5,47
Uitvoeren van praktische taken	21,20	5,51	21,37	4,65	22,31	5,68	23,82	4,73	25,08	5,64
Monitoren, reflecteren en evalueren van praktische taken	24,27	6,69	24,49	6,44	25,39	7,04	28,07	6,44	29,78	7,82
Aandachtig werken	21,19	6,26	21,25	5,63	22,74	6,03	24,33	5,34	25,26	6,48
Gemotiveerd werken	21,36	5,88	21,35	5,21	22,37	5,52	24,23	4,61	24,85	5,20
Werken zonder gedragsproblematiek	21,59	5,71	21,66	5,35	22,54	5,40	22,34	4,41	21,68	5,23
Zelfstandig werken en hulp vragen	21,33	5,52	21,62	4,83	22,44	5,40	23,48	4,88	25,03	5,69

Tabel 4.18 Effectgroottes van de gemiddelde scoreverschillen tussen opeenvolgende categorieën van uitstroomprofiel

	Pro minus Vmbo bb	Vmbo bb minus kb	Vmbo kb minus gt	Vmbo gt minus Havo/vwo
Oriënteren, voorbereiden en plannen van praktische taken	-0,15	-0,04	-0,51	-0,36
Uitvoeren van praktische taken	-0,03	-0,18	-0,29	-0,24
Monitoren, reflecteren en evalueren van praktische taken	-0,03	-0,13	-0,40	-0,24
Aandachtig werken	-0,01	-0,26	-0,28	-0,16
Gemotiveerd werken	0,00	-0,19	-0,37	-0,13
Werken zonder gedragsproblematiek	-0,01	-0,16	0,04	0,14
Zelfstandig werken en hulp vragen	-0,05	-0,16	-0,20	-0,29

Het algebraïsche teken van de opeenvolgende effectgroottes in tabel 4.18 is overwegend negatief wat erop wijst dat de vaardigheid in leren leren toeneemt naarmate het 'niveau' van het uitstroomprofiel 'hoger' is.

De contrasten tussen de opeenvolgende uitstroomprofielen kunnen we als volgt beschrijven:

- De vaardigheidsverschillen tussen PrO en Vmbo-bb en tussen Vmbo bb en Vmbo kb zijn verwaarloosbaar klein (met uitzondering van Aandachtig werken waar er sprake is van een klein verschil in het voordeel van Vmbo-kb-leerlingen).
- Tussen vmbo kb en vmbo gt is het gemiddeld vaardigheidsverschil groter dan bij de drie overige contrasten. De effectgroottes variëren van een verwaarloosbaar verschil tot een middelgroot verschil in het voordeel van Vmbo-gt-leerlingen. Vmbo-gt-leerlingen scoren vooral relatief hoog bij Oriënteren, voorbereiden en plannen, Monitoren, reflecteren en evalueren en Gemotiveerd werken.
- Werken zonder gedragsproblematiek is niet of nauwelijks gerelateerd aan het uitstroomprofiel.

Resumerend lijkt er sprake van een zwakke en weinig consistente toename van de vaardigheid in leren leren naarmate het uitstroomprofiel een 'hoger' niveau impliceert. Vooralsnog is onduidelijk waarom het verwachte verschil zich vrijwel uitsluitend voordoet tussen vmbo gt en vmbo kb en veel minder tussen de overige contrasten (PrO versus Vmbo bb, Vmbo bb versus Vmbo kb en Vmbo gt versus Havo/vwo). De bestudering van de samenhang met uitstroomprofiel is hiermee niet in strijd met de van tevoren opgestelde verwachtingen, maar vormt zeker geen overtuigend bewijs voor de veronderstelde begripsvaliditeit.

4.7.4.4 Sbo versus so cluster 4

Van de 629 leerlingen zaten er 226 op een sbo-school en 403 een so cluster 4-school. Tabel 4.19 geeft inzage in de verschillen tussen beide schooltypen voor wat betreft de gemiddelde vaardigheid in leren leren.

Tabel 4.19 Verschillen tussen leerlingen van sbo- en so cluster 4-scholen

	sbo		so cluster 4		Effect-grootte	p	Eta ²
	Gem	SD	Gem	SD			
Oriënteren, voorbereiden en plannen van praktische taken	21,73	5,46	20,56	5,55	0,21	0,023	0,01
Uitvoeren van praktische taken	23,01	5,74	21,86	5,34	0,21	0,014	0,01
Monitoren, reflecteren en evalueren van praktische taken	27,03	7,40	25,13	7,27	0,26	0,006	0,01
Aandachtig werken	23,14	6,33	22,14	6,02	0,16	0,056	0,01
Gemotiveerd werken	23,36	5,55	21,94	5,34	0,26	0,002	0,02
Werken zonder gedragsproblematiek	23,17	5,91	21,12	4,97	0,38	< 0,001	0,03
Zelfstandig werken en hulp vragen	23,29	5,38	21,87	5,14	0,27	0,002	0,02

De effectgroottes van het vaardigheidsverschil tussen sbo- en so cluster 4-leerlingen zijn overwegend klein. Sbo-leerlingen zijn gemiddeld iets vaardiger dan so cluster 4-leerlingen in het cognitieve domein van leren leren; de effectgroottes wijzen echter in de richting van kleine verschillen. In het domein Werkhouding behalen sbo-leerlingen relatief gunstige scores voor Werken zonder gedragsproblematiek, Gemotiveerd werken en Zelfstandig werken en hulp vragen. Het grootste is het verschil in het voordeel van sbo-leerlingen bij Werken zonder gedragsproblematiek. Dit is in overeenstemming met de verwachtingen aangezien in so cluster 4 meer leerlingen met gedragsproblematiek zitten dan in sbo.

4.7.4.5 Doorverwijzing naar een praktische opdracht

De checklist heeft tot doel de minst vaardige leerlingen te selecteren die het meest voor afname van een (tijdrovende) praktische opdracht in aanmerking komen. De leerkrachten hebben aangegeven of zij bij de leerling waarvoor zij de checklist hebben ingevuld een praktische opdracht zouden afnemen.

Van de 629 beoordeelde leerlingen zou de leerkracht er 300 doorverwijzen naar de praktische opdracht (48%), 169 zou hij of zij niet doorverwijzen (27%) en van 160 leerlingen ontbrak dit gegeven (25%).

Tabel 4.20 De vaardigheid in leren leren van de groepen leerlingen bij wie de leerkracht al dan niet een praktische opdracht zou afnemen per schaal

	Ja		Nee		Effect-grootte	p	Eta ²
	Gem	SD	Gem	SD			
Oriënteren, voorbereiden en plannen	20,35	5,37	21,64	5,72	-0,23	0,020	0,01
Uitvoeren	21,78	5,47	22,65	5,69	-0,16	0,110	0,01
Monitoren, reflecteren en evalueren	24,60	7,30	26,41	7,24	-0,25	0,014	0,01
Aandachtig werken	22,05	6,14	23,38	6,29	-0,21	0,031	0,01
Gemotiveerd werken	21,98	5,42	23,26	5,66	-0,23	0,019	0,01
Werken zonder gedragsproblematiek	21,16	5,12	22,44	5,85	-0,23	0,018	0,01
Zelfstandig werken en hulp vragen	21,90	5,03	23,13	5,76	-0,23	0,019	0,01

De gemiddelde vaardigheid van de groep die de leerkracht zou doorverwijzen naar de praktische opdracht is iets lager dan die van groep leerlingen aan wie de leerkracht geen praktische opdracht zou voorleggen (zie tabel 4.20). Gezien de effectgroottes gaat het echter om een hooguit klein verschil. De richting van het verschil is zoals verwacht. Het verschil in vaardigheid is echter veel kleiner dan wij op voorhand voorzien hadden. De verwachting was immers dat leerkrachten de opdrachten alleen zouden voorleggen aan de kleine groep leerlingen die zeer lage prestaties voor leren leren zouden behalen en die dus het meest baat zouden hebben bij verdere ontwikkeling van hun vaardigheden middels deelname aan een praktische opdracht. Het onverwacht hoge percentage doorverwezen leerlingen van bijna de helft vormt een aannemelijke verklaring voor het nagenoeg ontbreken van een vaardigheidsverschil tussen de wel en niet doorverwezen leerlingen. Nader onderzoek in de 'echte' onderwijs- en toetspraktijk lijkt hier geboden.

5 Uitvoerbaarheid, betrouwbaarheid en validiteit van het beoordelingsinstrument

5.1 Doel, opzet en uitvoering van de tweede proefafname

Na de proefafname van de checklist in het vroege voorjaar van 2017 heeft een paar maanden later een tweede proefafname plaatsgevonden. Het onderwerp was de uitvoerbaarheid, betrouwbaarheid en validiteit van de praktische opdrachten en de bijbehorende beoordelingsformulieren. Aan deze proefafname deden 21 leerkrachten van 12 scholen mee. In totaal is bij 115 leerlingen een praktische opdracht afgenomen. Daarvan zaten 62 leerlingen op een sbo-school en 53 leerlingen op een so cluster 4-school. Vanwege budgettaire beperkingen zijn slechts enkele aspecten van betrouwbaarheid en validiteit onderzocht. Zo is geen onderzoek gedaan naar interbeoordelaarsbetrouwbaarheid, test-hertestbetrouwbaarheid, *Differential Item Functioning* (DIF) en invariantie van de interne structuur over relevante groepen (zoals jongens en meisjes). Omdat het pakket Leren leren vooralsnog enig in zijn soort is, is onderzoek naar de convergente en discriminante validiteit nog niet mogelijk.

5.1.1 Doel

In de tweede proefafname zijn gegevens verzameld ter verbetering van de praktische opdrachten, de beoordelingsformulieren en het gebruik ervan. De proefafname kende drie doelen:

- het uittesten van de uitvoerbaarheid van de aanwijzingen voor de afname, beoordeling en instructie tijdens de afname;
- het uittesten van de beoordelingsformulieren
- het onderzoeken van de betrouwbaarheid en de validiteit van de meting van leren leren met de praktische opdrachten.

Uittesten uitvoerbaarheid en aanwijzingen

In het pakket Leren leren is een nieuwe vorm van toetsing – *dynamic assessment* – toegepast (zie hoofdstuk 2). Het is de tweede keer dat Cito een pakket ontwikkelt waarin toetsen, onderwijzen en leren volledig geïntegreerd zijn. Het pakket Leren leren voor sbo- en so cluster 4-leerlingen bouwt voort op de ervaringen die zijn opgedaan in het kader van de ontwikkeling van het pakket Leren leren voor VSO-leerlingen (Kuhlemeier & Chermin, 2015). Omdat *dynamic assessment* voor sbo- en so cluster 4-leerkrachten een volledig nieuwe toetsvorm is, is het onderzoeken van de uitvoerbaarheid een eerste vereiste. Een eerste doel van de proefafname is het verzamelen van informatie over de uitvoerbaarheid van de aanwijzingen voor de afname, beoordeling en instructie tijdens de afname (met als doel de uitvoerbaarheid zo nodig te verbeteren). We willen immers een pakket ontwikkelen dat zo goed mogelijk aansluit bij de onderwijspraktijk.

De gebruikershandleiding bij het eerste pakket Leren leren voor VSO (Cito, 2015) bevatte uitgebreide suggesties voor de afname, beoordeling en instructie van vaardigheden van leren leren tijdens de afname. De conclusie was destijds dat de uitvoerbaarheid van de aanwijzingen weinig te wensen overliet.

In paragraaf 5.5.2 doen we verslag van een soortgelijk onderzoek naar de ervaringen en verbeteringsuggesties van de leerkrachten die deelnamen aan de proefafname van de praktische opdrachten en beoordelingsformulieren van het nieuwe pakket Leren leren sbo en so cluster 4. Daarbij laten we tevens zien hoe we de resultaten hebben benut ter verbetering van het nieuwe pakket Leren leren.

In de proefafname is ook informatie verzameld over het functioneren van de adaptieve toewijzing van varianten aan leerlingen. Van elke opdracht was er een makkelijke A-variant, een wat moeilijkere B-variant en een moeilijke C-variant. De leerkrachten is gevraagd de keuze van de variant af te stemmen op het vaardigheidsniveau van de leerling. De makkelijke A-variant was in principe bedoeld voor leerlingen in groep 5 of 6, de middelmatige moeilijke variant voor leerlingen in groep 6 of 7 en de moeilijkste C-variant

voor leerlingen in groep 7 of 8. Uiteraard konden de leerkrachten van dit advies afwijken als zij daar reden toe hadden. Een tweede doel van de proefafname was na te gaan in hoeverre de leerkrachten onze adviezen voor de adaptieve toewijzing van varianten aan leerlingen hebben opgevolgd. Van dit deel van het onderzoek is verslag gedaan in paragraaf 5.2 en 5.5.3.

Onderzoeken betrouwbaarheid en validiteit

Een derde doel van de proefafname is het verzamelen van informatie over de betrouwbaarheid en validiteit van de meting van leren leren. De betrouwbaarheid is vastgesteld aan de hand van de interne consistentie, de item-restcorrelaties en de standaardmeetfout. Een van de validiteitsvragen is in hoeverre het a priori onderscheid in cognitieve en non-cognitieve vaardigheden van leren leren is terug te vinden in de data. Van het onderzoek naar de betrouwbaarheid en validiteit is verslag gedaan in paragraaf 5.4. en 5.5.

5.1.2 Opzet en uitvoering

Aan de proefafname van de praktische opdrachten is deelgenomen door 21 leerkrachten van 12 scholen. In totaal is bij 115 leerlingen een praktische opdracht individueel afgenomen. Van de leerlingen was 78% jongen en 22% meisje. De verdeling over de leeftijden 8, 9, 10, 11, 12 en 13 jaar was respectievelijk 2%, 6%, 28%, 35%, 24% en 5%. Van de leerlingen zat 4% in groep 5, 35% in groep 6, 28% in groep 7 en 32% in groep 8. Tabel 5.1 toont de verdeling van de leerlingen naar groep en uitstroomprofiel. Hierbij tekenen we aan dat de leerkrachten de vraag naar het uitstroomprofiel alleen voor leerlingen uit groep 7 en 8 hoefden te beantwoorden.

Tabel 5.1 Verdeling van de leerlingen naar uitstroomprofiel in groep 5, 6, 7 en 8

Uitstroomprofiel	Groep					Totaal
	5	6	7	8	Missing	
PrO	0	4	7	1	3	15
vmbo-bb	1	3	5	7	5	21
vmbo-bb/kb	0	2	0	1	1	4
vmbo-kb	0	0	1	8	1	10
vmbo-gt	0	0	5	6	0	11
havo	0	0	2	8	0	10
havo/vwo	0	0	2	0	0	2
anders, namelijk	0	1	0	2	2	5
Missing	3	26	7	0	1	37
Totaal	4	36	29	33	13	115

5.1.3 Het onderzochte materiaal

De praktische opdrachten en varianten

In de proefafname zijn niet alle 10 praktische opdrachten (30 varianten) van het pakket Leren leren voor sbo en so cluster 4 onderzocht. Vanwege budgettaire beperkingen is volstaan met het uittesten van 2 karakteristieke opdrachten met in totaal 6 varianten die als exemplarisch worden beschouwd voor het pakket als geheel:

- de A-, B- en C-variant van de opdracht Meten;
- de A-, B- en C-variant van de opdracht PowerPoint.

De opdrachten Meten en PowerPoint zijn zo evenwichtig mogelijk aan de leerkrachten toegewezen, maar de leerkrachten bepaalden zelf of zij de leerling de A-, B- of C-variant van de toegewezen opdracht voorlegden. De leerkrachten namen de opdrachten individueel af.

De beoordelingsformulieren

Het gebruikte beoordelingsinstrument bevatte zestien aspecten die de leerkracht ter plekke moest observeren en beoordelen (zie tabel 5.2). Een deel van de aandachtspunten is geformuleerd als vragen die de leerkracht tijdens de leergesprekken moest stellen (zie ook paragraaf 3.3). Van de zestien beoordelingsaspecten hebben er drie betrekking op Oriënteren en voorbereiden, drie op Uitvoeren, vier op Reflecteren en evalueren, drie op de hoeveelheid benodigde hulp (als omgekeerde maat voor het leerpotentieel) en drie op werkhouding (aandachtig en gemotiveerd werken zonder gedragsproblematiek). Het beoordelingsinstrument is integraal opgenomen in bijlage 3.

De beoordelingsaspecten

Zoals weergegeven in tabel 5.2, zijn de zestien beoordelingsaspecten onderverdeeld in zelfregulerende vaardigheden van leren leren, benodigde hoeveelheid hulp (als omgekeerde maat voor het leerpotentieel) en werkhouding (aandachtig en gemotiveerd werken zonder gedragsproblematiek).

Tabel 5.2 Indeling van de zestien beoordelingsaspecten naar fase en type aspect

Vaardigheid en beoordelingsaspect	Type aspect
Oriënteren en voorbereiden	
Wat ga je doen?	Zelfregulering
Wat ga je eerst doen en wat daarna?	Zelfregulering
Heb je alle spullen klaargezet? en Wat moet je nog pakken?	Zelfregulering
Benodigde hulp tijdens oriënteren en voorbereiden	Leerpotentieel
Uitvoeren	
Werken volgens (eigen) plan	Zelfregulering
Zelfmonitoring	Zelfregulering
Controle eigen werk	Zelfregulering
Benodigde hulp tijdens de uitvoering	Leerpotentieel
Reflecteren en evalueren	
Lijkt wat je gemaakt hebt op het voorbeeld? Of: Past wat je hebt gedaan/gemaakt bij de opdracht?	Zelfregulering
Wat ging er goed (terugkijkvraag)?	Zelfregulering
Wat ging er niet zo goed of fout (terugkijkvraag)?	Zelfregulering
Wat kan er beter (vooruitkijkvraag)?	Zelfregulering
Benodigde hulp tijdens de reflectie en evaluatie	Leerpotentieel
Werkhouding	
Aandachtig werken	Aandacht
Gemotiveerd werken	Motivatie
Werken zonder gedragsproblematiek	Impulsbeheersing

Voor elk beoordelingsaspect was er een descriptieve schaal met vier vaardigheidsniveaus (gecodeerd als 4, 3, 2 en 1). De schaalpunten 4, 3, 2 en 1 waren ieder van een inhoudelijke niveauomschrijving voorzien.

5.1.4 Statistische analyse

De betrouwbaarheid en validiteit is geanalyseerd met behulp van betrouwbaarheidsanalyse (interne consistentie, item-restcorrelaties, standaardmeetfout), correlatieanalyse en exploratieve factoranalyse (principale componentenanalyse met varimax rotatie).

5.2 Uitvoerbaarheid

5.2.1 Ervaringen van gebruikers

In de proefafname is de deelnemende leerkrachten gevraagd naar de kwaliteit van de praktische opdrachten, de hanteerbaarheid van de beoordelingsformulieren en de uitvoerbaarheid van de procedures van afname, beoordeling en instructie. Hierna rapporteren we de belangrijkste resultaten en hoe we deze resultaten hebben benut ter verbetering van het pakket Leren leren.

Hoe functioneerden de praktische opdrachten?

- De individuele afname van een praktische opdracht nam gemiddeld 26 minuten in beslag.
- 61% van de leerkrachten gaf aan dat de leerlingen het leuk vonden om de opdrachten te maken; 39% gaf zelfs aan dat de leerlingen het zeer leuk vonden.
- De helft van de leerkrachten gaf aanwijzingen om de praktische opdrachten te verbeteren. Zo noemden zij onder andere dat het niveau iets hoger mag, er minder tekst gebruikt mag worden en dat een los antwoordblad (bij de praktische opdracht Meten) prettig zou zijn. Verder gaven leerkrachten aan dat ze de opdrachten bij voorkeur klassikaal wilden uitvoeren.

Wat heeft de afname van de praktische opdracht u opgeleverd?

- Leerkrachten noemden hier onder andere 'het is een bevestiging van hoe leerlingen opdrachten aanpakken en op welke punten er nog ontwikkeling moet zijn (leerstrategie)'. Verder gaven leerkrachten aan dat ze meer inzicht hebben gekregen in hoe oplossingsgericht leerlingen denken.
- De helft van de leerkrachten gaf aan dat het afnemen van een praktische opdracht nuttige informatie opleverde. Een aantal leerkrachten gaf daarbij wel aan dat het een te grote tijdsinvestering vraagt.

Hoe functioneerde de toewijzing van de praktische opdrachten aan leerlingen?

Van elke opdracht was er een makkelijke A-variant, een wat moeilijkere B-variant en een moeilijke C-variant. We vroegen de leerkrachten de keuze van de variant af te stemmen op het vaardigheidsniveau van de leerling. De makkelijke A-variant was in principe bedoeld voor leerlingen van groep 5/6, de wat moeilijkere B-variant voor leerlingen uit groep 6/7 en de moeilijke C-variant voor leerlingen uit groep 7/8. De leerkrachten bleken zich hier vrijwel zonder uitzondering aan gehouden te hebben. Eén leerkracht gaf aan dat ze ook naar kennis op het gebied van de inhoud van de praktische opdracht gekeken heeft. Haar argument was: "Als je vaardigheden van leren leren van een leerling wilt bekijken, heeft het niet zo veel zin om een opdracht aan te bieden die de leerling inhoudelijk al heel goed beheerst".

Hoe was het niveau van de opdrachten?

- Leerkrachten gaven aan dat zij de vaardigheden van hun leerlingen meestal als goed tot zeer goed beoordeelden. Veel leerlingen hadden maar weinig hulp nodig (denk daarbij aan hints, uitleg, instructie, voordoen, bemoedigen en corrigeren).
- We hebben leerkrachten gevraagd de moeilijkheidsgraad van de praktische opdrachten te beoordelen. Meten A, Powerpoint A en B worden een enkele keer als te makkelijk beschouwd. De meerderheid (83%) vond de opdrachten noch te moeilijk, noch te makkelijk.

Hoe organiseert u de praktische opdrachten?

We hebben leerkrachten gevraagd hoe zij de praktische opdrachten georganiseerd hebben. Genoemd werden: de opdrachten zijn afgenomen door de onderwijsassistent, de opdrachten zijn afgenomen tijdens de handvaardigheid- of gymles (les onder begeleiding van een vakleerkracht) en de opdrachten zijn afgenomen tijdens het zelfstandig werken. Ook werd genoemd dat de klas door een collega overgenomen werd.

Hoe functioneerden de afnamekaarten?

- 70% van de leerkrachten gaf aan dat de aanwijzingen op de afnamekaart duidelijk waren en 30% vond ze min of meer duidelijk.
- Als verbeteringsuggesties werden gegeven: minder herhaling, voorbeelden geven, afnamekaart per opdracht, niet per variant.

Hoe functioneerde de beoordeling?

- 83% van de leerkrachten vond de aanwijzingen voor het beoordelen duidelijk.
- 82% heeft de beoordeling naar eigen zeggen geheel volgens de instructies in de handleiding uitgevoerd.
- Suggesties ter verbetering van het beoordelingsformulier waren: fase 4 op 1 bladzijde en het aantal verschillende typen hints mag minder.

Hoe functioneerde het dynamisch toetsen?

- 64% van de leerkrachten staat positief tegenover het dynamisch toetsen en 36% staat hier neutraal tegenover. Ze ervoeren het als positief om individueel met leerlingen te werken. Ook gaven ze aan dat het erg leuk was, omdat leerlingen de afname niet ervaren als een toets. Je kijkt als leerkracht ook net even anders.
- 95% vond de suggesties voor het geven van hulp duidelijk.
- 73% van de leerkrachten gaven de hulp naar eigen zeggen geheel volgens de richtlijnen in de handleiding.

De feedback van de leerkrachten heeft onder meer aanleiding gegeven tot de volgende verbeteringen:

- De oorzaak voor de hoge scores in de proefafname heeft waarschijnlijk te maken met de instructies in de proefafnamehandleiding. In deze handleiding is de leerkrachten gevraagd de praktische opdrachten bij een diverse groep leerlingen af te nemen, dus niet alleen bij de leerlingen die dat het meest nodig hebben, maar ook bij degenen die al vaardig zijn op het gebied van leren leren. De verwachting is dat een groot deel van de leerkrachten een arbeidsintensieve praktische opdracht in de 'echte' onderwijspraktijk alleen aan de minst vaardige leerlingen zullen voorleggen. In de uiteindelijke gebruikershandleiding is de leerkrachten geadviseerd de praktische opdrachten vooral af te nemen bij de leerlingen die daar het meeste baat bij hebben. Welke leerlingen dat zijn, blijkt dan uit de checklist die leerkrachten voorafgaand aan de afname van de praktische opdrachten kunnen invullen.
- De tekst op de afnamekaarten hebben we ingekort waar mogelijk. Daarnaast hebben we wat aanpassingen in de formulering van de aanwijzingen gedaan om deze duidelijker te maken. De overige suggesties die leerkrachten gaven met betrekking tot de afnamekaarten waren niet concreet genoeg en/of niet haalbaar.
- Op de beoordelingsformulieren hebben we twee typen hints (suggestieve vraag; nee schudden) weggelaten, omdat deze niet of nauwelijks zijn gebruikt tijdens de proefafname. Verder willen leerkrachten graag dat het formulier van fase 4 op 1 bladzijde staat. Helaas past dit niet, dus hebben we niet aan deze wens kunnen voldoen.
- We hebben de constructiegroep om feedback gevraagd op de opbouw en inhoud van de VSO-handleiding. Deze feedback is meegenomen bij het opstellen van de handleiding voor sbo en so cluster 4. Aandachtspunten die uit de pilot naar voren kwamen hebben we waar mogelijk in de handleiding verwerkt (als aanvullende informatie of instructie voor de leerkracht). In de handleiding zijn bij wijze van voorbeeld en ter inspiratie concrete casussen opgenomen uit de mini-pilot van de constructiegroep. Verder heeft de constructiegroep handelingssuggesties per fase van de praktische opdrachten gegeven. Deze suggesties hebben we verwerkt in hoofdstuk 6 en aangevuld met suggesties vanuit onze eigen observaties (tijdens de (mini-)pilot). Ook op basis van de literatuur hebben we concrete suggesties geformuleerd. Tot slot hebben we een hoofdstuk met achtergrondinformatie over leren leren toegevoegd (hoofdstuk 7).

- Op de opdrachtkaarten van de praktische opdrachten hebben we waar mogelijk de teksten korter en/of concreter geformuleerd. Daarnaast hebben we in de handleiding suggesties opgenomen over de organisatie van de afname van de praktische opdrachten. Tevens hebben we de leerkracht geadviseerd om bij de keuze van een variant zo dicht mogelijk bij het niveau van de leerling aan te sluiten (zone van de naaste ontwikkeling) om te voorkomen dat de opdracht te makkelijk is en de leerling tijdens de afname helemaal niet tegen problemen aanloopt.

5.2.2 De toewijzing van varianten aan leerlingen

Er zijn twee opdrachten (Meten en PowerPoint) met ieder drie in moeilijkheidsgraad toenemende versies (A, B en C) aan de leerlingen voorgelegd. Elke leerling maakte slechts één van de zes versies. Tabel 5.3 laat zien hoe de varianten over de leerlingen verdeeld waren. Variant C van de opdracht PowerPoint is slechts aan zeven leerlingen voorgelegd. Een aannemelijke verklaring verwijst naar de hoge moeilijkheid van deze variant, wat ertoe heeft geleid dat leerkrachten deze variant alleen aan een kleine groep 'betere' leerlingen hebben voorgelegd.

Tabel 5.3 Verdeling van de leerlingen naar opdracht en variant

Opdracht	Variant				Totaal
	A	B	C	Missing	
Metten	16	19	24	1	60
PowerPoint	21	24	7	0	52
Missing	1	1	0	1	3
Totaal	38	44	31	2	115

In de proefafnamehandleiding zijn leerkrachten erop gewezen dat de makkelijke A-variant in principe bedoeld is voor leerlingen in groep 5 of 6, de middelmatige moeilijke variant voor leerlingen in groep 6 of 7 en de moeilijkste C-variant voor leerlingen in groep 7 of 8. Uiteraard konden de leerkrachten van deze globale richtlijn afwijken als daar zij daar reden toe hadden. Als de toewijzingsprocedure goed wordt uitgevoerd en iedere leerling de variant maakt die past bij zijn of haar vaardigheidsniveau, zouden de zes varianten gemiddeld ongeveer even goed gemaakt moeten zijn. Tabel 5.4 toont onder meer de gemiddelde oordelen per opdracht per variant (waarbij de gemiddelden voor de algehele vaardigheid in leren leren zijn berekend na vervanging van tot drie ontbrekende waarnemingen door waarden op basis van de wel beoordeelde aspecten). De tabel laat zien dat de gemiddelden voor de varianten A, B en C van Metten niet ver uit elkaar liggen. De verschillen in moeilijkheid bedragen hooguit 0,18 op een schaal met een theoretisch bereik van 1 tot en met 4. Van de opdracht PowerPoint lijkt de moeilijkste variant C naar verhouding slecht gemaakt, maar vanwege het zeer kleine aantal van slechts vijf afnames is bij de interpretatie extra voorzichtigheid geboden.

Tabel 5.4 Aantal leerlingen, gemiddelde oordelen en standaarddeviatie per opdracht per variant

Opdracht	Variant	N	Gem	Stddev
Meten	A	16	3,07	0,70
	B	19	3,14	0,47
	C	23	3,22	0,63
	Totaal	58	3,15	0,60
PowerPoint	A	21	2,95	0,67
	B	23	3,15	0,44
	C	6	2,78	0,71
	Totaal	50	3,02	0,58
Totaal	A	37	3,00	0,68
	B	42	3,14	0,45
	C	29	3,13	0,66
	Totaal	108	3,09	0,59

De mate waarin de leerkrachten het toewijzingsadvies hebben opgevolgd, kan met enige slagen om de arm ook worden afgeleid uit de samenstelling van de groepen leerlingen die de A-, B- en C-varianten gemaakt hebben. De veronderstelling is dat de samenstelling niet verschilt qua sekse, maar wel qua groep, leeftijd en uitstroomprofiel. De resultaten zijn als volgt:

- De verdeling van de groepen leerlingen die de A-, B- en C-varianten maakte, verschilt niet significant naar geslacht ($p = 0,159$).
- Variant A bevat relatief veel jongere leerlingen en variant C relatief veel oudere leerlingen ($p < 0,001$).
- Variant A is relatief vaak aan leerlingen uit de lagere groepen voorgelegd en variant C vaker aan leerlingen uit de hogere groepen ($p < 0,001$).
- Variant A is relatief vaak aan leerlingen met een 'lager' uitstroomprofiel voorgelegd en variant C vaker aan leerlingen met 'hoger' uitstroomprofiel ($p = 0,002$).

Al met al zijn er geen aanwijzingen dat leerkrachten de aanwijzingen voor de afname in de wind geslagen hebben. De resultaten zijn in overeenstemming met de verwachtingen en geven aanleiding tot vertrouwen in de adaptieve toewijzing van de varianten.

5.2.3 Hanteerbaarheid van het beoordelingsinstrument

De leerkrachten beoordeelden de vaardigheden van leren leren aan de hand van zestien beoordelingsaspecten. Een belangrijke indicatie voor de hanteerbaarheid van de beoordelingsformulieren is het percentage ontbrekende waarnemingen. Tabel 5.5 geeft de verdeling van het aantal niet-beoordeelde aspecten. Driekwart van de leerlingen is op alle zestien aspecten beoordeeld, bij 9% ontbreekt één oordeel, bij 13% ontbreken drie oordelen en bij 4% vier of meer oordelen. Zoals we later zullen zien, komen de missings vooral voor bij de schalen voor werkhouding.

Tabel 5.5 Verdeling van het aantal ontbrekende oordelen

Aantal ontbrekende waarnemingen	Aantal leerlingen	Percentage leerlingen
0	86	75
1	10	9
3	15	13
4	1	1
12	1	1
13	1	1
14	1	1
Totaal	115	100

5.3 Beschrijvende statistische gegevens

5.3.1 Beschrijvende gegevens per beoordelingsaspect

Enkele beschrijvende statistische gegevens van de zestien beoordelingsaspecten van het beoordelingsinstrument zijn weergegeven in tabel 5.6. Behalve de percentages per vaardigheidsniveau bevat de tabel het gemiddelde, de standaarddeviatie en de scheefheid van de verdelingen. De oordelen van de leerkrachten blijken zonder uitzondering scheef verdeeld te zijn met meer oordelen in het 'positieve' gedeelte van de schaal dan in het 'negatieve' gedeelte (uitgedrukt op een gemiddeldenschaal met een theoretisch bereik van 1 tot en met 4). De scores 3 en 4 voor een goede en uitstekende vaardigheid komen veel vaker voor dan de scores 1 en 2 voor een zwakke en redelijke vaardigheid.

Tabel 5.6 Enkele verdelingskenmerken van de zestien beoordelingsaspecten van het beoordelingsinstrument

	N	Min	Max	Gem	Stddev	Scheefheid	Percentuele verdeling over de vier niveaus			
							1	2	3	4
Oriënteren en voorbereiden										
Wat ga je doen?	112	1	4	3,20	0,84	-0,57	2%	21%	32%	45%
Wat ga je eerst doen en wat daarna?	112	1	4	3,12	0,90	-0,38	2%	29%	24%	45%
Heb je alle spullen klaargezet? en Wat moet je nog pakken?	113	1	4	3,39	0,77	-0,92	1%	15%	28%	56%
Hoeveel hulp heb je de leerling bij de oriënteren en voorbereiden in totaal moeten geven?	111	1	4	3,02	0,85	-0,57	5%	19%	44%	32%
Uitvoeren										
Werken volgens (eigen) plan	113	1	4	2,92	0,92	-0,38	7%	26%	37%	30%
Zelfmonitoring	112	1	4	2,97	0,85	-0,39	4%	24%	41%	30%
Controle eigen werk	113	1	4	2,88	0,93	-0,43	9%	23%	40%	28%
Hoeveel hulp heb je de leerling bij de uitvoering in totaal moeten geven?	108	1	4	2,73	0,93	-0,28	11%	27%	40%	22%
Reflecteren en evalueren										
Lijkt wat je gemaakt hebt op het voorbeeld. Of: Past wat je hebt gedaan gemaakt bij de opdracht product?	109	1	4	3,23	0,79	-0,66	2%	17%	39%	43%
Wat ging er goed?	112	1	4	3,22	0,80	-0,77	4%	13%	41%	42%
Wat ging er niet zo goed of fout?	112	1	4	3,04	1,01	-0,84	13%	11%	37%	40%
Wat kan er beter?	112	1	4	3,02	1,03	-0,69	12%	17%	29%	42%
Hoeveel hulp heb je de leerling bij de reflectie en evaluatie in totaal moeten geven?	112	1	4	2,83	0,82	-0,30	6%	26%	47%	21%
Werkhouding										
Aandachtig werken	97	2	4	3,31	0,73	-0,54	0%	18%	35%	47%
Gemotiveerd werken	97	2	4	3,38	0,68	-0,66	0%	11%	39%	49%
Werken zonder gedragsproblematiek	97	1	4	3,30	0,83	-0,72	1%	21%	26%	53%

Vergelijken we de gemiddelden per beoordelingsaspect, dan blijkt 'Welke spullen heb je klaargezet? en Wat moet je nog pakken?' tot de hoogste oordelen te leiden. Het gemiddelde oordeel voor dit aspect zit met 3,39 relatief dicht tegen de maximale score van 4 aan. Kennelijk hebben de meeste leerlingen met dit aspect van de fase oriënteren en voorbereiden weinig moeite. Daarna volgen de drie aspecten van werkhouding. Dit wijst erop dat de leerlingen de opdrachten over het algemeen aandachtig, gemotiveerd en zonder gedragsproblematiek uitgevoerd hebben. De meeste moeite hebben de leerlingen met de drie aspecten van de cognitieve vaardigheid Uitvoering: zelfmonitoring, werken volgens (eigen) plan en controle eigen werk.

Het rechter gedeelte van tabel 4.6 toont de verdeling van de oordelen over de vier antwoordmogelijkheden. Enkele leerkrachten hadden zeer sporadisch twee antwoordmogelijkheden tegelijkertijd aangekruist. In de

analyses is hiervoor het gemiddelde genomen over de beide aangekruiste scores. Bij het berekenen van de percentuele verdeling over de vier niveaus zijn deze gemiddelden naar beneden afgerond op een geheel getal. De globale betekenis van de scores voor de vaardigheden en de drie aspecten van werkhouding is 1 = zwakke vaardigheid; 2 = redelijke vaardigheid; 3 = goede vaardigheid; 4 = uitstekende vaardigheid. De scores voor de hoeveelheid benodigde hulp hebben de volgende betekenis: 1 = voortdurend hulp; 2 = regelmatig hulp; 3 = af en toe hulp; 4 = geen hulp. De tabel laat onder meer zien dat ongeveer een kwart tot ruim een derde van de leerlingen regelmatig of voortdurend hulp nodig had. Ruim een vijfde tot bijna een derde maakte de opdracht zonder daarbij door de leerkracht geholpen te worden.

5.3.2 Beschrijvende gegevens per schaal

Voor elk van de hoofdvaardigheden van leren leren is een gemiddelde score geconstrueerd door de scores over de desbetreffende beoordelingsaspecten te middelen. Op dezelfde manier is een schaal geconstrueerd voor de hoeveelheid benodigde hulp. Het gemiddelde over alle 16 aspecten is berekend op basis van de records waarvoor er scores voor alle 16 beoordelingsaspecten beschikbaar waren. Behalve deze totaalscore is een tweede gemiddelde berekend waarbij voor ontbrekende waarnemingen in records met tot 3 missings vervangende waarden zijn geïmputeerd op basis van de wel beoordeelde aspecten. Daarmee stijgt het aantal leerlingen waarvoor een totaalscore beschikbaar is van 86 tot 111 van de 115 leerlingen. Enkele verdelingskenmerken van de schalen van het pakket Leren leren zijn weergegeven in tabel 5.7. Te zien is onder meer dat het voor de verdeling van de totaalscores weinig uitmaakt of de berekening zonder of met vervanging van ontbrekende waarnemingen is uitgevoerd.

Tabel 5.7 Enkele verdelingskenmerken van de schalen van het definitieve instrument

	Aantal uitspraken	Aantal leerlingen	Min	Max	Gem	Stddev	Scheefheid
Oriënteren en voorbereiden	3	111	1,33	4,00	3,23	0,67	-0,59
Uitvoeren	3	112	1,00	4,00	2,93	0,76	-0,68
Reflecteren en evalueren	4	109	1,38	4,00	3,12	0,70	-0,71
Hoeveelheid benodigde hulp	3	106	1,33	4,00	2,87	0,71	-0,51
Werkhouding	3	97	1,83	4,00	3,33	0,64	-0,57
Totaalscore zonder imputaties	16	86	1,69	3,94	3,09	0,60	-0,69
Totaalscore met imputaties	16	111	1,69	3,94	3,09	0,59	-0,67

5.4 Betrouwbaarheid en meetnauwkeurigheid van het beoordelingsinstrument

De uitgave Leren leren is niet bedoeld om er belangrijke beslissingen mee te onderbouwen. Het doel reikt niet verder dan het in kaart brengen en volgen van de vaardigheid in leren leren. In dit hoofdstuk gaan we na in hoeverre de vastgestelde betrouwbaarheid voldoet aan de eisen die het COTAN Beoordelingssysteem hieraan stelt (Evers, Lucassen, Meijer & Sijtsma, 2010).

5.4.1 Interne consistentie en item-restcorrelaties

Mede op basis van de uitkomsten van de exploratieve factoranalyses zijn vijf schalen samengesteld door de toegekende scores over de beoordelingsaspecten te middelen. Drie schalen betreffen zelfregulerende vaardigheden van leren leren, één schaal betreft de hoeveelheid benodigde hulp en één schaal gaat over Werkhouding. Tabel 5.8 toont de interne consistentie en de item-restcorrelaties.

Tabel 5.8 Aantal aspecten, Cronbach's alfa en gemiddelde item-restcorrelatie (RIR) voor het beoordelingsinstrument

	Aantal aspecten	Alfa	Gemiddelde RIR (range)
Oriënteren en voorbereiden	3	0,73	0,55 (0,45 - 0,86)
Uitvoeren	3	0,79	0,63 (0,59 - 0,71)
Reflecteren en evalueren	4	0,75	0,56 (0,41 - 0,71)
Hoeveelheid benodigde hulp (leerpotentieel)	3	0,76	0,59 (0,55 - 0,61)
Werkhouding	3	0,82	0,67 (0,63 - 0,70)

Gelet op het zeer kleine aantal beoordelingsaspecten kan de betrouwbaarheid in termen van interne consistentie als zeer hoog worden aangemerkt ($0,73 < \text{alfa} < 0,82$). Volgens de COTAN-norm voor de betrouwbaarheid voor het nemen van minder zwaarwegende beslissingen op individueel niveau, zoals voortgangscntrole, hebben de schalen Oriënteren en voorbereiden, Reflecteren en evalueren, Uitvoeren en Hoeveelheid benodigde hulp een voldoende betrouwbaarheid en Werkhouding een goede betrouwbaarheid (Evers, Lucassen, Meijer & Sijtsma, 2010). De betrouwbaarheid van de totale schaal met alle zestien uitspraken bedraagt overigens 0,93.

5.4.2 Standaardmeetfout

Tabel 5.9 toont de standaardmeetfout met het bijbehorend 90%- en 95%-betrouwbaarheidsinterval. De gepresenteerde standaardmeetfouten zijn klein. Het bereik van plusminus (bijna) twee standaardmeetfouten rond de gemiddelde score geeft aan waar de ware score zich met 95%-zekerheid zal bevinden.

Tabel 5.9 Standaardmeetfout met bijbehorend 90%- en 95%-betrouwbaarheidsinterval

	Standaard-meetfout	Betrouwbaarheidsinterval	
		90%	95%
Oriënteren en voorbereiden	0,34	0,57	0,67
Uitvoeren	0,32	0,53	0,63
Reflecteren en evalueren	0,34	0,56	0,67
Hoeveelheid benodigde hulp (leerpotentieel)	0,35	0,57	0,68
Werkhouding	0,28	0,46	0,55

5.5 Validiteit van het beoordelingsinstrument

In dit hoofdstuk doen we verslag van de uitkomsten van de analyses die zijn uitgevoerd met het oog op de begripsvaliditeit van het beoordelingsinstrument. Het gebruiksdoel van het instrument reikt, net als bij andere instrumenten in het Cito Volgsysteem primair en speciaal onderwijs, niet verder dan het beschrijven en volgen van leerlingen, in dit geval van hun vaardigheid in leren leren. Omdat de uitgave Leren leren sbo en so cluster 4 niet bedoeld is voor 'voorspellend gebruik' blijven voorspellende kwaliteiten en criteriumvaliditeit in deze paragraaf verder onbesproken.

5.5.1 Factorzuiverheid

Hiervoor hebben we vastgesteld dat de betrouwbaarheid in termen van interne consistentie voldoende tot goed is. Naast homogeniteit is tijdens de constructie ook aandacht besteed aan het verkrijgen van factorzuivere schalen, dit wil zeggen schalen waarvan de inhoud zo weinig mogelijk overlapt met die van andere schalen. De meting van leren leren geldt als factorzuiver als de veronderstelde aspecten (schalen) in de gegevens terug te vinden zijn. In deze paragraaf onderzoeken we de factorzuiverheid ofwel de empirische onderscheidbaarheid van de geconstrueerde schalen. Dat doen we met behulp van exploratieve factoranalyse: principale componentenanalyse met varimax rotatie. We beschouwen een schaal als factorzuiver als de ladingen op de factor die de schaal pretendeert te meten hoog zijn (d.w.z. $\geq 0,40$) en de ladingen op de 'eigen' factor hoger zijn dan die op andere factoren die de schaal niet beoogd te meten. De resultaten van de exploratieve factoranalyse zijn weergegeven in tabel 5.10 (waarbij ladingen tussen $-0,40$ en $+0,40$ niet zijn weergegeven). Hierbij merken we op dat het aantal waarnemingen in verhouding tot het aantal uitspraken en het aantal te schatten parameters veel kleiner is dan de meeste vuistregels adviseren (Osborne & Costello, 2004). De resultaten gelden dan ook als zeer voorlopig en moeten met de grootst mogelijke voorzichtigheid geïnterpreteerd worden.

Tabel 5.10 Factorladingen van de beoordelingsaspecten van leren leren

	F1	F2	F3
Oriënteren en voorbereiden			
Wat ga je doen?		0,85	
Wat ga je eerst doen en wat daarna?		0,76	
Heb je alle spullen klaargezet? en Wat moet je nog pakken?		0,53	
Hoeveel hulp heb je de leerling bij de oriënteren en voorbereiden in totaal moeten geven?		0,78	
Uitvoeren			
Werken volgens (eigen) plan	0,76		
Zelfmonitoring	0,79		
Controle eigen werk	0,71		
Hoeveel hulp heb je de leerling bij de uitvoering in totaal moeten geven?	0,80		
Reflecteren en evalueren			
Lijkt wat je gemaakt hebt op het voorbeeld. Of: Past wat je hebt gedaan/gemaakt bij de opdracht?	0,53		
Wat ging er goed?			0,47
Wat ging er niet zo goed of fout?			0,85
Wat kan er beter?			0,87
Hoeveel hulp heb je de leerling bij de reflectie en evaluatie in totaal moeten geven?			0,75
Werkhouding			
Aandachtig werken	0,50	0,43	0,41
Gemotiveerd werken	0,58	0,43	
Werken zonder gedragsproblematiek			0,42

De exploratieve factoranalyse brengt 3 factoren aan het licht die tezamen 64% van de variantie verklaren (voor factor 1, 2 en 3 respectievelijk 49%, 8% en 7%). Het patroon van factorladingen kunnen we als volgt omschrijven:

- Zoals verwacht laadt een groot deel van de beoordelingsaspecten hoger op de 'eigen' factor dan op alle overige factoren.
- De eerste factor wordt gedomineerd door de 4 aspecten van de vaardigheid Uitvoeren.
- De tweede factor wordt grotendeels bepaald door de 4 aspecten van de vaardigheid Oriënteren en voorbereiden.

- Op de derde factor hebben 3 van de 5 aspecten van Reflecteren en evalueren hoge factorladingen (en 1 aspect een relatief lage lading).

Het patroon van factorladingen komt hiermee grotendeels overeen met de beoogde structuur. Het beeld wordt echter enigszins ontsierd door 2 afwijkingen van het ideaalbeeld (waarbij elk beoordelingsaspect uitsluitend hoog en op zijn 'eigen' factor zou moeten laden):

- Het aspect 'Lijkt wat je gemaakt hebt op het voorbeeld? Of: Past wat je hebt gedaan/gemaakt bij de opdracht?' laadt op de factor voor Uitvoeren in plaats van op de eigen factor voor Reflecteren en evalueren.
- De 3 non-cognitieve aspecten van Werkhouding vormen geen afzonderlijke factor. 2 van deze aspecten (Gemotiveerd werken en Werken zonder gedragsproblematiek) laden op de eerste en tweede factor en 1 aspect (Aandachtig werken) op alle 3 factoren van leren leren. In deze dataset vormen de aspecten van Werkhouding dus geen duidelijke vierde factor (naast de 3 cognitieve 'zelfregulerende' hoofdvaardigheden van leren leren).

De resultaten van de factoranalyse wijken enigszins af van de beoogde structuur. De afwijkingen zijn niet groot en geven daarmee weinig aanleiding tot twijfel aan de toewijzing van de beoordelingsaspecten aan de drie cognitieve vaardigheden voor Oriënteren en voorbereiden, Uitvoeren en Reflecteren en evalueren. Een duidelijk minpunt is wel dat de drie aspecten voor Werkhouding geen 'eigen' factor vormen. De gevonden structuur is minder duidelijk dan de structuur die werd gevonden in het onderzoek van Kuhlemeier en Chermin (2015). In een omvangrijkere responsgroep van VSO-leerlingen gaven de correlaties tussen de beoordelingsaspecten toen een duidelijke vier-factorstructuur te zien met goed onderscheidbare en interpreteerbare factoren voor Oriënteren en voorbereiden, Uitvoeren, Reflecteren en evalueren en Werkhouding. Mogelijk speelt het zeer kleine aantal waarnemingen ons hier parten.

Als extra check op het onderscheid in 3 (schalen voor) zelfregulerende vaardigheden van leren leren is een aanvullende exploratieve factoranalyse uitgevoerd met alleen de tien beoordelingsaspecten op het gebied van zelfregulering. Zoals blijkt uit tabel 5.11 kunnen de correlaties tussen de tien aspecten met drie factoren afdoende beschreven worden. De verklaarde variantie bedraagt 67% waarvan voor factor 1, 2 en 3 respectievelijk 48%, 10% en 9%. De gevonden drie-factorstructuur is grotendeels in overeenstemming met de a priori indeling overeenkomstig de drie fasen van het leerproces (Butler & Winne, 1995; Den Besten & Vulpen, 2013; Pintrich & Zusho, 2002; Timmerman, 2002; Van Bon, 2009; Zimmerman, 1990; 2002; Zumbun, Tadlock & Roberts, 2011). De belangrijkste afwijking betreft het eerste beoordelingsaspect van Reflecteren en evalueren dat relatief hoog op de factor voor Uitvoeren laadt in plaats van op zijn 'eigen' factor. Omdat de huidige analyse op zeer weinig leerlingen gebaseerd is, neigen wij ertoe om de kleine afwijkingen van het veronderstelde patroon voor lief te nemen. Bij de constructie van de schalen en de rapportage aan de gebruiker gaan we derhalve uit van de vier-factorstructuur zoals deze destijds gevonden is bij het pakket Leren leren voor VSO (Kuhlemeier & Chermin, 2015).

Tabel 5.11 Factorladingen van de beoordelingsaspecten met betrekking tot zelfregulerende vaardigheden van leren leren

	F1	F2	F3
Oriënteren en voorbereiden			
Wat ga je doen?			0,86
Wat ga je eerst doen en wat daarna?			0,75
Heb je alle spullen klaargezet? en Wat moet je nog pakken?			0,55
Hoeveel hulp heb je de leerling bij de oriënteren en voorbereiden in totaal moeten geven?			0,79
Uitvoeren			
Werken volgens (eigen) plan	0,76		
Zelfmonitoring	0,80		
Controle eigen werk	0,71		
Hoeveel hulp heb je de leerling bij de uitvoering in totaal moeten geven?	0,81		
Reflecteren en evalueren			
Lijkt wat je gemaakt hebt op het voorbeeld. Of: Past wat je hebt gedaan/gemaakt bij de opdracht?	0,53		
Wat ging er goed?		0,50	
Wat ging er niet zo goed of fout?		0,85	
Wat kan er beter?		0,86	
Hoeveel hulp heb je de leerling bij de reflectie en evaluatie in totaal moeten geven?	0,40	0,76	

5.5.2 Interne structuur

5.5.2.1 Correlaties tussen de schalen

De begripsvaliditeit van de meting van de vaardigheid van leren leren moet in eerste instantie blijken uit de interne structuur ofwel de dimensionaliteit van de meting. Als de veronderstelde structuur hout snijdt, zou die ook in de gegevens terug te vinden moeten zijn. In de toetsingsfase van het valideringsproces (Cronbach, 1970) zijn er tenminste drie situaties die aanleiding geven om de begripsvaliditeit van de meting van een aspect van de vaardigheid van leren leren in twijfel te trekken:

- Ten eerste kunnen de schalen of aspecten die deze vaardigheid pretenderen te meten perfect samenhangen. Het onderscheid tussen deze aspecten of schalen is dan empirisch betekenisloos. Dit kan aanleiding zijn om de theorie (en het instrument) te vereenvoudigen door het onderscheid tussen de desbetreffende indicatoren te laten vallen.
- Ten tweede kan er sprake zijn van een niet perfecte samenhang van de schalen en aspecten, terwijl de richting van de samenhang anders is dan op grond van theorie te verwachten valt. Ook dan is er aanleiding om de validiteit van de theorie en/of de meting in twijfel te trekken.
- Ten derde kunnen de schalen of aspecten geen enkele of slechts een geringe samenhang vertonen. Ze kunnen dan niet meer worden beschouwd als behorend tot hetzelfde overkoepelende begrip. Het ware dan beter om te spreken van afzonderlijke, op zichzelf staande begrippen.

Er is sprake van convergente validiteit als een schaal of aspect relatief hoog samenhangt met verwante schalen en aspecten en van discriminante validiteit als er sprake is van een ontbrekende of een relatief lage samenhang met niet verwante of minder verwante schalen of aspecten (Campbell & Fiske, 1959). In deze proefafname is door middel van correlatieanalyse een eerste indruk verkregen van de interne structuur van de meting van de vaardigheid leren leren. De correlaties tussen de schalen voor de drie zelfregulerende vaardigheden, de schaal voor de hoeveelheid benodigde hulp en de non-cognitieve schaal voor Werkhouding zijn weergegeven in tabel 5.12. De afhankelijke variabele is uitgedrukt op een schaal van 1 tot

en met 4 waarbij een hogere score staat voor een hogere vaardigheid, minder hulp en een betere werkhouding.

Tabel 5.12 Interne structuur (correlaties)

	Oriënteren en voor- bereiden	Uitvoeren	Reflecteren en evalueren	Hoeveelheid benodigde hulp	Werk- houding	Totaal	Totaal geïmputeerd
Oriënteren en voorbereiden	1,00						
Uitvoeren	0,53	1,00					
Reflecteren en evalueren	0,56	0,61	1,00				
Hoeveelheid benodigde hulp	0,72	0,74	0,76	1,00			
Werkhouding	0,64	0,67	0,62	0,72	1,00		
Totaal	0,84	0,84	0,84	0,92	0,86	1,00	
Totaal geïmputeerd	0,80	0,84	0,84	0,92	0,83	1,00	1,00

De drie zelfregulerende vaardigheden van leren leren vertonen een middelmatige samenhang ($0,53 < r < 0,61$). Hun samenhang is niet zo hoog dat er getwijfeld moet worden aan de empirische onderscheidbaarheid en niet zo laag dat men zich moet afvragen of deze drie vaardigheden wel onder het paraplubegrip leren leren geordend kunnen worden.

De drie zelfregulerende vaardigheden hangen sterk positief samen met de hoeveelheid benodigde hulp ($0,72 < r < 0,76$). Kennelijk hebben leerlingen met zwakkere zelfregulerende vaardigheden meer instructie en begeleiding nodig dan degenen met hogere zelfregulerende vaardigheden.

De werkhouding hangt eveneens sterk samen met de overige aspecten van leren leren ($0,62 < r < 0,72$): hoe gunstiger de werkhouding, hoe beter de zelfregulerende vaardigheden, maar ook hoe minder hulp de leerkracht de leerling hoefde te geven en hoe gunstiger het leerpotentieel.

Alle besproken relaties zijn in overeenstemming met de theoretische noties die aan de constructie ten grondslag liggen en vormen een ondersteuning voor de begripsvaliditeit van de meting van de vaardigheid in leren leren.

5.5.2.2 Samenhang van leren leren met het geven van hints

De meting van het aantal hints

De praktijkopdrachten zijn zo ontwikkeld dat leerlingen de ruimte krijgen om fouten te maken en daarvan te leren. Als een leerling tegen een probleem aanloopt en er op eigen kracht niet uitkomt, mag de leerkracht de leerling helpen. Dat kan bijvoorbeeld door hints, uitleg of instructie te geven of door dingen voor te doen. Het beoordelingsformulier (zie bijlage 3) bevatte behalve de zestien beoordelingsschalen ook een tabel waarin de leerkrachten tijdens de afname konden turven hoeveel hints zij gaven en ook om welk type hints het dan ging. Uit een aan de proefafname voorafgaande mini-pilot was namelijk gebleken dat leerkrachten hier behoefte aan hadden. Op het formulier waren de hints geordend volgens het principe van de geleidelijke toename van algemene naar specifieke hulp (o.a. Den Besten & Vulpen, 2013). Er is een onderscheid gemaakt in verbale en non-verbale hints. De verschillende typen hints zijn samen met de toelichting voor de leerkrachten weergegeven in het volgende overzicht.

Type hints	Toelichting
Verbale hints	
Algemene hint	Als de leerling is vastgelopen, is het niet de bedoeling dat je meteen zegt of voordoet hoe het moet. Een algemene hint kan al voldoende zijn om het denken van de leerling te stimuleren en het werk weer op gang te brengen. Voorbeelden zijn: 'Kijk nog eens naar de opdracht', 'Wat is er aan de hand?' of 'Doe je dit wel goed?'
Open vragen stellen	Wat ook vaak goed werkt, is het stellen van open vragen. Open vragen nodigen uit tot een gesprek en stimuleren de leerling om na te denken over de eigen aanpak zodat hij/zij deze gericht kan verbeteren. Je kunt deze vragen herkennen doordat ze vaak beginnen met hoe, wanneer, wie, waar, waarom, welke en wat. Voorbeelden zijn: 'Hoe deed je dat ook weer bij die vorige opdracht?', 'Wat zou je kunnen doen als dit gebeurt?', 'Wat heb je allemaal al gedaan?' of 'Hoe komt het dat het er anders uitziet dan het voorbeeld of de afbeelding?'
Gesloten vragen stellen	Als open vragen niet helpen, kun je een gesloten vraag proberen. Een gesloten vraag begint vaak met een werkwoord. Bijvoorbeeld: 'Heb je ...? Wil je ...? Kun je dit uitleggen? Kun je een andere manier bedenken om dit op te lossen? Vertel eens ...'
Suggestieve vraag stellen	Als ook een gesloten vraag geen soelaas biedt, kun je een suggestieve vraag stellen. Bijvoorbeeld: 'Zou het iets te maken kunnen hebben met ...?', 'Vond jij ... ook zo ...?' of 'Had je misschien ...?'
Oplossingen voorleggen	Als een suggestieve vraag niet helpt, kun je een leerling verschillende oplossingen voorleggen en hem/haar vragen welke oplossing in de gegeven situatie de beste is. Bijvoorbeeld: 'Wat denk jij, zou het kunnen liggen aan A, B of C?' Voor de leerling is een meerkeuzevraag vaak makkelijker dan een open vraag. De leerling hoeft de oplossing immers niet geheel zelf te bedenken, maar kan volstaan met het nadenken over welke oplossing in de gegeven situatie het meest voor de hand ligt.
Vertellen of voordoen	Pas in allerlaatste instantie, als de leerling er zelf echt niet uitkomt, kun je vertellen of voordoen hoe het moet. Het is immers de bedoeling dat de leerling de opdrachten steeds zelfstandiger leert uit te voeren en je als leerkracht steeds meer naar de achtergrond verdwijnt
Non-verbale hints	In plaats van of aanvullend op verbale hints kun je ook non-verbale hints inzetten. Denk daarbij aan hints als vragend of afwachtend kijken, wijzen of iets aanwijzen, ja knikken, nee schudden.

Het gedeelte van het beoordelingsformulier waarop de hints gescoord konden worden is hieronder onverkort weergegeven. De leerkrachten vulden de tabel indien zij dat wensten drie keer in: tijdens of na fase 1 en 2 (Oriënteren en voorbereiden), tijdens of na fase 3 (Uitvoeren) en tijdens of na fase 4 (Reflecteren en evalueren).

Ruimte om het aantal & type hints te turven (indien gewenst); (zie § 5.2 van de handleiding)					Ruimte voor opmerkingen (voor uzelf)	
Verbaal	Turven	Totaal	Non Verbaal	Turven	Totaal	
1) Algemene hint		<input type="text"/>	• Schouders ophalen		<input type="text"/>	
2) Open vraag		<input type="text"/>	• Vragend kijken		<input type="text"/>	
3) Gesloten vraag		<input type="text"/>	• (Aan)Wijzen		<input type="text"/>	
4) Suggestieve vraag		<input type="text"/>	• Ja knikken		<input type="text"/>	
5) Oplossingen voorleggen		<input type="text"/>	• Nee schudden		<input type="text"/>	
6) Vertellen of voordoen		<input type="text"/>	• ...		<input type="text"/>	

In het vervolg van deze paragraaf bespreken we eerst het aantal gegeven hints per type hint. Vervolgens gaan we na in hoeverre het aantal hints een betrouwbare schaal vormt. Tot slot brengen we het aantal gegeven hints in verband met scores voor de vaardigheid in leren leren. In de analyse is voor de vaardigheid in leren leren de tweede totaalscore gebruikt (waarbij voor ontbrekende waarnemingen in records met maximaal drie missings vervangende waarden zijn geïmputeerd op basis van de wel beoordeelde aspecten).

Verdeling van het aantal gegeven hints per type hint

Van de 112 leerlingen waarvoor bruikbare gegevens beschikbaar waren, is voor 70 leerlingen bijgehouden hoeveel en welk type hints zij ontvingen. Binnen de totale responsgroep van 112 leerlingen varieerde het totale aantal gegeven hints van 0 tot 110 met een gemiddelde van 12,61 en een standaarddeviatie van 21,20. Binnen de groep van 70 leerlingen die 1 of meer hints kreeg, varieerde het aantal hints van 1 tot 110 met een gemiddelde van 20,71 en een standaarddeviatie van 23,83. Tabel 5.13 toont het aantal hints per type hint voor de groep leerlingen die 1 of meer hints kreeg.

Tabel 5.13 Verdeling van het aantal hints per type voor de groep die hints ontvingen

	N	Min	Max	Gem	Stddev
Totaal aantal hints	70	1	110	20,17	23,83
Verbaal					
Algemene hints	70	0	16	3,10	3,52
Open vraag	70	0	13	3,36	2,99
Gesloten vraag	70	0	8	1,01	1,65
Suggestieve vraag	70	0	13	1,41	2,54
Oplossingen voorleggen	70	0	17	2,30	3,50
Vertellen of voordoen	70	0	13	2,14	3,38
Non-verbaal					
Schouders ophalen	70	0	19	0,96	2,68
Vragend of afwachtend kijken	70	0	20	1,80	3,44
Wijzen of iets aanwijzen	70	0	15	1,81	3,41
Ja-knikken	70	0	20	2,09	3,58
Nee-schudden	70	0	2	0,19	0,49

Betrouwbaarheid van de meting van het aantal hints

De mate waarin de leerkracht de leerling hints heeft gegeven blijkt betrouwbaar gemeten te worden. Cronbach's alfa over de elf typen hints bedraagt namelijk 0,92. Tabel 5.14 toont de item-rest-correlaties die behoudens twee uitzonderingen zeer hoog zijn. De (zeer) weinig gegeven hints 'Suggestieve vraag' en 'Nee schudden' blijken minder goed in de schaal te passen dan de overige hints en zijn dan ook niet in de definitieve versie van de beoordelingsformulieren opgenomen.

Tabel 5.14 Item-restcorrelaties per type hint

	Item-rest-correlatie
Algemene hints	0,86
Open vraag	0,70
Gesloten vraag	0,70
Suggestieve vraag	0,51
Oplossingen voorleggen	0,91
Vertellen of voordoen	0,87
Schouders ophalen	0,52
Vragend of afwachtend kijken	0,82
Wijzen of iets aanwijzen	0,70
Ja-knikken	0,76
Nee-schudden	0,46

Samenhang tussen de vaardigheid in leren leren en het aantal gegeven hints

Overeenkomstig het doel van dynamisch toetsen zouden leerlingen die zwak zijn in leren leren tijdens de afname meer hints moeten krijgen dan leerlingen die daar goed in zijn. Tabel 5.15 toont de correlaties tussen het totaal aantal gegeven hints en de vaardigheid in leren leren (voor de groep die één of meer hints kreeg). De correlaties zijn alle negatief waaruit we kunnen opmaken dat minder vaardige leerlingen tijdens de afname meer hints ontvingen dan vaardige leerlingen. De hoogte van de correlatie is sterk afhankelijk van het type hint. De hints Stellen van suggestieve vragen en Nee schudden blijken niet met de vaardigheid in leren leren samen te hangen en het Vertellen of voordoen en het Voorleggen van oplossingen het sterkst (r respectievelijk $-0,61$ en $-0,59$). Dit laatste wekt wellicht weinig verbazing als we bedenken dat leerkrachten deze twee typen hints alleen in allerlaatste instantie zullen toepassen, namelijk pas op het moment dat de leerling er zelf echt niet meer uitkomt en minder invasieve hints niet geholpen hebben.

Tabel 5.15 Samenhang tussen het aantal hints en de vaardigheid in leren leren

	r	p	N
Totaal aantal hints	-0,57	< 0,001	69
Verbaal			
Algemene hints	-0,44	< 0,001	69
Open vraag	-0,50	< 0,001	69
Gesloten vraag	-0,41	0,001	69
Suggestieve vraag	-0,16	0,180	69
Oplossingen voorleggen	-0,59	< 0,001	69
Vertellen of voordoen	-0,61	< 0,001	69
Non-verbaal			
Schouders ophalen	-0,39	0,001	69
Vragend of afwachtend kijken	-0,42	< 0,001	69
Wijzen of iets aanwijzen	-0,44	< 0,001	69
Ja-knikken	-0,36	0,003	69
Nee-schudden	-0,12	0,332	69

De hints 'Suggestieve vraag' en 'Nee schudden' dragen maar weinig bij aan de betrouwbaarheid en zijn nauwelijks gerelateerd aan de vaardigheid in leren leren. Vandaar dat deze twee typen hints niet in de definitieve versie van het beoordelingsinstrument zijn opgenomen.

De hoge negatieve correlaties tussen de vaardigheid in leren leren en de frequentie van de gegeven hints zijn in overeenstemming met het educatieve doel van dynamisch toetsen (d.w.z. het bieden van hulp aan degenen die daar het meest behoefte aan hebben) en vormen een ondersteuning voor wat wel de consequentiële validiteit van toetsing genoemd wordt (Messick, 1989). Het bewijs dat de hints daadwerkelijk effectief zijn en de vaardigheid in leren leren ten goede komen, is hiermee natuurlijk niet geleverd. Dat vereist een experimentele onderzoeksopzet met ten minste twee equivalente groepen leerlingen waarbij de ene groep de hints wel krijgt en de andere groep niet.

5.5.3 Verschillen tussen opdrachten en varianten qua moeilijkheidsgraad

Bij de constructie is ernaar gestreefd om overeenkomstige varianten van verschillende opdrachten ongeveer even moeilijk te maken (zie paragraaf 3.2). Zo is geprobeerd om de gemakkelijke A-variant van *Meten* ongeveer even moeilijk te maken als de gemakkelijke A-variant van *PowerPoint* en hetzelfde geldt ook voor de moeilijkere B- en C-varianten. Overeenkomstige varianten van verschillende opdrachten moeten met andere woorden 'nominaal parallel' zijn. De opdrachten en varianten verschillen onder meer qua vakgebied en onderwerp, maar moeten wel ongeveer even moeilijk zijn. Mede op grond van de adaptieve toewijzing van varianten aan leerlingen (zie paragraaf 3.3.2) verwachten we dat de gemiddelde oordelen over overeenkomstige varianten van verschillende opdrachten bij benadering gelijk zijn. We lichten deze verwachting als volgt toe. De leerkracht beoordeelt de vaardigheid van de leerlingen met generieke beoordelingschalen. Dat wil zeggen dat de beoordelingsaspecten en -criteria voor elke opdracht en voor elk van de varianten A, B en C gelijk zijn. Zoals uiteengezet in paragraaf 3.3, is de leerkracht gevraagd precies die variant te kiezen die naar zijn of haar inschatting het beste past bij wat de leerling aankan. De leerkracht geeft hiermee geen absoluut oordeel over de vaardigheid van de leerling, maar alleen een relatief oordeel gegeven de toewezen opdracht en variant. Als de opdrachten inderdaad ongeveer even moeilijk zijn en de adaptieve toewijzing doet wat het moet doen, zouden de gemiddelde oordelen voor de A-, B- en C-varianten ongeveer gelijk moeten zijn, ongeacht de opdracht. In het vervolg van deze paragraaf vragen we ons af in hoeverre soortgelijke varianten van verschillende opdrachten inderdaad even moeilijk zijn. Hierbij plaatsen we de kanttekening dat bij de interpretatie van de resultaten vanwege het kleine aantal waarnemingen de grootst mogelijke voorzichtigheid geboden is.

De opdrachten en varianten

In het onderzoek zijn niet alle tien opdrachten en 30 varianten van het pakket *Leren leren* onderzocht. Om financiële en logistieke redenen is volstaan met het uittesten van twee karakteristieke opdrachten met in totaal zes varianten die als exemplarisch worden beschouwd voor het pakket als geheel:

- de A-, B- en C-variant van de opdracht *Meten*;
- de A-, B- en C-variant van de opdracht *PowerPoint*.

Twee onderzoeksvragen

We maken een onderscheid in twee in elkaar geneste onderzoeksvragen. De eerste onderzoeksvraag is in hoeverre de gemiddelde oordelen over overeenkomstige varianten van *Meten* en *PowerPoint* vergelijkbaar zijn. De moeilijkheid van *Meten* variant A moet dan ongeveer gelijk zijn aan die van *PowerPoint* variant A, *Meten* B aan *PowerPoint* B en *Meten* C aan *PowerPoint* C, maar *Meten* A, B en C mogen qua moeilijkheid sterker van elkaar verschillen; hetzelfde geldt voor *PowerPoint* A, B en C.

De tweede onderzoeksvraag gaat nog iets verder en betreft de verwachting dat de gemiddelde oordelen voor alle zes varianten bij benadering gelijk zijn (dus ongeacht de opdracht en het uitstroomprofiel).

Statistische analyse

De twee onderzoeksvragen zijn beantwoord met behulp van een univariate variantieanalyse met als afhankelijke variabele de gemiddelde scores over alle zestien beoordelingsaspecten en als onafhankelijke variabelen de opdracht, variant en groep (waarbij een volledig model is geanalyseerd). De afhankelijke variabele is hiermee uitgedrukt op een schaal van 1 tot en met 4 waarbij een hogere score staat voor een hogere vaardigheid, minder hulp en een betere werkhouding. De analyse is uitgevoerd met de GENLIN-procedure uit SPSS.

De eerste onderzoeksvraag zou ondersteuning vinden als er geen twee- of drieweginteracties van variant met opdracht en/of groep gevonden zouden worden. De oordelen mogen dan verschillen van opdracht tot opdracht (bijvoorbeeld Meten gemakkelijker dan PowerPoint), van variant tot variant (bijvoorbeeld variant A gemiddeld gemakkelijker dan B dan C) en van groep tot groep (bijvoorbeeld lagere oordelen voor leerlingen in de lagere groepen dan voor leerlingen in de hogere groepen), maar variant A van Meten is even moeilijk als variant A van PowerPoint en hetzelfde geldt voor de overeenkomstige varianten B en C (dus ongeacht de opdracht en de groep).

De tweede onderzoeksvraag zou ondersteuning vinden als er geen enkel hoofd- of interactie-effect van opdracht, variant en groep significant zou zijn. De oordelen zijn dan voor alle zes varianten ongeveer even hoog (dus ongeacht de opdracht, de variant en de groep).

Resultaten

De resultaten van de univariate variantieanalyse zijn weergegeven in tabel 5.16.

Tabel 5.16 *Significantietoetsing van de hoofd- en interactie-effecten van de regressie van vaardigheid op opdracht, variant en groep*

Effect	Wald Chi-square	df	p
(Intercept)	1876,656	1	< 0,001
Opdracht	2,136	1	0,144
Variant	2,997	2	0,223
Groep	3,091	2	0,213
Opdracht * Variant	0,724	2	0,696
Opdracht * Groep	9,406	2	0,009
Variant * Groep	8,209	4	0,084
Opdracht * Variant * Groep	5,081	2	0,079

Van de zeven geschatte effecten geeft er één significantie te zien ($p < 0,05$). Het betreft de twee-weg-interactie tussen opdracht en groep ($p = 0,009$). Deze resultaten interpreteren we als volgt:

- De afwezigheid van een hoofdeffect van opdracht betekent dat de opdracht Meten gemiddeld genomen niet moeilijker of makkelijker is dan de opdracht PowerPoint (overeenkomstig de verwachtingen).
- De afwezigheid van het hoofdeffect van variant betekent dat de varianten A, B en C gemiddeld ongeveer even moeilijk zijn (overeenkomstig onze verwachtingen).
- De afwezigheid van een hoofdeffect van groep betekent dat de vaardigheid van leerlingen in de lagere groepen niet systematisch lager of hoger beoordeeld is dan van leerlingen in de hogere groepen (overeenkomstig de verwachtingen).
- De afwezigheid van een twee-weginteractie tussen opdracht en variant betekent dat het voor de moeilijkheid van de variant niet uitmaakt van welke opdracht de variant afkomstig is (overeenkomstig onze verwachtingen).
- De afwezigheid van een significante twee-weginteractie van variant en groep betekent dat het verschil tussen de gemiddelden van varianten van dezelfde opdracht onafhankelijk is van de groep waarin de leerling zit (overeenkomstig onze verwachtingen).
- De afwezigheid van een drie-weginteractie tussen opdracht, variant en groep betekent dat het voor de gemiddelde scores op Meten en PowerPoint niet uitmaakt voor welke combinatie van variant en groep de opdracht is afgenomen (overeenkomstig onze verwachtingen).
- De aanwezigheid van een twee-weginteractie tussen opdracht en groep betekent dat het voor het verschil tussen de scores op Meten en PowerPoint uitmaakt of de leerling uit een lagere of hogere groep afkomstig is (in strijd met onze verwachtingen).

De conclusie is dat de resultaten van de variantie-analyse vrijwel geheel in overeenstemming zijn met onze verwachtingen omtrent het functioneren van de toewijzings- en beoordelingssystematiek.

De geconstateerde interactie tussen opdracht en groep doet vermoeden dat het verschil tussen de scores op de beide opdrachten voor leerlingen uit de lagere groepen anders is dan voor leerlingen uit de hogere groepen. Een nadere inspectie van de paarsgewijze contrasten tussen de geschatte marginale gemiddelden per opdracht per groep brengt aan het licht dat er alleen in groep 7 sprake is van een significant verschil (van 0,50 scorepunt) tussen de oordelen voor Meten en PowerPoint ($p = 0,015$).

Een verklaring waarom dit verschil zich juist in groep 7 voordoet, is moeilijk te geven. Waarschijnlijk spelen steekproeffluctuaties ten gevolge van het zeer kleine aantal waarnemingen ($N = 97$) ons hier parten.

6 Samenvatting en discussie

6.1 Samenvatting

In deze wetenschappelijke verantwoording staat het pakket Leren leren centraal. Dit pakket is bedoeld voor leerlingen uit de groepen 5 tot en met 8 van sbo en so cluster 4. Het materiaal omvat een checklist en tien praktische opdrachten met bijbehorende beoordelingsformulieren voor de vakgebieden Rekenen, Tekenen, Computer-/schoolse vaardigheden, Zaakvakken en Techniek. Voor elke opdracht is er een gemakkelijke A-variant, een middelmatig moeilijke B-variant en een moeilijke C-variant beschikbaar. De leerkracht kan ermee vaststellen in hoeverre de leerling in staat is om cognitieve en non-cognitieve vaardigheden van leren leren toe te passen in nieuwe situaties. Bij de cognitieve vaardigheden gaat het om zelfregulerende vaardigheden zoals het kunnen nadenken over eigen kennis en vaardigheden en het kunnen toepassen ervan bij het uitvoeren van uiteenlopende taken. De non-cognitieve vaardigheden betreffen de werkhouding van de leerling waarbij een onderscheid is gemaakt in aandachtig werken, gemotiveerd werken, zelfstandig werken en werken zonder gedragsproblematiek. Het pakket stelt de leerkracht in staat de ontwikkeling van vaardigheden van leren leren vast te stellen, te volgen en verder te ontwikkelen. Daarnaast krijgt de leerkracht informatie over het leerpotentieel van de leerling en zijn of haar werkhouding. Het materiaal sluit nauw aan bij de leergebiedoverstijgende kerndoelen van het nationaal expertisecentrum leerplanontwikkeling (Ministerie van OCW, 1998; 2006; 2009) en de uitwerking daarvan in de vakoverstijgende leerlijnen voor leren leren van de CED-Groep (CED-groep, 2009).

De kern van het pakket Leren leren bestaat uit een checklist, een serie praktische opdrachten met bijbehorende beoordelingsformulieren en handreikingen voor de afname, beoordeling en instructie van de vaardigheid leren leren. In hoofdstuk 2 beschreven we de meetpretentie van het pakket Leren leren en lichtten we de doelgroep, het gebruiksdoel en de functie van het instrumentarium toe. Daarbij is ruim aandacht geschonken aan de theoretische en onderwijskundige argumenten op basis waarvan de cognitieve en non-cognitieve vaardigheden van leren leren gekozen zijn. De keuze van relevante vaardigheden is gebaseerd op recente literatuur over zelfregulerend leren als effectief middel om leerprestaties duurzaam te verhogen (o.a. De Boer, Donker-Bergstra & Kostons, 2012). Daarnaast is de keuze gebaseerd op de bewezen effectieve BOL/EIM-methodiek voor het onderwijzen van vaardigheden van leren leren in het speciaal onderwijs (Timmer, 1994; Timmer, Dekker & Voortman, 2003), de zogenoemde 'Beertjes-methodiek' (Meichenbaum, 1981; Timmerman, 2002; Van Bon, 2009) en recente literatuur over zogeheten executieve functies (Cooper-Kahn & Dietzel, 2011; Dawson & Guare, 2013; Smidts & Huizinga, 2011).

Het materiaal is ontwikkeld volgens de principes van *dynamic assessment*, een vorm van toetsend leren waarbij toetsen, onderwijzen en leren volledig geïntegreerd zijn (Tzurriel, 2001). *Dynamic assessment* is een vrij onbekende toetsvorm die Cito pas één keer eerder heeft toegepast (Kuhlemeier & Chermin, 2015). Vandaar dat we uitgebreid op *dynamic assessment* ingingen, waarbij we ook aandacht besteedden aan de vraag waarom juist deze toetsvorm bij uitstek geschikt is voor leerlingen in het speciaal onderwijs.

In hoofdstuk 3 stond de constructie en beschrijving van het pakket Leren leren centraal. Eerst bespraken we de algemene overwegingen die ten grondslag lagen aan de constructie. Daarbij verantwoordden we onder meer de keuze voor praktische in plaats van schriftelijke opdrachten en voor individuele in plaats van klassikale afname. Vervolgens gingen we in op de manier waarop we het materiaal overeenkomstig de uitgangspunten ontwikkeld hebben. Tot slot presenteerden we het uiteindelijk instrumentarium, inclusief een aantal beschrijvende statistische gegevens. Behalve aan de inhoud en opbouw van het pakket besteedden we ook aandacht aan de manier waarop de integratie van toetsen, onderwijzen en leren in de praktijk gestalte kreeg en hoe de testresultaten gerapporteerd worden. Uitgebreide informatie over het pakket en het gebruik ervan is te vinden in de gebruikershandleiding (Cito, 2018).

In hoofdstuk 4 deden we verslag van de constructie van de checklist. Deze vragenlijst heeft tot doel snel de leerlingen te selecteren die het meest voor afname van een arbeidsintensieve praktische opdracht in aanmerking komen. Daarnaast is de checklist een hulpmiddel om te bepalen welke vaardigheden extra (klassikale) aandacht vereisen en bij welke vaardigheden dat niet of minder nodig is. De checklist is samengesteld op basis van een proefafname waaraan 68 leerkrachten van 17 scholen de vragen beantwoordden voor in totaal 629 leerlingen uit 53 groepen. In de proefafname zijn gegevens verzameld over de uitvoerbaarheid, betrouwbaarheid en validiteit van de checklist. De ervaringen van de proefafnamedocenten waren grotendeels positief. De meeste leerkrachten konden goed uit de voeten met de checklist en vonden de checklist bruikbaar voor het selecteren van leerlingen die het meest in aanmerking komen voor afname van een praktische opdracht. Op basis van de feedback van de leerkrachten hebben we de checklist ingekort van 87 naar 58 vragen. Daarnaast hebben we in de handleiding een extra aanwijzing opgenomen over het doel en het invullen van de kwalitatieve beoordelingsschaal.

De definitieve versie van de checklist bevat 58 uitspraken die verdeeld zijn over 7 schalen. 3 schalen betreffen cognitieve vaardigheden van leren leren: Oriënteren, voorbereiden en plannen van praktische taken, Uitvoeren van praktische taken en Monitoren, reflecteren en evalueren van praktische taken. De overige schalen zijn non-cognitief van aard en hebben betrekking op de werkhouding tijdens het maken van praktische opdrachten: Aandachtig werken, Gemotiveerd werken, Werken zonder gedragsproblematiek en Zelfstandig werken en hulp vragen.

De betrouwbaarheid waarmee de 7 aspecten van leren leren gemeten werden, was goed.

Een factoranalyse op de uitspraken per schaal liet zien dat 5 van de 7 schalen uni-dimensioneel waren; de overige 2 schalen deden een beroep op 2 factoren.

De factoranalyses naar de empirische onderscheidbaarheid ofwel de factorzuiverheid bracht aan het licht dat de 3 cognitieve vaardigheden voldoende factor-zuiver gemeten werden. De onderscheidbaarheid van de 4 non-cognitieve vaardigheden (werkhouding) bleek enigszins te wensen over te laten. Dit gold in het bijzonder voor de schaal Zelfstandig werken en hulp vragen waarvan de uitspraken relatief hoge ladingen hadden op de factoren voor de overige drie aspecten van werkhouding (d.w.z. aandachtig werken, gemotiveerd werken en werken zonder gedragsproblematiek).

Uit de analyse van de interne structuur kwam naar voren dat de correlaties tussen sommige schalen weliswaar hoog waren, maar niet zo hoog dat aan de onderscheidbaarheid getwijfeld moest worden. De verwachte samenhang van leren leren met achtergrondkenmerken van leerlingen bleek vaak niet aanwezig of lastig te interpreteren. Een van de verwachtingen die niet uitkwamen, was dat oudere leerlingen in de hogere leerjaren gemiddeld duidelijk vaardiger zouden zijn dan jongere leerlingen in de lagere leerjaren (als gevolg van onderwijs en van rijping met het stijgen van de leeftijd). Bij een volledig nieuw onderzoeksgebied als leren leren in het speciaal onderwijs hoeft dat niet per se te wijzen op een gebrekkige validiteit. In het speciaal onderwijs hebben leerlingen vaak te kampen met een of meer beperkingen. Een van de vele mogelijke verklaringen verwijst naar de grote invloed van de handicap van de leerling op de vaardigheid in leren leren. Dat sterke effect zou met andere woorden het effect van kenmerken zoals geslacht, leeftijd en leerjaar kunnen overschaduwen. Een andere verklaring verwijst naar het subjectieve karakter van de beoordeling van de vaardigheid in leren leren. Van oudere leerlingen zullen leerkrachten hogere verwachtingen hebben dan van jongere leerlingen. Bij het invullen van de checklist zijn zij bij een jongere leerling vaak milder bij het beoordelen van hun vaardigheid. Bij oudere leerlingen zullen zij kritischer zijn. Dit leidt ertoe dat de vaardigheid van jongere leerlingen overschat worden en de vaardigheid van oudere leerlingen onderschat. Deze onbedoelde vertekening zou ervoor kunnen zorgen dat de gemiddelden van oudere en jongere leerlingen als het ware “naar elkaar toe kruipen” waardoor deze groepen op basis van hun gemiddelde scores nauwelijks meer van elkaar te onderscheiden zijn.

In hoofdstuk 5 rapporteerden wij over de constructie van de praktische opdrachten met bijbehorende beoordelingsformulieren. De ontwikkeling van een nieuwe toetsvorm voor een moeilijk meetbare vaardigheid als leren leren betekent dat de implementatie in de onderwijspraktijk extra aandacht vraagt.

In een tweede proefafname hebben 21 leerkrachten van 12 scholen bij in totaal 115 leerlingen een *dynamic assessment* individueel afgenomen. Dit gebruikersonderzoek heeft meer inzicht gegeven in de uitvoerbaarheid van de aanwijzingen voor de adaptieve afname, de beoordeling en de handreikingen voor de instructie. De duidelijkheid van de aanwijzingen voor de afname, beoordeling en instructie bleek weinig te wensen over te laten. De tekst op de afnamekaarten kon nog wat korter volgens de leerkrachten. In de definitieve versie hebben de tekst waar mogelijk wat ingekort en hier en daar wat aanpassingen in de formulering gedaan om de aanwijzingen nog duidelijker te maken. Daarnaast hebben we aandachtspunten die uit de pilot naar voren kwamen waar mogelijk in de definitieve handleiding verwerkt (als aanvullende informatie of instructie voor de leerkracht). Op de opdrachtenkaarten van de praktische opdrachten hebben we waar mogelijk de teksten korter/concreter geformuleerd. Leerkrachten hebben tijdens de pilot de adviezen voor de adaptieve toewijzing van varianten aan leerlingen goed opgevolgd. Bovendien blijken zij de beoordelingsformulieren zeer volledig en nauwkeurig te hebben ingevuld.

De tweede proefafname is gecombineerd met een onderzoek naar de betrouwbaarheid en de validiteit van de meting van vaardigheden van leren leren. In hoofdstuk 5 rapporteerden wij ook over de betrouwbaarheid en validiteit van het beoordelingsinstrument. De definitieve versie bevatte zestien beoordelingsaspecten die tezamen vijf aspecten van leren leren meten: a) zelfregulatie tijdens de fase van oriënteren en voorbereiden, b) zelfregulatie tijdens de uitvoering, c) zelfregulatie tijdens de fase van reflectie en evaluatie, d) het leerpotentieel geoperationaliseerd als de benodigde hulp tijdens de fasen van oriëntatie/voorbereiding, uitvoering en reflectie/evaluatie en e) de werkhouding met de aspecten aandachtig werken, gemotiveerd werken en werken zonder gedragsproblematiek.

De betrouwbaarheid van de vijf schalen in termen van interne consistentie bleek voldoende tot goed te zijn. Andere aspecten van de betrouwbaarheid, zoals de interbeoordelaarsbetrouwbaarheid, de test-hertestbetrouwbaarheid en de stabiliteit, konden vanwege budgettaire beperkingen niet onderzocht worden. Het onderzoek naar de empirische onderscheidbaarheid ofwel de factorzuiverheid bracht aan het licht dat het onderscheid in drie cognitieve vaardigheden van leren leren – Oriënteren en voorbereiden, Uitvoeren en Reflecteren en evalueren – ondersteuning vindt in de data. De factoranalyse liet echter ook zien dat de meting van deze drie aspecten niet geheel onafhankelijk is van de meting van het leerpotentieel en de werkhouding. Een aannemelijke verklaring is dat leerlingen die zwak zijn in zelfregulering en een minder goede werkhouding bezitten meer hulp en begeleiding nodig hebben dan degenen die daar beter in zijn. In de analyse van de interne structuur gaven de drie cognitieve vaardigheden van leren leren een relatief hoge samenhang te zien. De samenhang was echter niet zo hoog dat er getwijfeld moet worden aan de empirische onderscheidbaarheid.

Vanwege de beperkte omvang van de responsgroep aan wie de praktische opdrachten zijn voorgelegd kon er geen onderzoek worden gedaan naar andere aspecten van de validiteit, zoals *Differential Item Functioning* (DIF) en invariantie van de interne structuur over relevante groepen (zoals jongens en meisjes). Evenmin is onderzoek gedaan naar de veronderstelde positieve invloed van de *dynamic assessment* op de ontwikkeling van de vaardigheid leren leren. Omdat het pakket Leren leren vooralsnog enig in zijn soort is, was onderzoek naar de convergente en discriminante validiteit nog niet mogelijk.

6.2 Discussie

Vervolgonderzoek

De resultaten van de tweede proefafname zijn slechts gebaseerd op 115 *dynamic assessments*. Ook zijn niet meer dan 2 opdrachten met elk 3 varianten in het onderzoek betrokken. De resultaten van het onderzoek naar de uitvoerbaarheid, betrouwbaarheid en validiteit moeten derhalve als zeer voorlopig worden beschouwd en met het grootst mogelijke voorbehoud geïnterpreteerd worden. Het lijkt raadzaam de betrouwbaarheid en validiteit met meer opdrachten en leerlingen te onderzoeken en de uitkomsten ervan in de wetenschappelijke verantwoording en de gebruikershandleiding te rapporteren.

De hoge negatieve correlaties tussen de vaardigheid in leren leren en de frequentie van de gegeven hints zijn in overeenstemming met het educatieve doel van dynamisch toetsen (d.w.z. het bieden van hulp aan

degenen die daar het meest behoefte aan hebben) en vormen een ondersteuning voor wat wel de consequentiële validiteit van toetsing genoemd wordt (Messick, 1989). Het bewijs dat de hints daadwerkelijk effectief zijn en de vaardigheid in leren leren ten goede komen, is hiermee natuurlijk niet geleverd. Dat vereist een experimentele onderzoeksopzet met ten minste twee equivalente groepen leerlingen waarbij de ene groep de hints wel krijgt en de andere groep niet. Het verdient aanbeveling de effectiviteit van het pakket voor het ontwikkelen van de vaardigheid in leren leren nader te onderzoeken.

Betrouwbaarheid

De checklist en de beoordelingsformulieren voor het vaststellen en volgen van de vaardigheid leren leren zijn gebaseerd op observatie; het zijn geen objectief scorebare (leervorderingen)toetsen. Het is niet onbelangrijk te weten in hoeverre subjectiviteit in de observatie en beoordeling van leerlingen een rol speelt. De vraag is dan in hoeverre beoordelaars onafhankelijk van elkaar tot vergelijkbare oordelen komen. Leerkrachten met een duobaan zijn beiden goed bekend met de leerlingen in hun groep en daardoor bij uitstek geschikt als informant in onderzoek naar de interbeoordelaarsbetrouwbaarheid. Het verdient aanbeveling hier nader onderzoek naar te doen.

Het uitgevoerde betrouwbaarheidsonderzoek betrof alleen de interne consistentie van de meting. De test-hertestbetrouwbaarheid en de stabiliteit zijn nog niet onderzocht. Onderzoek naar de test-hertestbetrouwbaarheid over een periode van enkele weken geeft een indicatie van de meetfouten, terwijl stabiliteitsonderzoek met een interval van bijvoorbeeld zes maanden inzicht geeft in veranderingen van de ware scores. Dit wil zeggen: de 'echte' veranderingen van de vaardigheid leren leren, bijvoorbeeld ten gevolge van de instructie in leren leren (Coleman, 1968; Heise, 1969). Het lijkt raadzaam de test-hertestbetrouwbaarheid in termen van meetfouten en stabiliteit nader te onderzoeken.

Validiteit

Overeenkomstig het principe van *dynamic assessment* heeft het pakket tot doel een bijdrage te leveren aan het ontwikkelen van vaardigheden in leren leren. Een belangrijk validiteitscriterium is dan in hoeverre de leerling van de afname en de interventie geleerd heeft. Als het pakket daadwerkelijk doet wat het beoogt te doen, zouden de vaardigheden van de leerling een stijgende lijn te zien moeten geven (als gevolg van de deelname aan afnames, leergesprekken, hints et cetera). Teneinde een uitspraak te kunnen doen over de effectiviteit van het pakket voor het verbeteren van leren leren lijkt onderzoek geboden waarbij dezelfde of parallelle opdrachten minimaal twee keer worden afgenomen door een groep leerkrachten en leerlingen die wel met het pakket werkt en een equivalente groep die dat niet doet.

De toewijzings- en beoordelingssystematiek laat het toe om de vaardigheid van individuele leerlingen binnen een gegeven groep vast te stellen (en in de tijd te volgen). Het is echter niet mogelijk om de vaardigheid van leerlingen uit verschillende groepen met elkaar te vergelijken. In de gebruikershandleiding (Cito, 2018) is de leerkracht gewezen op de beperkte gebruikspretenties van het pakket Leren leren als meetinstrument.

Literatuur

- Aalsvoort, G.M. van der, Resing, W.C.M., & Ruijsenaars, J.J.M. (Eds.), *Advances in cognition and educational practice: Vol. 7. Learning potential assessment and cognitive training: Actual research and perspectives on theory building and methodology*. New York: Elsevier.
- Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Thousand Oaks, CA: Corwin Press.
- Aukes, N., Dekker, K., Jong, Y. de, Mulder, A., & Timmer, J. (2003). *Docentenhandleiding EIM voor de Educatie van het beroepsonderwijs*. Groningen: Stichting Beheer Collectieve Middelen NOSW.
- Besten, C. den, & Vulpen, A. (2013). *Begeleid ontdekkend leren: Het Eigen Initiatief Model in zes stappen*. Amsterdam: Boom Nelissen.
- Black, P., & Wiliam, D. (1998a). *Inside the black box: raising standards through classroom assessment*. London: School of Education, King's College.
- Black, P., & Wiliam, D. (1998b). Assessment and classroom learning. *Assessment in Education*, 5, 1, 7–75.
- Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European Psychologist*, 1, 100-112.
- Boekaerts, M. (1997). Zelf-regulerend leren bevorderen met behulp van het zes-blokkenmodel. In: J.G.L.C. Lodewijks, P.R.J. Simons & J.G.G. Zuylen (ReMeid.), *Het reguleren van leren* (pp. 32-57). Tilburg: MesoConsult B.V.
- Boekaerts, M. & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review*, 54, 2, 199-231.
- Boekaerts, M. & Simons, R-J. (1995) *Leren en instructie. Psychologie van de leerling en het leerproces*. Assen: Van Gorcum.
- Boer, H. de, Donker-Bergstra, A.S., & Kostons, D.D.N.M. (2012). *Effective strategies for self-regulated learning: A meta-analysis*. Groningen: GION.
- Bon, N. van (2009). "Ssst, ik kan zo niet werken!". *Concentratieproblemen bij kinderen met ADHD. (Afstudeeronderzoek Pabo)*. Arnhem: Hogeschool van Arnhem en Nijmegen.
- Boxtel, H. van, Hollenberg, J., Horsels, L., & Kleintjes, F. (2012). *Wetenschappelijke verantwoording van de toetsen Taal voor zml-leerlingen*. Arnhem: Cito.
- Budoff, M., Meskin, J., & Harrison, R. H. (1971). Educational test of the learning-potential hypothesis. *American Journal of Mental Deficiency*, 76, 159–169.
- Butler, D. (1996). The strategic content learning approach to promoting self-regulated learning: An introduction to the coordinated symposium. Paper presented at the 1996 (April) meeting of American Educational Research Association in New York, New York.
- Butler, D.L., & Winne, P.H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65, 3, 245-281.

Campione, J. C. (1989). Assisted assessment: A taxonomy of approaches and an outline of strengths and weaknesses. *Journal of Learning Disabilities*, 22, 151-165.

Carlson, J.S., & Wiedl, K.H. (1978). Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. *American Journal of Mental Deficiencies*, 82, 6, 559-565.

Carver, C.S., & Scheier, M.F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, 97, 1, 19-35.

CED-groep (2009). *PO Leerlijn Leren Leren Vakoverstijgend*. Rotterdam: CED.

CED-groep (2013). *VSO leerlijnen voorbereiding op dagbesteding & arbeid*. Rotterdam: CED.

Chan, R.C.K., Shum, D., Touloupoulou, T., & Chen, E.Y.H. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, 23, 2, 201-216.

Cito (2018). Handleiding Cito Volgsysteem Primair en speciaal onderwijs Leren leren sbo en so voormalig cluster 4. Cito: Arnhem.

Coleman, J.S. (1968). The mathematical study of change. In: M.M. Blalock, Jr. & A.B. Blalock (eds), *Methodology in social research* (pp. 428-478). New York: McGraw-Hill.

Cooper-Kahn, J., & Dietzel, L. (2011). *Vergeten, kwijt en afgeleid. Opvoedwijzer om executieve functies bij kinderen te versterken*. Amsterdam: Hofrege Uitgevers.

Corno, L. (1993). The Best-laid plans: Modern conceptions of volition and educational research. *Educational Researcher*, 22, 2, 14-22.

Dawson, P., & Guare, R. (2009). *Slim maar... Help kinderen hun talenten te benutten door hun executieve functies te versterken*. Amsterdam: Hogrefe.

Dawson, P. & R. Guare (2010). *Executieve functies bij kinderen en adolescenten. Een praktische gids voor diagnostiek en interventie*. Amsterdam: Hogrefe.

Deci, E.L., & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum.

Deci, E.L., & Ryan, R.M. (1991). A motivational approach to self: Integration in personality. In: R. Dienstbier (Ed.), *Nebraska symposium on motivation: Vol. 38. Perspectives on motivation* (pp. 237-288). Lincoln: University of Nebraska Press.

Deci, E.L., & Ryan, R.M. (2000). The "What" and "Why" of goal pursuits: Human needs and the selfdetermination of behavior. *Psychological Inquiry*, 11, 4, 227-268.

DeShazo, B.T., Lyman, R.D., & Grofer-Klinger, L. (2002). Academic underachievement and attention-deficit/hyperactivity disorder: The negative impact of symptom severity on school performance. *Journal of School Psychology*, 40, 3, 259-283.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition Learning*, 3, 231-265.

- Dignath, C., Büttner, G., & Langfeldt, H. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educational Research Review*, 3, 101-129.
- Dinsmore, D.L., & Parkinson, M.M. (2013). What are confidence judgments made of? Students' explanations for their confidence ratings and what that means for calibration. *Learning and Instruction*, 24, 4-15.
- Feuerstein, R. (1979). *The dynamic assessment of retarded performers*. Baltimore: University Park Press.
- Feuerstein, R., Feuerstein, R. S., & Falik, L. H. (2010). *Beyond smarter: Mediated learning and the brain's capacity for change*. New York, NY: Teachers College Press.
- Floor, H., Zoete, T. de , Schaafsma, M., & Oetelaar, M. v.d. (2013). Specifiek s(b)o. Landelijk Expertise Centrum Speciaal onderwijs, PO Raad en SBO werkverband.
<http://www.sbowerkverband.nl/bestanden/206/Samenvatting-S%28B%29O.pdf>
- Gennep, A.T.G. van (1994). *De zorg om het bestaan. Over verstandelijke handicap, kwaliteit en ondersteuning*. Amsterdam/Meppel: Boom.
- Gennep, A.T.G. van (1989). *De kwaliteit van het bestaan van de zwaksten in de samenleving*. Amsterdam: Boom Meppel.
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. Odessa, FL: Psychological Assessment Resources.
- Gioia, G.A., Isquith, P.K., Guy, S.C., Kenworthy, I., & Barton, R. (2002). Profiles of everyday executive functioning in acquired and developmental disorders. *Child Neuropsychology*, 8, 2, 121-137.
- Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, 124, 75–111.
- Haller, E., Child, D.A., & Walberg, H.J. (1988). Can comprehension be taught? A quantitative synthesis of "metacognitive" studies. *Educational Researcher*, 17, 5-8.
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, 66, 99-136.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112.
- Heise, R. (1969). Separating reliability and stability in test-retest correlation. *American Sociological Review*, 34, 1, 93-100.
- Inspectie van het Onderwijs (2010). *Analyse en waardering van opbrengsten primair onderwijs*. Utrecht: Inspectie van het Onderwijs.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 2, 130–145.
- Kaniel, S., Tzuriel, D., Feuerstein, R., Ben-Shacher, N., & Eitan, T. (1991). Dynamic assessment: Learning and transfer of Ethiopian immigrants to Israel. In: R. Feuerstein, P. Klein, & A. Tannenbaum (Eds.), *Mediated learning experience* (pp. 197-209). London: Freund.

- Kester, E.S. Peña, E.D., & Gillam, R.B. (2001). Outcomes of dynamic assessment with culturally and linguistically diverse students: A comparison of three teaching methods within a test-teach-retest framework. *Journal of Cognitive Education and Psychology*, 2, 37-54.
- Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 2, 254–285.
- Kuhlemeier, H. (2011). Het construeren van praktijktoetsen. In P. Sanders (red.), *Toetsen op school* (pp. 125-135). Arnhem: Cito.
- Kuhlemeier, H., & Chermin, J. (2015). *Het gebruikersonderzoek naar het functioneren van het Pakket leren leren* (ongepubliceerde projectnotitie). Arnhem: Cito.
- Kuhlemeier, H., Knoop, H., Boxtel, H. van, Papenburg, I., & Hollenberg, J. (2016). *Wetenschappelijke verantwoording VISEON 2.0. Volginstrument voor sociaal-emotioneel functioneren*. Arnhem: Cito.
- Kuhlemeier, H., Boxtel, H. van, & Til, A. van (2012). *Balans van de sociale opbrengsten in het basisonderwijs. Eerste meting voorjaar 2011* (PPON-reeks nummer 48). Arnhem: Cito.
- Ledoux, G., Meijer, J., Veen, I. van der, & Breetvelt, I. (2013). *Meetinstrumenten voor sociale competenties, metacognitie en advanced skills*. Amsterdam: Kohnstamm Instituut.
- Lidz, C.S. (1987). *Dynamic assessment: An interactional approach to evaluating learning potential*. New York, NY: Guilford Press.
- Lidz, C.S. (1991). *Practitioner's guide to dynamic assessment*. New York, NY: Guilford Press.
- Lidz, C. S., & Peña, E. D. (1996). Dynamic assessment: The model, its relevance as a non-biased approach and its application to Latino American preschool children. *Language, Speech, and Hearing Services in Schools*, 27, 367-372.
- Maes, B., Penne, A., & Petry, K. (2005). Kinderen en jongeren met een verstandelijke beperking. In: H. Grietens, J. Vanderfaellie, W. Hellinckx & A. Ruijsenaars (Red.), *Handboek orthopedagogische hulpverlening. 1. Een orthopedagogisch perspectief op kinderen en jongeren met problemen* (pp. 127-160). Leuven/Voorburg: Acco.
- Meichenbaum, D. (1981). *Cognitieve gedragsmodificatie. Een integrale benadering*. Deventer: Van Loghum Slaterus.
- Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). Washington, DC: American Council on Education.
- Ministerie van OCW. (1998). *Kerdoelen basisonderwijs*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Ministerie van OCW. (2006). *Kerdoelen basisonderwijs*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Ministerie van OCW. (2009). *Kerdoelen speciaal onderwijs*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Mönks, F.J. & Knoers, A.M.P. (2009). *Ontwikkelingspsychologie. Inleiding tot de verschillende deelgebieden*. Assen: Van Gorcum.

- Nijntjes, W., & Diepen, M. van (2008). *Maatschappelijke taken in het onderwijs*. Amsterdam: TSO Nipo.
- Passend Onderwijs (z.d.). *Over passend onderwijs*. Geraadpleegd op 26 oktober 2017
<https://www.passendonderwijs.nl/over-passend-onderwijs/in-het-kort/>
- Penã, E., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language assessment: A nonbiased procedure. *The Journal of Special Education*, 26, 269-280.
- Pintrich, P.R. (2000). The role of goal orientation in self-regulated learning. In: M. Boekaerts, P.R. Pintrich & M. Zeidner (eds.) *Handbook of self-regulation* (pp. 451-502). San Diego: Academic Press.
- Pintrich P.R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16, 385-407.
- Pintrich, P.R., & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In: A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation* (pp. 249–284). San Diego, CA: Academic Press.
- Resing, W.C.M. (1993). Measuring inductive reasoning skills: The construction of a learning potential test. In: J.H.M. Hamers, K. Sijtsma, & A.J.J.M. Ruijsenaars (Eds.), *Learning potential assessment: Theoretical, methodological, and practical issues* (pp.219–242). Lisse, Netherlands: Swets & Zeitlinger B.V.
- Resing, W.C.M. (1997). Leerpotentieel onderzoek: wat is de meerwaarde? In: T. Engelen-Snaterse & R. Kohnstamm (Eds.), *Kinder- en jeugdpsychologie. Trends* (pp. 283-303). Lisse: Swets & Zeitlinger.
- Resing, W.C.M. (2006). *Zicht op potentieel. Over dynamisch testen, variabiliteit in oplossingsgedrag en leerpotentieel van kinderen* (Oratie). Universiteit Leiden: Leiden.
- Rijksoverheid (z.d.). *Organisatie van het speciaal onderwijs*. Geraadpleegd op 19 oktober 2017
<https://www.rijksoverheid.nl/onderwerpen/passend-onderwijs/speciaal-onderwijs>
- Scheltinga, F., Keuning, J., & Kuhlemeier, H. (2014). *Gericht werken aan opbrengsten in taal- en leesonderwijs: Een systematische review naar toetsvormen*. Arnhem/Nijmegen: Cito/Expertisecentrum Nederlands.
- Schunk, D.H., & Zimmerman, B.J. (2008) (Eds.). *Motivation and self-regulated learning: Theory, research, and application*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Sluijsmans, D., Joosten-ten Brinke, D., & Van der Vleuten, C. (2013). *Toetsen met leerwaarde: een reviewstudie naar de effectieve kenmerken van formatief toetsen*. Maastricht: Universiteit Maastricht.
- Smidts, D. (2018). *Zelfsturing in de klas. Over aandacht, executieve functies en rust*. Amsterdam: Uitgeverij Nieuwezijds.
- Smidts, D., & Huizinga, M. (2011). *Gedrag in uitvoering. Over executieve functies bij kinderen en pubers*. Amsterdam: Uitgeverij Nieuwezijds.
- Smidts, D. & Huizinga, M. (2017). *Gedrag in uitvoering. Over executieve functies bij kinderen en pubers. Tweede editie*. Amsterdam: Uitgeverij Nieuwezijds.
- Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, 19, 1, 80-90.

- Sternberg, R.J., & Grigorenko, E.L. (2002). *Dynamic testing: The nature and measurement of learning potential*. Cambridge: Cambridge University Press.
- Straetmans, G.J.J.M. (2006). *Bekwaam beoordelen en beslissen. Beoordelen in competentiegerichte beroepsopleidingen* (lectorale rede). Enschede: Saxion Hogescholen.
- Swanson, H. L., & Lussier, C. M. (2001). A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research*, 71, 321-363.
- Thijs, A., Fisser, P., & Hoeven, M. van der (2014). *21e eeuwse vaardigheden in het curriculum van het funderend onderwijs*. Enschede: SLO.
- Timmer, J.T. (1994). *Development of a training program for sheltered workshops in the Netherlands*. Rijksuniversiteit Groningen: Academisch proefschrift. Groningen: Stichting kinderstudies.
- Timmer, J., Dekker, K.A., & Voortman, H. (2009). *Eigen initiatief model: theorie en toepassing*. Utrecht: Nederlands Instituut voor Zorg en Welzijn.
- Timmer, J.T., & Gemert, G.H. van (1992). *Theoretical backgrounds of a training program. Training employees with mental retardation*. In: H. Nakken, G.H. Van Gemert & T. Zandberg (eds.), *Research on intervention in special education*. Lewiston, NY: The Edwin Mellen Press.
- Timmerman, K. (2002). *Kinderen met aandachts- en werkhoudingsproblemen*. Leuven: Acco.
- Timmerman, K. (2017). *Aandachts- en werkhoudingsproblemen bij kinderen en jongeren*. Leuven: Acco.
- Tzuriel, D. (2001). *Dynamic assessment of young children*. New York, NY: Kluwer Academic/Plenum.
- Voogt, J., & Pareja Roblin, N.P. (2010). 21st century skills. Discussienota. Zoetermeer: Kennisnet.
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wang, M.C., Haertel, G.D., & Walberg, H.J. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research*, 84, 30–43.
- Winne, P.H. (1995). Inherent details in self-regulated learning. *Educational Psychologist*, 30, 4, 173-188.
- Zimmerman, B.J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25, 3-17.
- Zimmerman, B.J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego, CA, US: Academic Press.
- Zimmerman, B.J. (2002). Becoming a self-regulated learner: an overview. *Theory into Practice*, 41, 64-70.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45, 1, pp. 166-183.
- Zumbrunn, S., Tadlock, J., & Roberts, E.D. (2011). *Encouraging self-regulated learning in the classroom: A review of the literature*. Richmond, Virginia: Metropolitan Educational Research Consortium (MERC).

Bijlagen

Bijlage 1 Overzicht van doelen per hoofdvaardigheid van Leren leren

In deze bijlage is aangegeven hoe de schalen van de Checklist en de aandachtspunten uit het Beoordelingsformulier bij de praktische opdrachten zijn gerelateerd aan de leergebiedoverstijgende kerndoelen en de vakoverstijgende leerlijnen voor leren leren van de CED-groep (2009).

Deze opsomming van doelen is niet uitputtend maar alleen illustratief bedoeld.

Doelen bij vaardigheden tijdens de oriëntatie en voorbereiding van de opdracht

Voor leerlingen die zwak zijn in oriëntatie en voorbereiding zijn de volgende doelen mogelijk van belang:

- Luistert en kijkt naar de uitleg van een opdracht in een een-op-een situatie (bron: CED-leerlijn 1)
- Wijst aan waar hij moet beginnen (bron: CED-leerlijn 1)
- Pakt en ruimt zijn materialen op (bron: CED-leerlijn 1)
- Herhaalt de opdracht die hem is gegeven (wat ga je nu doen?) (bron: CED-leerlijn 1)
- Begint na de uitleg met zijn taak (bron: CED-leerlijn 1)
- Bepaalt zelf in welke volgorde hij een taak uitvoert (bron: CED-leerlijn 1)
- Leest van het bord of taakkaart wat en hoeveel hij moet doen (bron: CED-leerlijn 1)
- Bedenkt en vertelt hoe hij een taak gaat aanpakken (bron: CED-leerlijn 1)
- Weet wanneer hij hulp nodig heeft (bron: CED-leerlijn 1)
- Maakt een plan voor de aanpak van een klein project (volgorde van taken, duur van taken, tijdsindeling, benodigdheden) (bron: CED-leerlijn 1)
- Geeft aan wat hij verwacht te kunnen: qua eigen vaardigheden (bron: CED-leerlijn 1)

Doelen bij vaardigheden tijdens de uitvoering van de opdracht inclusief monitoren tijdens de opdracht

Voor leerlingen die zwak zijn in vaardigheden tijdens de uitvoering van de opdracht zijn de volgende doelen mogelijk van belang:

- Voert een simpele enkelvoudige taak tot het einde uit (bron: CED-leerlijn 1)
- Verbeterd na aanwijzingen de manier waarop hij een taak uitvoert (bron: CED-leerlijn 1)
- Achterhaalt bij een veelvoorkomende (of belangrijke) fout waar de fout zit (bron: CED-leerlijn 1)
- Probeert de fout herstellen (bron: CED-leerlijn 1)
- Lost zelf praktische zaken/problemen op (bron: CED-leerlijn 2)
- Maakt een langdurige opdracht af als daar af en toe aanwijzingen voor worden gegeven (bron: CED-leerlijn 3)
- Maakt gebruik van hulp (bron: CED-leerlijn 3)
- Werkt harder door om een taak af te krijgen als de leerkracht dit aangeeft (bron: CED-leerlijn 3)
- Volgt een stappenplan om een grotere taak uit te voeren (bron: CED-leerlijn 3)
- Werkt netjes en/of snel afhankelijk van de eisen die aan de taakuitvoering worden gesteld (bron: CED-leerlijn 3)
- Zoekt een oplossing bij tegenslag met een taak (aanwijzingen onleesbaar, spullen onvindbaar) (bron: CED-leerlijn 3)
- Past zijn werkstijl (kwaliteit t.o.v. hoeveelheid) aan waar nodig (bron: CED-leerlijn 5)
- Geeft tijdens een activiteit aan dat hij het (te) makkelijk of moeilijk vindt (bron: CED-leerlijn 5)
- Controleert zijn eigen taak en corrigeert waar nodig (bron: CED-leerlijn 1)
- Controleert of de taak klaar is (alles gebruikt) (bron: CED-leerlijn 1)

Doelen bij vaardigheden van reflectie en evaluatie

Voor leerlingen die zwak zijn in vaardigheden van reflectie en evaluatie zijn de volgende doelen mogelijk van belang:

- Geeft aan wanneer hij toe is aan een moeilijkere taak (bron: CED-leerlijn 1)
- Vertelt na afloop van een taak wat hij heeft gedaan (bron: CED-leerlijn 5)
- Geeft aan welke taken hij moeilijk vond en welke makkelijk (bron: CED-leerlijn 5)
- Controleert het gemaakte werk (bron: CED-leerlijn 5)
- Geeft aan of de taak makkelijk of moeilijk was en of hij goed heeft gewerkt (bron: CED-leerlijn 5)
- Vertelt in een paar stappen hoe hij zijn taak heeft aangepakt (bron: CED-leerlijn 5)
- Zoekt een fout op als de leerkracht vraagt nog eens goed te kijken (bron: CED-leerlijn 5)
- Beoordeelt samen met de leerkracht hoe hij zijn taak gedaan heeft werkwijze, snel/langzaam, resultaat) (bron: CED-leerlijn 5)
- Onderbouwt zijn mening over de kwaliteit van een gemaakte taak (goed, omdat) (bron: CED-leerlijn 5)
- Kijkt naar en praat over overeenkomsten en verschillen tussen eigen en andermans werk/aanpak (bron: CED-leerlijn 5)
- Geeft aan op welk gebied zich een probleem voordeed en welke oplossing is gekozen (bij materiaal, taak of samenwerking) (bron: CED-leerlijn 5)
- Vertelt achteraf waarom iets goed, minder goed of fout is gegaan (bron: CED-leerlijn 5)
- Legt samen met leerkracht een relatie tussen deze beoordeling en afspraken voor de volgende keer (bron: CED-leerlijn 5)
- Evalueert de periode van zelfstandig werken gericht op de voorbereiding, het proces en het resultaat (bron: CED-leerlijn 5)
- Reflecteert zowel op het resultaat van een taak als op het proces (hoe heb ik het aangepakt) (bron: CED-leerlijn 5)
- Onderscheidt daarbij persoonlijke (kun je zelf beïnvloeden) en externe oorzaken (bron: CED-leerlijn 5)
- Verbindt consequenties voor de volgende keer aan zijn beoordeling (bron: CED-leerlijn 5)
- Beoordeelt of hij een uitgevoerde taak goed had voorbereid en uitgevoerd en verbindt hier consequenties aan voor de volgende keer (bron: CED-leerlijn 5)
- Beoordeelt of hij de juiste prioriteiten heeft gesteld en geeft aan welke gevolgen dit heeft voor een volgende planning (bron: CED-leerlijn 5)
- Komt met suggesties voor aanpassingen op inhoudelijke (wat wil/moet ik leren) en procesmatige aspecten van het onderwijsleerproces (bron: CED-leerlijn 5)

Doelen bij werkhouding: aandachtig werken, gemotiveerd werken en werken zonder gedragsproblematiek

Voor leerlingen met een minder adequate werkhouding zijn de volgende doelen mogelijk van belang:

- Begint na de uitleg met zijn taak (bron: CED-leerlijn 1)
- Gaat zorgvuldig met materialen om (bron: CED-leerlijn 1)
- Zet door en past oplossingen toe bij een taak die hij moeilijk of eng vindt (bron: CED-leerlijn 1)
- Hanteert de afgesproken regels om problemen op te lossen (bv. hulp van medeleerling of leerkracht, opzoeken) (bron: CED-leerlijn 2)
- Voert een korte, voorgestructureerde opdracht uit en maakt die af zonder te stoppen (bron: CED-leerlijn 3)
- Probeert de taak nog een keer wanneer het de eerste keer niet lukt (bron: CED-leerlijn 3)
- Werkt door bij interne afleiding (schrijft vragen en gedachten die afleiden op) (bron: CED-leerlijn 3)
- Werkt door bij externe afleiding (onderbouw speelt buiten, maken lawaai) (bron: CED-leerlijn 3)
- Lost een klein materiaalprobleem op zonder anderen te storen (papier vol, ander potlood nodig) (bron: CED-leerlijn 3)
- Zet zich langere tijd in voor een taak die hij niet leuk vindt of die niet lukt (bron: CED-leerlijn 3)
- Reageert zonder boos te worden op kritiek op zijn werk (bron: CED-leerlijn 5)

Bijlage 2 De uiteindelijke versie van de checklist

Checklist Leren leren

		1	2	3	4
		(bijna) nooit	soms	vaak	(bijna) altijd
1 Oriënteren, voorbereiden en plannen van praktische taken					
1.1	Is betrokken bij de instructie in een groepssituatie.				
1.2	Leest de opdracht vóóordat hij start met de opdracht.				
1.3	Legt het benodigde materiaal voor de opdracht klaar.				
		1	2	3	4
		zeer zwak ontwikkeld	zwak ontwikkeld	matig ontwikkeld	sterk ontwikkeld
1.4	Kan uitleggen wat hij gaat doen.				
1.5	Kan uitleggen hoe hij de opdracht gaat aanpakken.				
1.6	Maakt een realistische planning voor een opdracht.				
1.7	Is in staat prioriteiten te stellen.				
1.8	Kan vooraf benoemen met welk eindresultaat hij tevreden is.				

		1	2	3	4
		(bijna) nooit	soms	vaak	(bijna) altijd
2 Uitvoeren van praktische taken					
2.1	Raadpleegt de opdracht opnieuw wanneer hij vastloopt				
		1	2	3	4
		zeer zwak ontwikkeld	zwak ontwikkeld	matig ontwikkeld	sterk ontwikkeld
2.2	Voert de stappen van een opdracht in de aangegeven volgorde uit.				
2.3	Probeert een gemaakte fout te herstellen.				
2.4	Past na aanwijzing van de leerkracht zijn werkhouding aan.				
2.5	Lost een taakinhoudelijk probleem zelf op.				
2.6	Gebruikt de beschikbare tijd op een efficiënte manier.				
2.7	Kijkt hoe de leerkracht een taak aanpakt en neemt deze aanpak over.				
2.8	Vult zijn werk aan als hij iets vergeten is.				

		1	2	3	4
		zeer zwak ontwikkeld	zwak ontwikkeld	matig ontwikkeld	sterk ontwikkeld
3 Monitoren, reflecteren en evalueren van praktische taken					
3.1	Houdt in de gaten of de opdracht volgens plan verloopt.				
3.2	Vraagt zich bij een fout af: hoe komt het dat ik dit fout doe?				
3.3	Beoordeelt of zijn voorbereiding goed is geweest.				
3.4	Vergelijkt zijn eigen eindproduct met de opdracht (of het voorbeeld).				
3.5	Beoordeelt het eindresultaat.				
3.6	Onderbouwt zijn beoordeling van het eindresultaat.				
3.7	Benoemt wat goed is gegaan.				
3.8	Benoemt wat niet goed is gegaan.				
3.9	Geeft aan wat hij heeft geleerd van de opdracht.				
3.10	Benoemt wat hij de volgende keer anders zou doen.				

		1	2	3	4
		(bijna) nooit	soms	vaak	(bijna) altijd
4 Aandachtig werken					
4.1	Luistert geconcentreerd naar een langere instructie.				
4.2	Richt zijn aandacht - na het starten - op het uitvoeren van de opdracht.				
4.3	Richt na afleiding zijn aandacht opnieuw op de opdracht.				
		1	2	3	4
		zeer zwak ontwikkeld	zwak ontwikkeld	matig ontwikkeld	sterk ontwikkeld
4.4	Kan zijn aandacht bij het werk houden tot de opdracht af is. (spanningsboog)				
4.5	Werkt aandachtig in een rustige groepssituatie.				
4.6	Kan geconcentreerd aan een korte opdracht werken.				
4.6	Werkt geconcentreerd aan opdrachten die hij niet leuk vindt.				
4.7	Kan geconcentreerd aan een langere opdracht werken.				

		1	2	3	4
		(bijna) nooit	soms	vaak	(bijna) altijd
5 Gemotiveerd werken					
5.1	Werkt met enthousiasme aan een opdracht die aansluit bij zijn interesse.				
5.2	Werkt met betrokkenheid aan een opdracht die niet aansluit bij zijn interesse.				
		1	2	3	4
		zeer zwak ontwikkeld	zwak ontwikkeld	matig ontwikkeld	sterk ontwikkeld
5.3	Reageert op vragen die de leerkracht stelt tijdens de instructie.				
5.4	Werkt zelfstandig zonder aansporing van de leerkracht.				
5.5	Zet door ook als de opdracht moeilijk is.				
5.6	Zet door op het moment dat hij tegenslag ervaart.				
5.7	Maakt af waar hij aan begonnen is.				
5.8	Kan zichzelf motiveren voor de opdracht.				

		1	2	3	4
		(bijna) nooit	soms	vaak	(bijna) altijd
6 Werken zonder gedragsproblematiek (houding)					
6.1	Stoort anderen tijdens het werken.				
6.2	Reageert impulsief.				
6.3	Vraagt op ongepaste wijze hulp wanneer hij vastloopt.				
6.4	Vraagt continu om bevestiging van de leerkracht.				
6.5	Vertoont storend gedrag als hij tegen een werkprobleem aanloopt.				
		1	2	3	4
		zeer zwak ontwikkeld	zwak ontwikkeld	matig ontwikkeld	sterk ontwikkeld
6.6	Kan omgaan met kritiek op gemaakt werk en/of de werkhouding.				
6.7	Zet door op het moment dat hij gefrustreerd raakt.				
6.8	Kan emoties in bedwang houden, zodat ze de uitvoering van de opdracht niet beïnvloeden.				

		1	2	3	4
		(bijna) nooit	soms	vaak	(bijna) altijd
7 Zelfstandig werken en hulp vragen					
7.1	Past de afgesproken regels van het zelfstandig werken toe.				
7.2	Probeert eerst zelf het probleem op te lossen voor hij hulp vraagt.				
7.3	Vraagt hulp op een passend moment.				
		1	2	3	4
		zeer zwak ontwikkeld	zwak ontwikkeld	matig ontwikkeld	sterk ontwikkeld
7.4	Maakt bij de start van een opdracht een inschatting: Kan ik aan de slag of heb ik eerst hulp nodig?				
7.5	Werkt zelfstandig.				
7.6	Kan een hulpvraag goed verwoorden.				
7.7	Kan omgaan met uitgestelde aandacht.				
7.8	Werkt aan een alternatieve opdracht als hij niet verder kan.				

Bijlage 3 De uiteindelijke versie van de beoordelingsformulieren

Formulier 1: fase 1 oriëntatie & fase 2 voorbereiding

Beoordelingsaspect	Niveau 4	Niveau 3	Niveau 2	Niveau 1
1.1 'Wat ga je doen?'	Leerling is zich zichtbaar bewust van/ kan vertellen wat hij/zij gaat doen met 0 of 1 hints <input type="checkbox"/> 4	Leerling is zich zichtbaar bewust van/ kan vertellen wat hij/zij gaat doen met 2 hints <input type="checkbox"/> 3	Leerling is zich zichtbaar bewust van/ kan vertellen wat hij/zij gaat doen met 3 of meer hints <input type="checkbox"/> 2	Leerling is zich niet bewust van/ kan niet vertellen wat hij/zij gaat doen ondanks vele hints <input type="checkbox"/> 1
1.2 'Waar begin je mee?' en 'Hoe ga je verder?'	Volgorde stappen juist	Volgorde stappen juist	Volgorde stappen juist	Volgorde stappen meer fout dan goed
NB Het gaat erom dat de leerling globaal de stappen van de opdracht in beeld heeft.	met 0 of 1 hints <input type="checkbox"/> 4	met 2 hints <input type="checkbox"/> 3	met 3 of meer hints <input type="checkbox"/> 2	ondanks vele hints <input type="checkbox"/> 1
1.3 'Heb je alle spullen klaargezet?' (*Wat moet je nog pakken?*)	Heeft alle benodigde spullen gepakt en geen onnodige spullen gepakt	Heeft alle benodigde spullen gepakt en geen onnodige spullen gepakt	Heeft alle benodigde spullen gepakt en geen onnodige spullen gepakt	Het pakken van de juiste spullen lukt niet
NB Te beoordelen nadat je de leerling in de gelegenheid hebt gesteld om eventuele fouten te herstellen.	met 0 of 1 hints <input type="checkbox"/> 4	met 2 hints <input type="checkbox"/> 3	met 3 of meer hints <input type="checkbox"/> 2	ondanks vele hints <input type="checkbox"/> 1
1.4 Hoeveel hulp heb je de leerling bij de oriëntatie en voorbereiding in totaal moeten geven?	geen hulp <input type="checkbox"/> 4	af en toe <input type="checkbox"/> 3	regelmatig <input type="checkbox"/> 2	voortdurend <input type="checkbox"/> 1

Ruilte om het aantal & type hints te turven (indien gewenst); (zie § 3.3 van de handleiding)					Ruilte voor opmerkingen (voor jezelf)																																									
Verbaal	Turven	Totaal	Non Verbaal	Turven	Totaal																																									
A) Algemene hint B) Open vraag C) Gesloten vraag D) Oplossingen voorleggen E) Vertellen of voordoen		<table border="1"><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr></table>																					F) Schoudera ophalen G) Vragend kijken H) (Aan)Wijzen I) Ja knikken		<table border="1"><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr></table>																					

Formulier 2: fase 3 uitvoering

Beoordelingsaspect	Niveau 4	Niveau 3	Niveau 2	Niveau 1
2.1 Werken volgens (eigen) plan	Werkt systematisch volgens (eigen) plan met 0 of 1 hint <input type="checkbox"/> 4	Werkt systematisch, maar niet geheel volgens (eigen) plan met 1 of 2 hints <input type="checkbox"/> 3	Werkt min of meer systematisch met 2 of 3 hints <input type="checkbox"/> 2	Werkt chaotisch zonder aandacht voor plan ondanks vele hints <input type="checkbox"/> 1
2.2 Zelfmonitoring	Denkt regelmatig na voor iets te doen zonder hints <input type="checkbox"/> 4	Denkt regelmatig na voor iets te doen met 1 of 2 hints <input type="checkbox"/> 3	Denkt af en toe na voor iets te doen met 3 hints <input type="checkbox"/> 2	Denkt zelden na voor iets te doen ondanks vele hints <input type="checkbox"/> 1
2.3 Controle eigen werk	Controleert eigen werk regelmatig zonder hints <input type="checkbox"/> 4	Controleert eigen werk regelmatig met 1 of 2 hints <input type="checkbox"/> 3	Controleert eigen werk af en toe met 3 hints <input type="checkbox"/> 2	Controleert eigen werk nooit ondanks vele hints <input type="checkbox"/> 1
2.4 Hoeveel hulp heb je de leerling bij de uitvoering in totaal moeten geven?	nauwelijks <input type="checkbox"/> 4	af en toe <input type="checkbox"/> 3	regelmatig <input type="checkbox"/> 2	voortdurend <input type="checkbox"/> 1

Ruimte om het aantal & type hints te turven (indien gewenst); (zie § 3.3 van de handleiding)															
Verbaal	Turven	Totaal	Non Verbaal	Turven	Totaal										
A) Algemene hint B) Open vraag C) Gesloten vraag D) Oplossingen voorleggen E) Vertellen of voordoen		<table border="1"><tr><td></td><td></td><td></td><td></td><td></td></tr></table>						F) Schouderaars ophalen G) Vragend kijken H) (Aan)Wijzen I) Ja knikken		<table border="1"><tr><td></td><td></td><td></td><td></td><td></td></tr></table>					

Formulier 3: fase 4 reflectie & evaluatie

Beoordelingsaspect	Niveau 4	Niveau 3	Niveau 2	Niveau 1
3.1 'Lijken ze op elkaar?' (product) (Leerkracht wijst werk leerling aan en het voorbeeld)	Is zich zichtbaar bewust van/kan uitleggen in hoeverre zijn/haar product lijkt op het voorbeeld/past bij de opdracht zonder hints <input type="checkbox"/> 4	Is zich zichtbaar bewust van/kan uitleggen in hoeverre zijn/haar product lijkt op het voorbeeld/past bij de opdracht met 1 hint <input type="checkbox"/> 3	Is zich zichtbaar bewust van/kan uitleggen in hoeverre zijn/haar product lijkt op het voorbeeld/past bij de opdracht met 2 of meer hints <input type="checkbox"/> 2	Leerling geeft geen zinnig antwoord/ reactie ondanks vele hints <input type="checkbox"/> 1
Bij geen voorbeeld: 'Heb je de opdracht goed gedaan?' (product)				
3.2 'Wat vond je moeilijk?' (terugkijkvraag) (proces)	Geeft aan wat hij/zij moeilijk vond zonder hints <input type="checkbox"/> 4	Geeft aan wat hij/zij moeilijk vond met 1 hint <input type="checkbox"/> 3	Geeft aan wat hij/zij moeilijk vond met 2 of meer hints <input type="checkbox"/> 2	Kan niet aangeven wat hij/zij moeilijk vond ondanks vele hints <input type="checkbox"/> 1
3.3 'Wat vond je makkelijk?' (terugkijkvraag) (proces)	Geeft aan wat hij/zij makkelijk vond zonder hints <input type="checkbox"/> 4	Geeft aan wat hij/zij makkelijk vond met 1 hint <input type="checkbox"/> 4	Geeft aan wat hij/zij makkelijk vond met 2 of meer hints <input type="checkbox"/> 2	Kan niet aangeven wat hij/zij makkelijk vond ondanks vele hints <input type="checkbox"/> 1
3.4 'Als je deze opdracht nog een keer moet doen, wat zou je dan anders doen?' (voorkijkvraag) (proces)	Geeft aan wat hij/zij de volgende keer anders zou doen zonder hints <input type="checkbox"/> 4	Geeft aan wat hij/zij de volgende keer anders zou doen met 1 hint <input type="checkbox"/> 4	Geeft aan wat hij/zij de volgende keer anders zou doen met 2 of meer hints <input type="checkbox"/> 2	Geeft geen enkel verbeterpunt aan of geeft geen zinnig antwoord (terwijl daar wel aanleiding voor is) ondanks vele hints <input type="checkbox"/> 1
3.5 Hoeveel hulp heb je de leerling bij de reflectie en evaluatie in totaal moeten geven?	geen hulp <input type="checkbox"/> 4	af en toe <input type="checkbox"/> 3	regelmatig <input type="checkbox"/> 2	voortdurend <input type="checkbox"/> 1

Vervolg Formulier 3: fase 4 reflectie & evaluatie

Ruimte om het aantal & type hints te turven (indien gewenst); (zie § 3.3 van de handleiding)				Ruimte voor opmerkingen (voor jezelf)																																	
Verbaal	Turven	Totaal	Non Verbaal	Turven	Totaal																																
A) Algemene hint B) Open vraag C) Gesloten vraag D) Oplossingen voorleggen E) Vertellen of voordoen		<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>																	F) Schouder ophalen G) Vragend kijken H) (Aan)Wijzen I) Ja knikken		<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>																

Formulier 4: Werkhouding

Beoordelingsaspect	Niveau 4	Niveau 3	Niveau 2	Niveau 1
4.1 Aandachtig werken	Werkt geconcentreerd zonder daartoe aangemoedigd te hoeven worden <input type="checkbox"/> 4	Werkt geconcentreerd maar heeft wel enige aanmoediging nodig <input type="checkbox"/> 3	Werkt alleen geconcentreerd als de leerkracht hem daartoe aanmoedigt <input type="checkbox"/> 2	Werkt ongeconcentreerd ondanks intensieve aanmoediging <input type="checkbox"/> 1
4.2 Gemotiveerd werken	Is zeer gemotiveerd <input type="checkbox"/> 4	Is gemotiveerd <input type="checkbox"/> 3	Is weinig gemotiveerd <input type="checkbox"/> 2	Is niet gemotiveerd <input type="checkbox"/> 1
4.3 Werken zonder gedragsproblematiek (het kan om externaliserend gedrag en/of internaliserend gedrag gaan)	Werk wordt niet gehinderd door gedragsproblematiek <input type="checkbox"/> 4	Werk wordt minimaal gehinderd door gedragsproblematiek <input type="checkbox"/> 3	Werk wordt enigszins gehinderd door gedragsproblematiek <input type="checkbox"/> 2	Werk wordt regelmatig gehinderd door gedragsproblematiek <input type="checkbox"/> 1
Ruimte voor opmerkingen (voor uzelf)				

Bijlage 4 De omzetting van somscores op de checklist naar percentielen

Tabel 1 Omzettingstabel totaalscore checklist

Somscore	Percentiel	Somscore	Percentiel	Somscore	Percentiel	Somscore	Percentiel
58	0,22	108	9,07	158	46,44	208	92,87
59	0,22	109	9,72	159	47,52	209	93,09
60	0,22	110	9,72	160	48,60	210	93,30
61	0,22	111	10,37	161	50,76	211	93,52
62	0,22	112	10,58	162	51,19	212	94,38
63	0,22	113	11,88	163	52,05	213	94,60
64	0,22	114	12,53	164	52,48	214	94,82
65	0,22	115	12,96	165	53,78	215	95,90
66	0,22	116	13,39	166	55,29	216	96,33
67	0,22	117	13,82	167	55,72	217	96,54
68	0,22	118	14,47	168	56,80	218	96,54
69	0,43	119	15,12	169	58,10	219	96,54
70	0,43	120	15,98	170	59,83	220	96,54
71	0,43	121	16,41	171	61,56	221	97,41
72	0,43	122	17,06	172	63,71	222	98,49
73	0,43	123	17,93	173	64,79	223	98,70
74	0,65	124	18,14	174	66,74	224	98,92
75	0,65	125	19,44	175	67,60	225	98,92
76	0,65	126	20,30	176	68,47	226	99,35
77	0,65	127	21,81	177	68,68	227	99,35
78	1,08	128	22,68	178	70,19	228	99,78
79	1,30	129	23,97	179	70,63	229	99,78
80	1,30	130	24,84	180	72,14	230	100,00
81	1,30	131	25,70	181	72,79	231	100,00
82	1,51	132	26,78	182	73,65	232	100,00
83	1,51	133	27,65	183	74,30		
84	1,51	134	28,51	184	75,16		
85	1,51	135	29,16	185	76,46		
86	1,51	136	30,45	186	77,54		
87	2,16	137	30,67	187	78,40		
88	2,16	138	33,69	188	78,83		
89	2,16	139	34,56	189	79,05		
90	2,38	140	35,42	190	79,91		
91	2,59	141	35,42	191	80,78		
92	3,24	142	35,64	192	82,51		
93	4,10	143	36,29	193	82,94		
94	4,75	144	36,72	194	83,59		
95	5,18	145	37,80	195	84,67		
96	5,40	146	38,66	196	85,10		
97	5,40	147	39,09	197	85,53		
98	5,83	148	39,31	198	85,96		
99	6,05	149	39,96	199	86,61		
100	6,26	150	40,82	200	87,26		
101	6,48	151	41,47	201	87,90		
102	7,13	152	41,90	202	89,20		
103	7,34	153	42,55	203	90,28		
104	7,78	154	43,63	204	90,93		
105	8,21	155	44,06	205	91,36		
106	8,21	156	44,92	206	92,22		
107	8,64	157	45,57	207	92,66		

Tabel 2 Omzettingstabel somscores cognitieve vaardigheden checklist naar percentielen

Oriënteren, voorbereiden en plannen	
Somscore	Percentiel
8	0,50
9	1,30
10	3,70
11	5,50
12	7,90
13	10,40
14	12,80
15	17,60
16	22,00
17	28,20
18	35,30
19	40,30
20	46,90
21	52,90
22	60,40
23	65,80
24	71,60
25	77,10
26	81,70
27	86,80
28	91,20
29	93,60
30	96,30
31	98,50
32	100,00

Uitvoeren	
Somscore	Percentiel
8	0,30
9	0,50
10	1,30
11	3,50
12	4,70
13	6,50
14	8,70
15	12,70
16	16,40
17	19,70
18	25,60
19	30,90
20	37,30
21	45,00
22	50,30
23	58,00
24	64,90
25	68,60
26	75,30
27	79,40
28	84,30
29	88,80
30	93,50
31	96,70
32	100,00

Monitoren, reflecteren en evalueren	
Somscore	Percentiel
10	1,50
11	2,10
12	2,60
13	4,30
14	7,20
15	10,00
16	13,00
17	15,70
18	18,10
19	22,30
20	26,10
21	29,70
22	34,00
23	37,80
24	44,60
25	49,00
26	54,40
27	60,10
28	63,30
29	66,70
30	71,60
31	75,40
32	80,00
33	83,70
34	86,80
35	90,20
36	91,70
37	93,80
38	95,70
39	98,10
40	100,00

Tabel 3 Omzettingstabel somscores werkhouding checklist naar percentielen

Aandachtig werken		Gemotiveerd werken		Werken zonder gedragsproblematiek		Zelfstandig werken en hulp vragen	
Score	Percentiel	Score	Percentiel	Score	Percentiel	Score	Percentiel
8	1,20	8	0,50	8	0,50	8	0,50
9	1,70	9	0,50	9	0,50	9	0,50
10	2,70	10	2,15	10	2,10	10	1,90
11	4,80	11	3,47	11	3,50	11	3,70
12	6,60	12	4,46	12	4,50	12	5,80
13	8,80	13	6,61	13	6,60	13	9,10
14	11,90	14	8,93	14	8,90	14	11,40
15	15,40	15	12,73	15	12,70	15	13,50
16	18,90	16	16,03	16	16,00	16	16,90
17	22,20	17	19,01	17	19,00	17	22,10
18	27,70	18	23,14	18	23,10	18	27,10
19	33,00	19	29,26	19	29,30	19	33,00
20	36,70	20	34,88	20	34,90	20	39,20
21	41,10	21	40,66	21	40,70	21	45,50
22	48,60	22	48,60	22	48,60	22	52,70
23	53,40	23	56,20	23	56,20	23	58,50
24	60,50	24	61,65	24	61,70	24	65,90
25	63,70	25	65,62	25	65,60	25	72,80
26	69,00	26	73,72	26	73,70	26	78,70
27	74,30	27	79,34	27	79,30	27	84,50
28	79,30	28	86,45	28	86,40	28	89,10
29	84,40	29	90,74	29	90,70	29	93,30
30	90,40	30	94,05	30	94,00	30	95,60
31	93,90	31	96,86	31	96,90	31	97,70
32	100,00	32	100,00	32	100,00	32	100,00

Cito

Amsterdamseweg 13
6814 CM Arnhem
Postbus 1034
6801 MG Arnhem
T (026) 352 11 11

Fotografie: Gijs Versteeg

