

Onderwijzen van inductief redeneren in groep 3 van de basisschool*

J.H.M. Hamers, E. de Koning en K. Sijsma

Samenvatting

De resultaten worden gepresenteerd van een formatieve en een summatieve evaluatie van een trainingsprogramma voor inductief redeneren, gebaseerd op het Duitse programma 'Denk-training für Kinder I' (Klauer, 1989). In de formatieve studie werd ten eerste nagegaan of het programma uitvoerbaar was in de natuurlijke setting van de schoolklas en ten tweede werden de leereffecten van het programma bepaald. Leerlingen van twee experimentele groepen kregen lessen in inductief redeneren aangeboden volgens twee verschillende didactische werkwijzen, een 'banende' en een 'sturende'. De resultaten van lesobservaties en rapportages van betrokken leraren lieten zien dat leraren het programma volgens de voorgeschreven instructies konden uitvoeren. Ten tweede behaalden beide experimentele groepen significant hogere natest- en vervolgttest-scores dan de controle-groep. Verdere analyses wezen uit dat de leereffecten het grootst waren in de experimentele groep waarin het programma volgens de sturende methode werd onderwezen. In de summatieve studie werd hetzelfde programma aangeboden aan een grotere groep van leerlingen. De instructie van het programma was op onderdelen aangepast op basis van de resultaten van de eerste studie. De experimentele groepen bleken significant vooruit te gaan in vergelijking met de controlegroepen.

1 Inleiding

In de geschiedenis van het onderwijs is wellicht nog nooit zoveel aandacht geweest voor

stimulering van het denken als heden ten dage (o.a. Klauer & Phye, 1995; Resnick, 1987). Deze aandacht is niet nieuw. Er zijn altijd psychologen en pedagogen geweest die de ontwikkeling van de denkbekwaamheid van kinderen hebben gezien als een belangrijk onderwijsdoel. Onderwijs in rekenen en wiskunde bijvoorbeeld zijn in dit kader veel besproken (zie Hamers, Van Luit & Csapó, in voorbereiding). Een van de eersten die wees op de wenselijkheid van denkstimulering was Dewey (1911/1991). Andere belangrijke impulsen uit het verleden komen voort uit de Würzburger School (Külpe, Selz) en uit de Gestaltpsychologie (Wertheimer, Duncker, Maier). In 1930 schreef Kohnstamm over Selz: "Aanvankelijk was het hem alleen te doen om de 'probleemoplossingsmethoden' zelf te vinden, maar in den jonsten tijd is hij nu ook de vraag gaan stellen naar de 'leerbaarheid' van deze methoden" (p. 28). Selz (1935) verrichtte onder meer onderzoek naar inductief redeneren. Hij hanteerde daarbij het uitgangspunt dat het denkproces bestaat uit het toepassen van ordeningsmiddelen, denkschema's, die door hun op de oplossing van het denkproces anticiperende karakter het verloop van het denkproces bepalen. Uijtgaande hiervan moest het volgens Selz mogelijk zijn de menselijke intelligentie op een hoger niveau van functioneren te brengen door de daartoe benodigde denkmiddelen aan te reiken. Het hanteren van oplossingsmethoden was naar zijn visie leerbaar. In zijn experimenten 'Versuche zur Hebung des Intelligenzniveaus' werkte hij volgens het principe van de 'kleinstmögliche Hilfe', in hedendaagse termen de heuristische oplossingsmethode.

Resnick (1987) meent dat de behoefte aan denkstimulering gezocht moet worden in de snelle veranderingen in de huidige maatschappij. Kennis en informatie wordt complexer en zijn snel weer verouderd. Het is vrijwel onmogelijk om alle nieuwe informatie op te nemen, te beoordelen, te verwerken en vast te houden.

* Een deel van deze studies werd uitgevoerd met subsidie van het C. Wright Mills Fonds van de Faculteit der Sociale Wetenschappen van de Universiteit Utrecht.

Kinderen zouden daarom moeten worden toe-gerust om door logisch redeneren keuzes te kunnen afwegen en beargumenteren en proble-men te kunnen identificeren en oplossen. Dit betekent dat het niet voldoende is te beschik-ken over aanzienlijke hoeveelheden kennis (declaratieve kennis), maar dat ook de vraag belangrijk is naar efficiënte wijzen van kennis vergaren en het effectief toepassen van deze kennis (procedurele kennis).

Voorts zijn anderen (Halpern, 1992; Res-nick & Klopfer, 1989) van mening dat een beperkte beheersing van denkvaardigheden één van de oorzaken van het achterblijven van schoolprestaties is, daar in het rekenen, het lezen en het schrijven allerlei activiteiten voor-komen waarin denkvaardigheden een centrale rol spelen. Voorbeelden van denkvaardigheden zijn: het kunnen beschrijven en vergelijken van objecten, het groeperen van objecten, het kun-nen leggen van verbanden, het vormen van begrippen en het maken van generaliseringën. Vanuit dit gezichtspunt zouden mentale proces-sen, die wij gewoonlijk met het begrip 'den-ken' associëren, zich niet beperken tot een stadium van 'hogere orde' in de mentale ont-wikkeling. Integendeel, denkvaardigheden spe-len een rol in een breed scala van leeractivitei-ten en zelfs in de elementaire niveaus van lezen, schrijven en rekenen. Het betekent dat stimulering van deze vaardigheden een inte-graal onderdeel zou moeten vormen van het schoolcurriculum.

In de basisschool wordt doorgaans geen expliciete aandacht besteed aan de stimulering van denkvaardigheden. Gewoonlijk wordt ervan uitgegaan dat deze vaardigheden zich spontaan ontwikkelen als 'bijproduct' van onderwijs in de reguliere schoolvakken zoals lezen, schrijven en rekenen. Mede omdat hier-over de nodige twijfels bestaan (Resnick, 1987) zijn vele voorstellen gedaan voor expli-ciete denkstimulering van kinderen (Baron & Sternberg, 1987) en is een variatie aan denk-trainingsprogramma's ontwikkeld (voor over-zichten raadplege men o.a. Coles & Robinson, 1991; Costa, 1991; Hamers & Overtoom, 1997; McGuinness & Nisbet, 1991; Nickerson, Per-kins & Smith, 1985; Nisbet & Davies, 1990).

In deze bijdrage doen we verslag van twee evaluatiestudies van het trainingsprogramma 'Denktraining für Kinder I' (Klauer, 1989) in

groep 3 van de basisschool met overwegend allochtone leerlingen. Het programma traint het proces van het inductief redeneren. Ten behoeve van de toepassing in reguliere school-klassen is het Duitse programma vertaald en bewerkt tot Inductief Redeneren I (De Koning & Hamers, 1995). Alle deelnemende kinderen vertonen een achterstand in de beheersing van het inductief redeneren.

2 Theorie en programma

Voor onderzoek naar het onderwijzen of stimu-leren van het denken is veel aandacht sedert de kritieken op Piagets theorie over de ontwikke-ling van het denken (Brown & Desforges, 1979; Donaldson, 1978). Deze kritieken spit-ten zich toe op de vermeende begrenzing in het vermogen tot redeneren en abstract denken van kinderen van verschillende leeftijden. Ook werd verondersteld dat kinderen met een ogen-schijnlijk beperkte capaciteit over meer 'poten-tieel' kunnen beschikken (Guthke & Wiedl, 1996; Hamers, Sijtsma & Ruijssenaars, 1993). Deze nieuwe inzichten hebben een aanzienlijke toename teweeggebracht van het aantal pogin-gen het denken te trainen en voorts te compen-seren en te remediëren voor gebleken achter-standen (Sternberg, 1984).

In de literatuur over denktraining worden gewoonlijk twee 'scholen' onderscheiden (Hamers & Overtoom, 1997). De eerste school stelt dat denkoperaties expliciet en onafhanke-lijk van het reguliere curriculum onderwezen kunnen worden (de 'skills' of de 'across-the-curriculum' benadering). In deze visie wordt uitgegaan van universele denkvaardigheden die een grote invloed hebben op de prestaties in diverse schoolvakken. Een voorwaarde voor een positief effect op schoolprestaties is dat er tijdens de training van de universele denkvaar-digheden een 'brug' geslagen wordt naar de schoolvakken. Een voorbeeld van dergelijk programma is Thinking Science (Adey, Shayer & Yates, 1995). Vele andere voorbeelden zijn te vinden in Nickerson et al. (1985) en in Hamers en Overtoom (1997). De tweede school stelt dat het leren denken het beste kan worden ingebed in de schoolvakken (de 'infu-sion' of de 'within-the-curriculum' benade-ring). In deze benadering worden denkvaardig-

	Kenmerken		Relaties	
	item-klasse	item-typen	item-klasse	item-typen
overeenkomsten	generalisatie	- groepen vormen - groepen afmaken - overeenkomsten vinden	relatie-herkenning	- series ordenen - series afmaken - simpele analogieën
verschillen	discriminatie	- niet-passende voorbeelden vinden	relatie-onderscheiding	- verstoorde relaties
overeenkomsten+ verschillen	kruisclassificatie	- 4,6,9 matrix-figuren	systeem-vorming	- matrix-figuren met complexe analogieën

Figuur 1. Taxonomie van opgavenklassen voor het zoeken en vergelijken van kenmerken en relaties

heden onderwezen binnen de schoolvakken. Deze werkwijze vergt ingrijpende veranderingen in de inhoud en de presentatie van de leerstof. Voorbeelden van deze benadering worden gevonden in het rekenen, het stellen en het begrijpend lezen (Halpern, 1992; Hamers & Overtoom, 1997; Resnick & Klopfer, 1989).

In dit onderzoek werden de processen van inductief redeneren systematisch getraind met het programma Inductief Redeneren 1 (De Koning & Hamers, 1995). In de intelligentiepsychologie wordt aan dit redeneerproces een centrale plaats toegekend (Spearman, 1923/1973; Horn & Cattell, 1966). Volgens deze auteurs wordt de algemene intelligentie (de *g*-factor) voor het grootste deel bepaald door het inductief redeneren (analogieën, classificaties, series en matrijzen). Ook recentere psychometrische studies (o.a. Snow, Kyllonen & Marshalek, 1984) tonen dat het inductief redeneren beschouwd kan worden als een centrale denkvaardigheid die ten grondslag ligt aan het maken van een groot aantal inhoudelijk verschillende taken. Sternberg (1977) en Klauer (1989) vestigen de aandacht op het inductief redeneren vanuit het perspectief van de informatieverwerkingspsychologie en tonen aan dat processen van inductief redeneren getraind kunnen worden.

Het programma Inductief Redeneren 1 wordt in de hier gepresenteerde studies gebruikt voor leerlingen in achterstandssituaties in de remediërende en compenserende zin. Het programma is een voorbeeld van de 'across-the-curriculum' benadering. De inhoud van de taken is betekenisvol en het is verder herkenbaar in dagelijkse situaties (Klauer, 1989).

2.1 Structuur en inhoud van het programma

Inductief redeneren wordt door Klauer (1989) gedefinieerd als het systematisch en analytisch vergelijken van objecten met het oog op het ontdekken van regelmatigheden in schijnbare wanorde en het ontdekken van onregelmatigheden in schijnbare orde. (On)regelmatigheden worden vastgesteld door te letten op kenmerken van objecten (bijvoorbeeld kleur) of relaties tussen objecten (bijvoorbeeld grootte). Zowel voor kenmerken als relaties is door Klauer (1989) een drietal opgavenklassen bepaald met respectievelijk het accent op het vinden van overeenkomsten, verschillen en combinaties van overeenkomsten en verschillen. De twee maal drie opgavenklassen beslaan het hele gebied van het inductief redeneren.

In Figuur 1 is te zien dat er drie itemtypen (generalisatie, discriminatie en kruisclassificatie) zijn voor het groeperen van objecten op basis van kenmerken en drie itemtypen (relatie-herkenning, relatie-onderscheiding en systeem-vorming) voor het seriëren van objecten op basis van hun onderlinge relaties. Gemeenschappelijk aan deze zes itemtypen of opgavenklassen is dat ze alle een beroep doen op de vaststelling van overeenkomsten of verschillen in kenmerken of relaties. Volgens Klauer (1989) kunnen kruisclassificatie en systeem-vorming worden beschouwd als eindstadia van het inductief redeneren. Verder kunnen opgaven uit de klasse kruisclassificatie alleen opgelost worden wanneer het generaliseren en discrimineren worden beheerst. Analooz hieraan kunnen opgaven uit de klasse systeem-vorming alleen opgelost worden wan-

neer het herkennen en onderscheiden van relaties worden beheerst.


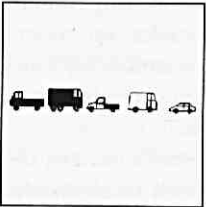
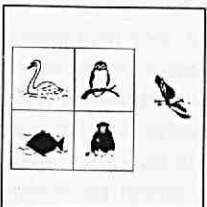
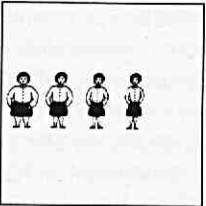

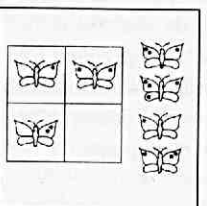
Het programma bestaat in totaal uit 120 opgaven (platen). Voor elke opgavenklasse zijn 20 platen opgenomen. Er zijn aldus 60 opgaven die betrekking hebben op het vergelijken van kenmerken van objecten (mensen, dieren, dingen, situaties) en er zijn 60 opgaven waarin het vergelijken van relaties tussen objecten (mensen, dieren, dingen, situaties) centraal staat. Iedere plaat stelt de leerlingen voor een probleem dat door het toepassen van inductieve redeneerprocessen opgelost kan worden (voor voorbeelden van opgaven, zie Figuur 2). De opgaven zijn verdeeld over 20 lessen zodat er per les zes opgaven worden behandeld. Elke les duurt een half uur tot drie kwartier en wordt door de leraar van de groep gegeven (voor verdere details, zie De Koning & Hamers, 1995).

2.2 Instructie

Centraal staat dat leerlingen onderwezen moeten worden in het vinden van overeenkomsten en verschillen in kenmerken of relaties. Essentieel is derhalve dat leerlingen leren zoeken naar kenmerken van objecten en relaties tussen objecten. Analoog aan het proces van inductief redeneren worden vier instructiestappen onderscheiden. Deze stappen vertonen overeenkomst met de door Sternberg (1985) onderscheiden

deelprocessen in het inductief redeneren: encoding, inference, respons en justification. In de eerste stap (Stap 1: 'zoeken') wordt bekeken wat de relevante kenmerken of relaties zijn. Vervolgens moeten er in Stap 2 ('vergelijken') overeenkomsten en verschillen in kenmerken of relaties van deze objecten worden gevonden door objecten met elkaar te vergelijken. Een belangrijk onderscheid tussen het vergelijken van kenmerken en relaties is dat bij kenmerken telkens twee objecten tegelijk bekeken worden, terwijl bij het vergelijken van twee relaties minimaal drie objecten betrokken zijn. Op basis van Stap 2 wordt in Stap 3 ('oplossen') het probleem opgelost. Ten slotte wordt in Stap 4 ('controleren') de oplossing gecontroleerd.

Het stappenplan kan volgens twee methoden worden uitgevoerd, de 'sturende' methode en de 'banende' methode (Van Parreren, 1981). De sturende methode wordt gekenmerkt door het systematisch en analytisch zoeken en vergelijken van kenmerken of relaties. Deze methode is systematisch daar het gaat om een uitputtend zoek- en vergelijkingsproces. Het analytische aspect drukt uit dat alleen objectkenmerken en relaties tussen objecten worden beschouwd. De banende methode wordt gekenmerkt door een meer globale wijze van bekijken en vergelijken van objecten. Hier is sprake van cycli van hypothesevorming betreffende de essentiële kenmerken of relaties, toetsing en

	Overeenkomsten	Verschillen	Overeenkomsten en verschillen
Kenmerken van objecten	Generalisatie: 	Discriminatie: 	Kruisclassificatie: 
Relaties tussen objecten	Relatie-herkenning: 	Relatie-onderscheiding: 	Systeem-constructie: 

Figuur 2. Voorbeelden van opgaven

besluitvorming. Bij een juiste hypothese is deze methode minder tijdrovend dan de sturende. De vraag of in het onderwijs de voorkeur gegeven moet worden aan sturing of aan baning van leerprocessen, is een punt van discussie (Hamers, De Koning & Pennings, 1995). In deze studie hebben we de keuze niet gemaakt maar zijn de beide methoden opgenomen als condities in de experimentele opzet.

Het onderwijs volgens de beide methoden wordt vormgegeven in drie fasen. De inleidende fase heeft tot doel om leerlingen aan de hand van eenvoudige opgaven, waarin met blokken wordt gemanipuleerd, een indruk te geven van de soort van opgaven en de wijze van oplossen. In de tweede fase, de verwerking, ligt het accent op het expliciet onderwijzen van het inductief redeneren met behulp van opgaven die door hun eenvoudige structuur de relevante eigenschappen van de opgavenklasse ongemaskeerd weergeven. Het is tevens in deze fase dat de eigenschappen van algemene abstracte begrippen als kenmerken en relaties aan de orde komen in het zogenaamde 'reflectie-deel' van de instructie. De toepassingsfase, tenslotte, staat in het teken van transfer naar inductie-opgaven in alledaagse situaties en in schoolsmateriaal. Het in verschillende situaties leren onderscheiden van de diverse opgavenklassen wordt bevorderd door een toenemende afwisseling in de opgavenklassen. Een ander doel binnen deze fase is het automatiseren van de vier stappen door regelmatige herhaling van gelijksoortige opgaven. Gaandeweg wordt de rol van de leraar minder prominent. De zelfstandigheid van de leerlingen wordt gestimuleerd door het opzettelijk wisselen van rollen (Palinscar & Brown, 1984), waardoor zij worden aangezet te reflecteren over hun eigen denkproces en dat van anderen.

De lessen verlopen in dialoogvorm met de hele groep. De leerlingen zitten in een halve cirkel zodat ze elkaar, de leraar en de voor de klas opgehangen plaat goed kunnen zien. Kern van de dialoog is dat de leraar de leerlingen weet aan te zetten tot redeneren. De leerlingen worden daartoe gestimuleerd doordat de leraar hen vragenderwijs begeleidt naar de oplossing van een opgave.

3 Het onderzoek

3.1 Opzet en populatie

Studie 1 was een formatieve evaluatie van de implementatie van het programma Inductief Redeneren I in groep 3 van de basisschool. Studie 2 was een summatieve evaluatie van hetzelfde programma. Voor de onderzoeken werden de volgende vragen geformuleerd: (a) Kan het programma in de natuurlijke setting – de klas – geïmplementeerd worden? (Studie 1); (b) Gaat de inductieve redeneervaardigheid van leerlingen die het programma volgen meer vooruit dan de inductieve redeneervaardigheid van leerlingen die het programma niet volgen? (Studie 1 en 2)

In Studie 1 werd het programma door leraren aangeboden aan twee experimentele groepen (bestaande klassen). In de eerste experimentele groep werd het programma volgens de sturende methode gegeven ($n=17$; gemiddelde leeftijd 84.0 maanden) en in de tweede experimentele groep volgens de banende methode ($n=13$; gemiddelde leeftijd 84.6 maanden). De controlegroep ($n=9$; gemiddelde leeftijd 85.6 maanden) ontving geen training maar nam deel aan de reguliere lessen. Er waren geen significante verschillen in leeftijd tussen de drie groepen. Alle groepen kregen voor-, na- en vervolgtests aangeboden. Aan de experimentele groepen werd het gehele programma aangeboden gedurende 20 lessen van drie kwartier, twee lessen per week. Het interval tussen de voor- en natests bedroeg 10 weken en tussen de natests en de vervolgtests $3\frac{1}{2}$ maand.

Studie 2 werd uitgevoerd in 18 groepen 3, zes experimentele en twaalf controlegroepen (bestaande klassen). In de experimentele groepen ($n=99$; gemiddelde leeftijd 85.5 maanden) werd een aangepaste versie van Inductief Redeneren 1 gebruikt, die werd aangeboden volgens de sturende methode. 'Aangepast' betekent dat op basis van de resultaten van Studie 1 de instructie van het programma op enkele details was gewijzigd. De controlegroepen ($n=232$; gemiddelde leeftijd 88.7 maanden) volgden het reguliere programma. Er waren geen significante verschillen in leeftijd tussen de groepen. In alle groepen werden voor- en natests afgenomen. Gelet op de vraag naar de duurzaamheid van de leereffecten werden de

natests 3 maanden na de beëindiging van het programma afgenomen.

De leerlingen van Studie 1 en 2 kwamen overwegend uit gezinnen met een sociaal-economische achterstand. Dit begrip werd gekwantificeerd in een zogenaamde leerling-score. Deze score werd bepaald op basis van het opleidingsniveau, de sociaal-economische status en de etnische achtergrond van de ouders: 1.25 voor kinderen uit de Nederlandse arbeidersklasse, 1.40 voor schipperskinderen in een internaat of in gezinsvervangend tehuis, en 1.90 voor kinderen waarvan tenminste een van beide ouders niet de Nederlandse nationaliteit had en tevens een laag opleidingsniveau had (Sijtsma, 1992).

Ten behoeve van de selectie van scholen en leerlingen werd eerst een pool van scholen samengesteld waarvan de meerderheid (95%) van de leerlingen voldeden aan het criterium van 1.90. Uit deze pool werden vervolgens scholen geselecteerd die voldeden aan de volgende additionele criteria: openbare scholen, geen vernieuwingscholen en niet betrokken bij fusies of andere veranderingsprocessen. In Studie 1 en 2 werd dezelfde selectieprocedure voor scholen gebruikt.

3.2 Instrumenten

In Studie 1 werd de inductieve redeneervaardigheid van de leerlingen vastgesteld met behulp van Ravens Standard Progressive Matrices (Raven-SPM, 1958). De SPM is in deze studie te beschouwen als een far-transfer-test. De opgaven in de test contrasteren met de trainingsopgaven van het programma in zoverre deze laatste betekenisvolle problemen voorstellen, ontleend aan het dagelijks leven of aan het schoolcurriculum. De test van Raven daarentegen bestaat uit betekenisloos figuratief materiaal.

Iedere opgave van de SPM doet een beroep op redeneerprocessen van het 'encoderen' en het 'afleiden en toepassen van gevonden relaties' (Spearman, 1923/1973; Sternberg, 1977). De opgaven vertonen sterke overeenkomst met de opgavenklassen kruisclassificatie en systeem-vorming in het trainingsprogramma. De analyses van de leereffecten hadden betrekking op de Raven-SPM totaalscores en op vijf opgaventypen uit deze test (Bereiter & Scardamalia, 1979). Deze auteurs bepaalden

deze vijf opgaventypen door voor de opgaven van de SPM vast te stellen wat het aantal 'mentale handelingen' is dat gelijktijdig geactiveerd moet worden voor correcte beantwoording. Deze kwantificering van taakkenmerken werd uitgedrukt in een 'Mental-Demand'. Samen met 'Mental-Capacity', dat de hoeveelheid mentale handelingen aangeeft die iemand gelijktijdig kan activeren, vormt Mental-Demand de basis voor het Mental-Construct concept van Pascual-Leone (1970). Items met een Mental-Demand van M1 meten vooral visuele waarneming; M2, M3, M4 en M5 opgaven worden geleidelijk complexer en meten in toenemende mate het eigenlijke inductief redeneren (Bereiter & Scardamalia, 1979).

Een inhoudelijke analyse (Hamers et al., 1995) van het aantal en de soort van essentiële kenmerken en relaties van de opgaven in Inductief Redeneren I liet zien dat er sprake is van een Mental-Demand variërend van M1 in de eenvoudige opgavenklasse tot M4 in de meest complexe opgaven aan het einde van het programma. In dit onderzoek werd ervan uitgegaan dat de leereffecten zich vooral zullen manifesteren op de M2, M3 en M4 opgaventypen van de SPM.

In Studie 2 werd de inductieve redeneervaardigheid van de leerlingen bepaald met behulp van de zogenaamde Programma-gerelateerde Test (Pg-T) (De Koning, Hamers & Sijtsma, 1996; De Koning, Hamers & Sijtsma, in voorbereiding). Deze test is te beschouwen als een near-transfer-test gebaseerd op de door Klauer (1989) gedefinieerde zes opgavenklassen in Figuur 1.

3.3 Procedure

De opzet van Studie 1 bood de mogelijkheid de leereffecten van beide experimentele groepen te bepalen. Daartoe dient ten eerste helderheid verschaft te worden over de vergelijkbaarheid van de verschillende groepen van de verschillende scholen en ten tweede over de wijze waarop leraren het programma uitvoerden. Ten behoeve van het eerste punt is getracht vergelijkbare scholen en vergelijkbare individuele leerlingen te selecteren (zie paragraaf 4.1). Verder is gebruikt gemaakt van een vragenlijst (De Koning & Hamers, 1994c) over enkele karakteristieken van de leerlingen zoals etnische achtergrond, verblijfsduur in Nederland en taalontwikkeling.

Ten behoeve van het tweede punt zijn de volgende maatregelen genomen. Ten eerste werd voor leerkrachten een handleiding geschreven met een uitgebreide beschrijving van de theoretische achtergrond, het doel, de inhoud en de structuur van Inductief Redeneren 1 (De Koning & Hamers, 1995). Ten tweede werden alle leraren vooraf uitvoerig geïnstrueerd over de uitvoering van het programma zoals beschreven in de handleiding. Ten derde werden alle lessen geobserveerd met behulp van een observatielijst (De Koning & Hamers, 1994a) en werden de lessen geëvalueerd met behulp van een vragenlijst voor leerkrachten (De Koning & Hamers, 1994b). De observatielijst, ingevuld door een onafhankelijk observator, had betrekking op het handelen van de leerkracht en de leerlingen tijdens elke les. De evaluatie-vragenlijst had betrekking op de duidelijkheid en de uitvoerbaarheid van de voorgeschreven instructies. De observatielijst en de evaluatie-vragenlijst overlappen gedeeltelijk zodat afwijkende interpretaties omtrent het verloop van de lessen aan het licht konden komen.

Voorafgaand aan de beantwoording van de onderzoeksvragen werd nagegaan of de onderzoeksgroep een achterstand had in inductief redeneren. Daartoe werden de ruwe scores op Raven-SPM op de voortest omgezet in percentielscores (Raven, Court & Raven, 1992). Op basis van deze scores werden drie niveaugroepen gedefinieerd. In de groep met zwak scorende leerlingen (score beneden P_{25}) viel 41% van de leerlingen en in de groep met gemiddeld scorende leerlingen (score tussen de P_{25} en P_{75}) viel 49% van de leerlingen. De groep goed scorende leerlingen (score boven de P_{75}) tenslotte

bestond dus uit slechts 10% van de leerlingen. Aangezien de verdeling van de leerlingen in dit experiment duidelijk niet conform de normering was, werd uitgegaan van een achterstand in het inductief redeneren van de onderhavige onderzoeksgroep.

4 Resultaten van Studie 1

4.1 Implementatie

Kan het programma Inductief Redeneren 1 in de natuurlijke setting van de schoolklas worden ingevoerd? De observatielijst en de evaluatie-vragenlijst toonden positieve reacties van leraren over de helderheid van de instructie en de uitvoerbaarheid van het programma. De leraren bleken soms geneigd gebruik te maken van beide instructiemethoden, de banende en de sturende, afhankelijk van de moeilijkheid van de opgaven. Voor de leraar in de groep met de sturende methode bleek het vooral in het begin van het programma lastig vast te houden aan de systematische, analytische oplossingsprocedure, aangezien de oplossing van deze eenvoudiger opgaven soms evident was. De leraar in de groep met de banende methode bleek daarentegen moeite te hebben met de toepassing bij moeilijker opgaven. In de groep met de sturende methode kostte het onderwijzen van de systematische, analytische oplossingsprocedure meer tijd waardoor de leraar niet altijd toekwam aan de reflectiefase. Verder bleek ook dat de beoogde rolwisseling in geen van beide groepen volledig tot stand kwam. Een mogelijke verklaring is dat rolwisseling geen geïntegreerd onderdeel vormde van de dagelijkse onderwijspraktijk en dat implementatie van

Tabel 1

Univariate covariantie analyses met de factoren groep en de Raven-SPM scores van de post- en follow-up-tests als onafhankelijke variabele

SPM Raven	Exp. Groep - Contr. Groep				Exp. Gr. 1 ^a - Exp. Gr. 2 ^a			
	Posttest		Follow-up-test		Posttest		Follow-up-test	
	F	P	F	P	F	P	F	P
Totaal score	7.64	.01	7.27	.01	2.07	.08	1.23	.14
M2	4.15	.03	3.86	.03	1.35	.13	.17	.34
M3	4.19	.02	5.90	.01	1.43	.12	3.93	.03
M4	6.45	.01	9.02	.00	1.72	.10	3.44	.04

^a Exp. Gr. 1: Experimentele groep, sturende methode. Exp. Gr. 2: Experimentele groep, banende methode

Tabel 2
Effectgrootten (*d*) van de Raven-SPM totaal scores en scores van de verschillende typen

SPM Raven	Exp. Groepen 1 + 2		Exp. Groep 1 ^a		Exp. Groep 2 ^a	
	Posttest - Pretest	Follow-up test - Pretest	Posttest - Pretest	Follow-up test - Pretest	Posttest - Pretest	Follow-up test - Pretest
Totaal score	.79**	.83**				
M2	.28**	.21**				
M3	1.57**	1.59**	1.48**	1.59**	1.41*	1.26*
M4	1.43**	1.09**	1.02**	1.20**	0.93**	0.98**

^a Exp. Groep 1: Experimentele groep, sturende methode. Exp. Groep 2: Experimentele groep, banende methode

**⁾ $p \leq 0.05$

⁾ $p \leq 0.10$

$d = (M_{EG} - M_{CG}) / s_p$

een nieuwe werkvorm in combinatie met een nieuw programma veel gevraagd was van leraren.

4.2 Leereffecten

Het experiment vond plaats in bestaande groepen leerlingen (schoolklassen) zodat van randomisatie van leerlingen over de experimentele en controle groepen geen sprake was. Voor effectmeting werden de voortest scores van de Raven-SPM daarom als covariaat opgenomen in variantieanalyses¹. Teneinde systematische verschillen tussen de groepen te elimineren werden twee covariaten toegevoegd: het aantal jaren dat leerlingen reeds in Nederland verbleven en het aantal jaren dat ze onderwijs in Nederland volgden.

In Tabel 1 is te zien dat, uitgaande van de totaalscores op de natest en de vervolgtest, de experimentele groepen significant verschilden van de controlegroep ($p=0.01$). Op geen van de tests verschilden de experimentele groepen significant van elkaar (respectievelijk $p=0.08$ en $p=0.14$). De covariaten hadden elk een significante invloed op de natestscore en vervolgtest score van de Raven-SPM. Er waren op de nameting en de vervolgmeting geen significante hoofdeffecten van sekse noch significante interactie-effecten van sekse met de groepen.

Uit univariate covariantieanalyses op de Mental-Demand opgaventypen bleek dat de verschillen op de natests en vervolgtests tussen de experimentele groepen en de controlegroep zich toespitsen op de M2-, M3- en M4-opgaven. Testresultaten op de M1- en M5-opgaven zijn daarom niet in deze beschrijving opgenomen. Kolmogorov-Smirnov toetsen op

de natestscores en de vervolgtest scores van de M2-, M3- en M4-opgaven toonden dat er in alle gevallen sprake was van normaliteit behalve bij de M4 natest scores. Op de natest verschilden de experimentele groepen significant van de controlegroep op elk van de opgaventypen (M2: $p=0.03$ / M3: $p=0.02$ / M4: $p=0.01$). Vergelijkbare resultaten werden gevonden voor de vervolgtest (M2: $p=0.03$ / M3: $p=0.01$ / M4: $p=0.00$). De groep met de banende methode had op de M3- en de M4-opgaven van de SPM vervolgtest een significant lagere score (respectievelijk $p=0.03$ en $p=0.04$) dan de groep met de sturende methode. De groepen verschilden niet op de SPM natest.

De (co)variantieanalyses toonden welke groepen significant van elkaar verschillen op de natests en de vervolgtests. Significante verschillen in combinatie met kleine effectgrootten kunnen de relevantie van de gevonden verschillen in twijfel trekken (Hager, in voorbereiding). In Tabel 2 zijn de effectgrootten weergegeven op de Raven-SPM totaalscore en de Mental-Demand typen M2, M3 en M4. De effectgrootten voor de typen M1 en M5 zijn niet opgenomen aangezien de groepen op deze variabelen geen significante verschillen vertoonden.

In Tabel 2 is te zien dat de hoogste effectgrootten werden gevonden voor de M3 en M4 opgaventypen. De effecten op de M2-opgaven waren relatief klein². Verder is te zien dat de effecten op de vervolgtest die van de natest meestal evenaarden. Vanwege de significante verschillen tussen de experimentele groepen op de M3- en M4-opgaven, werden de effectgrootten ook berekend voor deze groepen

afzonderlijk. Tabel 2 laat zien dat de effectgrootten van de groep met de sturende conditie in alle gevallen licht hoger zijn dan in de groep met de banende conditie.

Indien beide experimentele groepen tezamen worden genomen, kan de onderzoeksvraag of de inductieve redeneervaardigheid van leerlingen die het programma volgden meer vooruit gaan dan de inductieve redeneervaardigheid van leerlingen die het programma niet volgden, positief worden beantwoord. In zoverre we in staat waren tot controle van de invloed van de variabelen school en leerkracht, kan op basis van de gevonden resultaten worden geconcludeerd dat de sturende methode betere effecten sorteert dan de banende. Daarom werd in Studie 2 het programma met de directe instructiemethode gebruikt.

5 Resultaten van Studie 2

Kan de inductieve redeneervaardigheid van leerlingen die het programma volgen, verbeterd worden in vergelijking met leerlingen die het reguliere schoolprogramma volgen? De Scheffé toets met factor groep en de totaalscore op de voortest als afhankelijke variabele vertoonde geen significante verschillen tussen de combinatie van de twee groepen ($p=.05$). Om die reden werden de zes experimentele groepen samengevoegd, evenals de twaalf controlegroepen. De toetsresultaten van univariate analyses met factor groep en met als afhankelijke variabele de totaalscore op de voor- en de natest zijn, respectievelijk, $F=.81$; $p=.27$ en $F=27.04$; $p=.00$. Dit betekent dat de natestscores van de leerlingen van de experimentele groepen significant hoger zijn dan de scores van de controlegroepen. Er werden geen significante hoofd- of interactie-effecten gevonden van sekse. De (gecorrigeerde) effectgrootte van de totaalscore was 1.04.

6 Conclusies en discussie

De vragen in deze studies waren: (a) Is het mogelijk denktrainingsprogramma's in de natuurlijke setting te implementeren en (b) kunnen we kinderen inductief leren denken of kunnen we deze vorm van denken stimuleren?

De resultaten geven aan dat leraren in staat zijn om te gaan met de didactische vereisten die gesteld worden aan de implementatie van dit soort programma's. Hiermee is de eerste vraag positief beantwoord.

Het antwoord op de tweede vraag is ook positief. De onderzoeksresultaten tonen aan dat leerlingen die het programma Inductief Redeneren I hebben doorlopen zowel direct na afloop van het programma als na ruim drie maanden significant hogere scores behalen op de Raven-SPM in vergelijking met controlegroepen zonder training. Voorts zijn de effectgrootten in vergelijking met die van trainingsexperimenten van Klauer (1989) met dit programma overtuigend, zeker wanneer in ogeschouw wordt genomen dat de trainingen zijn uitgevoerd in kleine groepen. Een mogelijke verklaring hiervoor kan zijn dat leraren in onderhavig onderzoek het programma aan de hand van een gedetailleerder handleiding hebben kunnen uitvoeren. Het antwoord op de vraag naar de mogelijkheid tot het onderwijzen van inductief redeneren blijft ondanks deze positieve resultaten met twijfels omgeven, maar enig optimisme is gerechtvaardigd.

In Studie 1 werd aangetoond dat instructie volgens de sturende methode de beste leerresultaten geeft. Gal'perin (1967) heeft vooral de sturende methode verdedigd, de werking ervan in experimenteel onderzoek aangetoond, en er een theoretische analyse van gegeven. Het begrip 'volledige oriënteringsbasis' neemt in deze theorie een centrale plaats in. Hij verstaat eronder het geheel van kennis en van actuele gegevens waarop iemand zich bij de uitvoering van een taak baseert. Een verklaring voor de gevonden verschillen in deze studie is wellicht de behoefte van leerlingen in achterstandssituaties aan een volledige oriënteringsbasis, althans in de beginfase van het oplossingsproces. Dit gegeven heeft ons ertoe gebracht de dichotome keuze voor de banende of de sturende methode te verlaten en de vraag te stellen onder welke omstandigheden en voor welke leerlingen de banende, dan wel de sturende methode de voorkeur verdient.

In deze studies was 95% van de leerlingen van etnische afkomst, i.c. Turkse en Marokkaanse kinderen. We gaan ervan uit dat vooral de sociaal-economische verschillen met de dominante cultuur de belangrijkste verklaring

geven voor de gevonden leerachterstanden bij deze leerlingen (zie ook Van Langen & Jungbluth, 1990). Deze verschillen hebben hun uitwerking op de ontwikkeling van algemene vaardigheden (bijvoorbeeld inductief redeneren) en de schoolprestaties van kinderen. Met het hier gebruikte programma wordt geprobeerd een deel van die relatieve achterstand te compenseren.

Een vraag is of in het onderwijs een behoefte bestaat aan 'across-the-curriculum' programma's. Het antwoord bepaalt voor een deel hoe denken in het algemeen en inductief redeneren in het bijzonder onderwezen moet worden. Indien denken wordt onderwezen met behulp van een 'across-the-curriculum' programma, dan zijn denkvaardigheden de leerdoelen. Indien het denken in de context van schoolvakken wordt geleerd, dan zijn de inhoud van de schoolvakken de leerdoelen. Voorstanders van de eerste methode verwachten dat leerlingen overbelast zullen raken wanneer zij naast denkvaardigheden de inhoud van schoolvakken moeten leren (Resnick, 1987). Voorstanders van de tweede methode zijn van mening dat de meeste denkvaardigheden inhoud-specifiek zijn en dat het moeilijk zal zijn transfer te bewerkstelligen naar andere vakken. We zijn van mening dat in het algemeen genomen 'within-the-curriculum' programma's te prefereren zijn tenzij leerlingen grote moeite hebben met de inhoud van de schoolvakken. Het bewerkstelligen van transfer verdient dan wel de volledige aandacht.

In dit onderzoek is eveneens aandacht voor transfer. De resultaten op de natests en de vervolgtests tonen aan dat transfer optreedt. Een volgende stap is transfer te onderzoeken naar andere gebieden van het schoolcurriculum, zoals begrijpend lezen. In een ander onderzoeksproject (De Koning & Hamers, in voorbereiding) is dit onderwerp van onderzoek geweest.

Noten

1 Aan de assumpties van homoscedasticiteit, homogene regressies in geval van covariaten en de normaalverdeling van de scores was voldaan.

2 We hebben voor het beschrijven van effectgrootten Cohens (1988) vuistregel overgenomen: $ES \geq$

$.20$ = klein effect; $ES = .50$ -. 79 = gemiddeld effect; $ES \geq .80$ = groot effect.

Literatuur

- Adey, P.S., Shayer, M., & Yates, C. (1995). *Thinking sciences*. London: Thomas Nelson and Sons.
- Baron, J.B., & Sternberg, R.J. (Eds.) (1987). *Teaching thinking skills theory and practice*. New York: Freeman.
- Bereiter, C., & Scardamalia, M. (1979). Pascual-Leone's M-construct as a link between cognitive development and psychometric concepts of intelligence. *Intelligence*, 3, 41-63.
- Brown, D., & Desforges, C. (1979). *Piaget's theory. A psychological critique*. London: Routledge.
- Cohen, J. (1988). *Statistical power analysis for the behavior sciences*. Hillsdale, NJ.: Erlbaum.
- Coles, M.J., & Robinson, W.D. (1991). *Teaching thinking. A survey of programmes in education*. London: Bristol Classical Press.
- Costa, A.L. (Ed.) (1991). *Developing minds: Programs for teaching thinking*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Dewey, J. (1991). *How we think*. Buffalo, NY: Prometheus Books (Original work published in 1911).
- Donaldson, M. (1978). *Childrens' minds*. London: Fontana.
- Gal'perin, P.J. (1967). *Zu Grundfragen der Psychologie*. Berlin: Volk und Wissen.
- Guthke, J., & Wiedl, K.H. (1996). *Dynamisches Testen*. Göttingen: Hogrefe.
- Hager, W. (in voorbereiding). Evaluating the effectiveness of cognitive programs: Some methodological issues. In J.H.M. Hamers, J.E.H. van Luit & B. Csapó (Eds.), *Teaching thinking and thinking skills*. Lisse: Swets & Zeitlinger.
- Halpern, D.F. (Ed.) (1992). *Enhancing thinking skills in the sciences and mathematics*. Hillsdale, NJ.: Erlbaum.
- Hamers, J.H.M., Koning, E. de, & Pennings, A.H. (1995). *Stimulering van denkvaardigheden van allochtone leerlingen*. Utrecht: Universiteit Utrecht ISOR.
- Hamers, J.H.M., Luit, J.E.H. van, & Csapó, B. (Eds.) (in voorbereiding). *Teaching thinking and thinking skills*. Lisse: Swets & Zeitlinger.

- Hamers, J.H.M., & Overtoom, M. (Eds.) (1997). *Teaching thinking in Europe. Inventory of European programmes*. Utrecht: SARDES.
- Hamers, J.H.M., Sijtsma, K., & Ruijsseenaars, A.J.J.M. (Eds.) (1993). *Learning potential assessment. Theoretical, practical and methodological issues*. Amsterdam: Swets & Zeitlinger.
- Horn, J.L., & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized ability intelligence. *Journal of Educational Psychology*, 57, 253-270.
- Klauer, K.J. (1989). *Denktraining für Kinder 1. Ein Program zur intellektuellen Förderung*. Göttingen: Hogrefe Verlag.
- Klauer, K.J., & Phye, G.D. (1995). *Cognitive training for children. A developmental program of inductive reasoning and problem solving*. Seattle: Hogrefe & Huber Publishers.
- Kohnstamm, Ph. (1930). Over de 'vormende waarde' der leervakken. *Pedagogische Studiën*, 11, 129-150.
- Koning, E. de, & Hamers, J.H.M. (1994a). *Observatieleerkrachten*. Utrecht: Universiteit Utrecht ISOR.
- Koning, E. de, & Hamers, J.H.M. (1994b). *Vragenlijst voor de instructiestijl van leraren*. Utrecht: Universiteit Utrecht ISOR.
- Koning, E. de, & Hamers, J.H.M. (1994c). *Vragenlijst leerlingen*. Utrecht: Utrecht Universiteit ISOR.
- Koning, E. de, & Hamers, J.H.M. (1995). *Inductief redeneren 1*. Utrecht: Universiteit Utrecht ISOR.
- Koning, E. de, & Hamers, J.H.M. (in voorbereiding). Theoretical background and educational implications of inductive reasoning. In J.H.M. Hamers, J.E.H. van Luit & B. Csapó (Eds.), *Teaching thinking and thinking skills*. Lisse: Swets & Zeitlinger.
- Koning, E. de, Hamers, J.H.M., & Sijtsma, K. (1996). *Test voor inductief redeneren. Test en handleiding*. Den Haag: HCO.
- Koning, E. de, Hamers, J.H.M., & Sijtsma, K. (in voorbereiding). Designing a test for inductive reasoning: Integrative use of several data analyses methods.
- Langen, A. van, & Jüngbluth, P. (1990). *Onderwijskansen van migranten*. Lisse: Swets & Zeitlinger.
- McGuinness, C., & Nisbet, J. (1991). Teaching thinking in Europe. *British Journal of Educational Psychology*, 61, 174-186.
- Nickerson, R.S., Perkins, D.N., & Smith, E.E. (1985). *The teaching of thinking*. Hillsdale, NJ.: Erlbaum.
- Nisbet, J., & Davies, P. (1990). The curriculum redefined: Learning to think - thinking to learn. *Research Papers in Education*, 5, 49-72.
- Palincsar, A.S., & Brown, A.L. (1984). Reciprocal teaching of comprehension monitoring activities. *Cognition and Instruction*, 2, 117-175.
- Parreren, C.F. van (1981). *Onderwijsproceskunde*. Groningen: Wolters-Noordhoff.
- Pascual-Leone, J.A. (1970). A mathematical model for the transition rule in Piaget's development stages. *Acta Psychologica*, 63, 301-345.
- Raven, J.C. (1958). *Standard Progressive Matrices*. London: Lewis.
- Raven, J.C., Court, J.H., & Raven, J. (1992). *Standard Progressive Matrices. Raven Manual: Section 3*. Oxford: Psychologists Press.
- Resnick, L.B. (1987). *Education and learning to think*. Washington: National Academic Press.
- Resnick, L.B., & Klopfer, L.E. (1989). *Toward the thinking curriculum. Current cognitive research*. Alexandria, VA.: Association for Supervision and Curriculum Development.
- Selz, O. (1935). Versuche zur Hebung des Intelligenzniveaus. *Zeitschrift für Psychologie*, 134, 236-301.
- Snow, R.E., Kyllonen, P.C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence*. Hillsdale, NJ.: Erlbaum.
- Spearman, C. (1923/1973). *The nature of intelligence and the principles of cognition*. New York: Arno Press.
- Sternberg, R.J. (1977). *Intelligence, information processing and analogical reasoning*. Hillsdale, NJ.: Erlbaum.
- Sternberg, R.J. (1984). How can we teach intelligence? *Educational Leadership*, 42, 38-48.
- Sternberg, R.J. (1985). *Beyond IQ. A triarchic theory of human intelligence*. New York: Cambridge University Press.
- Sijtsma, J. (1992). *Balans van het taalonderwijs halverwege de basisschool*. Arnhem: CITO.

Manuscript aanvaard 11-6-1998

Auteurs

J.H.M. Hamers is universitair hoofddocent Pedagogiek, Universiteit Utrecht.

E. de Koning is assistent in opleiding, Pedagogiek, Universiteit Utrecht.

K. Sijtsma is hoogleraar Methoden en Technieken, Katholieke Universiteit Brabant.

Correspondentieadres: J.H.M. Hamers, Universiteit Utrecht, Vakgroep Pedagogiek, Postbus 80.140, 3508 TC Utrecht

Abstract

Teaching inductive reasoning in the first school grade

J.H.M. Hamers, E. de Koning & K. Sijtsma. Pedagogische Studiën, 1998, 75, 250-261.

This article presents the results of two evaluation studies with respect to a programme which aims at enhancing the inductive reasoning ability of third

grade students. The programme is a classroom version of the German programme 'Denktraining für Kinder I' (Cognitive training for children) (Klauer, 1989). In the first formative evaluation study, two experimental groups with 30 students and one control group with 9 students were involved. The outcomes of the observations during the lessons and of the teacher reports showed that the teachers were able to implement the programme in the third grade of primary schools. Both experimental groups significantly outperformed the control group on a posttest immediately after the programme and on a follow-up test 3½ months later. Further analyses of the data revealed tentative evidence of the superiority of a direct teaching method for these students. In the second, summative evaluation study, the same programme was applied once again in a much larger sample of third grade students. The programme instructions were slightly changed following the results of Study 1. The experimental groups scored significantly higher on a posttest 3 months after completion of the programme.