CONTEMPORARY EDUCATIONAL PSYCHOLOGY **23,** 132–148 (1998)
ARTICLE NO. EP980966

Inductive Reasoning in Third Grade: Intervention Promises and Constraints

J. H. M. Hamers and E. de Koning

Utrecht University, The Netherlands

and

K. Sijtsma

Tilburg University, The Netherlands

The results of two evaluation studies with respect to a programme for enhancing inductive reasoning ability of third grade students are presented. The programme is a classroom version of the German programme 'Denktraining für Kinder 1' (Cognitive training for children; Klauer, 1989). In the first formative evaluation study, two experimental groups with 30 students in total and one control group with 9 students were involved. Observations during the lessons, and teachers' reports showed that teachers were able to implement the programme. 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. In the second summative evaluation study, the same programme was applied to a larger sample (experimental groups: n = 99 in total; and control groups: n = 232 in total) of third grade students. On the basis of Study 1, the programme instructions were slightly changed. The experimental groups scored significantly higher on a posttest three months after completion of the programme. © 1998 Academic Press

INTRODUCTION

The systematic teaching of general thinking skills is considered important by many authors (see, for example, Klauer & Phye, 1995; Resnick, 1987). According to Resnick (1987), the need for teaching these skills is created by the rapid changes taking place in today's society. Knowledge and information are becoming ever more complex and soon may become dated. Children therefore have to be equipped with the skills of evaluating choices, and identifying and solving problems using logical reasoning. Thus, it is

This research was financed by the C. Wright Mills Fund of the Faculty of Social Sciences of Utrecht University.

Address correspondence and reprint requests to J. H. M. Hamers, Department of Education, Utrecht University, P.O. Box 80.140, 3508 TC Utrecht, The Netherlands.

0361-476X/98 $25.00
Copyright © 1998 by Academic Press
All rights of reproduction in any form reserved.

not enough to have a considerable amount of knowledge at one's disposal (declarative knowledge), but the questions of how to acquire knowledge, and how to apply this knowledge are also important (procedural knowledge).

It is also claimed (e.g., Halpern, 1992) that a limited command of thinking skills is one of the reasons for falling behind in school. This can be seen in mathematics, reading, and writing, where all sorts of activities come to the fore in which thinking skills play a central role. Examples are the ability to describe and to compare objects, to group objects, to associate one thing with another, to form concepts, and to generalize. Thus, mental processes which are normally associated with the concept of 'thinking' are not limited to some kind of 'higher order' of mental development. On the contrary, thinking processes play a role in a broad range of learning activities in school. This means that these thinking processes should form an integral part of the school curriculum.

In primary school, the stimulation of thinking skills is not pursued explicitly. It is usually assumed that these skills develop spontaneously as a 'by-product' of teaching regular school subjects. However, new theories about the learning potential of students with apparently limited capacity (Guthke & Wiedl, 1995; Hamers, Sijtsma, & Ruijssenaars, 1993), and criticism on the supposed delimitation of the students' capacity for abstract thinking in the Piagetian sense (Brown & Desforges, 1979; Donaldson, 1978), have brought about an increase in the number of attempts to remedy evident backwardness in thinking or to train thinking beyond the supposed limits (Baron & Sternberg, 1987). As a consequence, a range of programmes for training of thinking have been developed (e.g., Costa, 1991; Hamers & Overtoom, 1997; Nickerson, Smith, & Perkins, 1985).

We report on two evaluation studies concerning teaching of a thinking programme (Klauer, 1989), called 'Denktraining für Kinder 1' (Cognitive training for children). This programme trains the processes of inductive reasoning. For application in the natural setting of the classroom, the original programme was translated and adapted by De Koning and Hamers (1995). The studies were carried out in the third grade of primary schools. All children showed substantial backwardness in main school subjects and in general thinking skills. The majority of the children came from ethnic minorities.

THEORY AND PROGRAMME

Two 'schools' on teaching of thinking can be distinguished. Researchers of the first school believe that thinking skills can be taught explicitly, independent of the regular curriculum (the 'skills' or 'across-the-curriculum' approach). It is assumed that certain more or less universal thinking skills exist which can be generalized and applied to school subjects. Examples of programmes are Cognitive Research Trust programme (de Bono, 1983), Cognitive Training Program for Children (Klauer & Phye, 1995), and Instrumental

Enrichment programme (Feuerstein, 1980). Researchers from the second school believe that learning of thinking skills should be embedded in the school subjects (the 'infusion' or 'within-the-curriculum' approach). Examples of programmes can be found in mathematics, in comprehensive reading, and in text production (see, Halpern, 1992; Resnick & Klopfer, 1989).

In the present studies, the processes of inductive reasoning were trained using Klauer's (1989) adapted programme Inductive Reasoning 1 (De Koning & Hamers, 1995). The kind of reasoning trained by this programme can be considered as a central thinking process of intelligence. Support for this idea can be found in the early factor analytic studies by Spearman (1923) and Horn and Cattell (1966). In their opinion, inductive reasoning tasks (analogies, classifications, series completion, and matrices) determine for the most part the *g*-factor loadings. Also, recent studies (e.g., Snow, Kyllonen, & Marshalek, 1984) indicated that inductive reasoning can be seen as 'a domain-general thinking skill', which influences a broad variety of tasks in several content areas. Sternberg (1977) used the information processing view to study inductive reasoning, and claimed that processes underlying inductive reasoning can be trained. Klauer (1989) assumed the existence of formal content-free paradigms of inductive reasoning, which are related and which can be trained. Mastering these paradigms means the ability to apply them to a broad variety of tasks. In the context of education, Curtis (1988) showed that the acquisition of inductive reasoning is applicable even in classroom instruction.

Inductive Reasoning 1 (De Koning & Hamers, 1995) was used here with children identified as 'at risk' (children with low social-economic-status backgrounds) in the remedial and compensating sense. The programme can be characterized as an 'across-the-curriculum' programme because it stimulates the general thinking processes of inductive reasoning and problem solving. The content of the tasks, however, is meaningful and is recognizable in daily situations (see also Klauer & Phye, 1995).

Programme Structure and Content

Klauer (1989) defined inductive reasoning as systematic and analytic comparison aimed at discovering regularity in apparent chaos and irregularity in apparent order. Regularities and irregularities are recognized through the attributes of objects (for example, colour) or the relations among objects (for example, size). Objects in items measuring inductive reasoning are people, animals, objects, and situations. For both attributes and relations, Klauer (1989) determined three classes of tasks characterized by finding similarities, dissimilarities, or a combination of both, respectively. This yielded six classes in total, which constitute the whole area of inductive reasoning (Fig. 1).

Figure 1 shows three classes of tasks for grouping objects on the basis of

	Attributes		Relations	
	item-class	item-types	item-class	item-types
similarity	generalization	- class formation - class expansion - finding common attributes	recognition of relations	- order series - series completion - simple analogies
dissimilarity	discrimination	- identifying irregularities	differentiation of relations	- disrupted series
dissimilarity and similarity	cross-classification	- 4,6,9 matrix-figures	system-construction	- matrix-figures with complex analogies

FIG. 1. Taxonomy of classes of tasks for search and comparison of attributes and relations.

attributes (generalization, discrimination, and cross-classification) and three classes of tasks for seriation of objects based on their mutual relations (recognition of relationships, differentiation of relationships, and system-construction). Common to these six classes is the requirement of an ability to determine similarities and/or dissimilarities among attributes or relations. Mastering cross-classification and system-construction is considered to be the final stage of inductive reasoning (Klauer, 1989).

The programme consists of 120 tasks (problem plates). Each class of tasks consists of 20 plates. Each plate presents the student with a problem which can be solved by application of inductive reasoning processes (Fig. 2 exhibits examples of problem plates). The tasks are divided across 20 lessons, and six tasks are dealt with in each lesson. There are two lessons per week, and each lesson lasts 45 minutes. The programme is presented by teachers (De Koning & Hamers, 1995).

Programme Instruction

Students have to be instructed in finding similarities and/or dissimilarities in attributes and relations. This way, they learn how to seek and compare attributes of objects or relations between objects. Analogous to the process of inductive reasoning, four instruction steps are differentiated: (1) ‘searching’ for the relevant attributes or relations; 2) ‘comparing’ similarities or dissimilarities in attributes or relations; (3) ‘solving’ the problem on the basis of comparison; and (4) ‘controling’ the solution. There are similarities between these steps and the processes in inductive reasoning: encoding, inference, answering, and justification (Sternberg, 1985).

The steps can be performed by two procedures. The first procedure (see Fig. 3) is characterized by systematic and analytic searching and comparison of attributes or relations. The procedure is systematic because it is exhaustive. The procedure is analytical because only object attributes and relations between objects are considered. Instruction of this inductive reasoning pro-

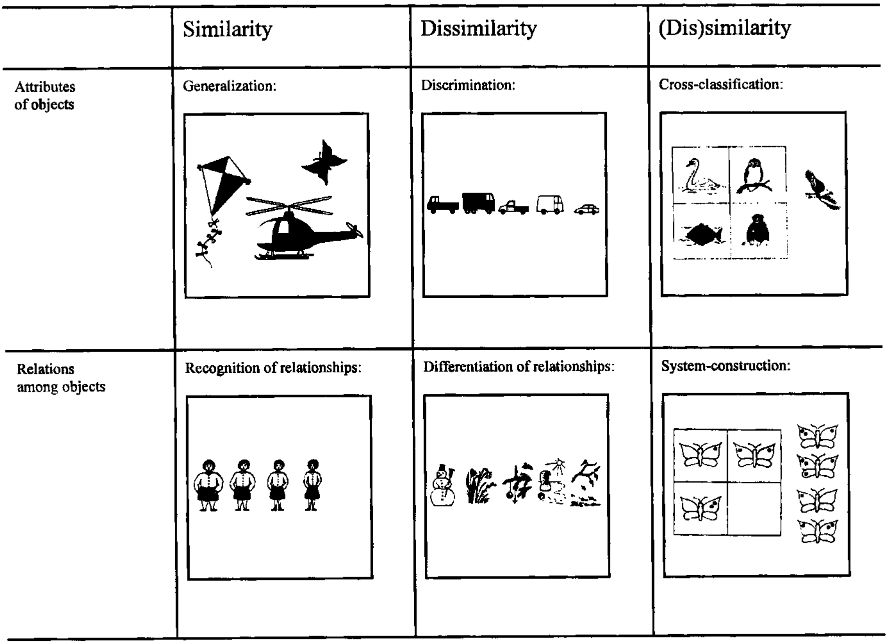

FIG. 2. Examples of problem plates.

cess is realized by means of a ‘direct teaching’ method (Van Parreren, 1981), known as ‘guided discovery’.

The second procedure (see Fig. 4) consists of a more global examining and comparison of objects and relations. Hypotheses with respect to the essential attributes or relations, and testing and decision-making, are formed in cycles. If a hypothesis is correct, this procedure is less time-consuming than the first procedure. The second procedure is taught by the ‘indirect teaching’ method (Van Parreren, 1981). The differences between the two procedures is demonstrated in the solution part in the Fig. 3 and 4: in Fig. 3 Step 1 to Step 4, and in Fig. 4 Resolve and Evaluation, respectively. Furthermore, the procedure differs in the reflection phase; see Fig. 4 where the so-called ‘Description of the solution process’ is added. This study also was concerned with the applicability and the effectiveness of both instruction methods.

Teaching both procedures consists of three phases. The introductory phase uses simple tasks to give students an impression of the tasks and the way to solve them. This phase precedes both the direct and the indirect teaching methods. In the second phase (see Fig. 3 and 4), instruction and task characteristics according to the two procedures are given. In the ‘reflection’ part of the second phase, the properties of general abstract concepts such as attri-

PROBLEM TABLE 4: GENERALIZATION TASK

Aim:	discovery of an attribute (looking for regularity)
Question:	what have these objects in common
Essential attribute:	flying in the air
Problem table:	* table with a helicopter, butterfly, kite * table 1 and table 3
Instruction:	
Recognizing problem type	*Again we are going to form groups, but we are not using blocks this time. Look at this table. First we name the objects on the table. This is a ...* (point to all the objects and give them a name).
Question	*These three objects fit together. They look more or less like each other. We should find out ...* why they fit together.
Step 1: search	*What do we have to do? All right, we are going to describe each object, just as we did with the blocks. What can you tell me about this ...* (point to a object) *and this ...*
Step 2: compare	*What do we have to do now? Yes we have to compare the objects. We look now for similarities* (point to all three the objects).
Step 3: resolve	*Well, the question was: why is it that these object belong to one group? Yes, that's right, because they all can fly. And what must be the other reason? All right, they also belong together because they have a tail.*
Step 4: check	*Now we must check our solution. The helicopter .....* flies in the air, *the butterfly ....* flies in the air *and the kite ...* flies in the air. *Yes, it's okay, they belong together in one group* (same for the tail).
Reflection	Describing in general terms: the concept attribute *Now I want to tell you something. We saw that all these objects (point) are a little bit the same: they all fly in the air. We say: these objects have the same attribute. The attribute of the kite is that it flies in the air. The attribute of the butterfly is that it flies in the air. The attribute of the helicopter is that it flies in the air. An attribute tells something about an object. For instance, whether it can fly, or what colour is has.* Compare with blocks of table 1 and 3 *In case of the blocks, we looked at the colour. So we grouped them by colour* (point out). *Who knows another attribute of the block? Very good, the form of the blocks.* Description of the problem type *Objects with the same attribute fit together in one group. Objects that do not share the same attribite do not fit together in one group. Does a car fit in this group? No, a car doesn't have the same attribute: it can not fly. So a car does not belong to this group*

FIG. 3. Part of teacher's manual of the direct teaching method with respect to the inductive reasoning problem generalization (see Fig. 2).

butes and relations are dealt with. The application phase follows after both procedures and deals with transfer to induction tasks in realistic situations and in school material. Students are trained to differentiate between the various classes of tasks in different situations. This is stimulated by an increasing alternation of tasks from different classes. The training also aims at automation of the four steps by regularly repeating similar tasks, while the role of the teacher becomes less prominent. Students are expected to be able to solve the problems themselves. To stimulate independence, in this phase students take over the role of the teacher (reciprocal teaching; Palinscar & Brown, 1984). As a result, they are forced to reflect on their own thinking in order to gain insight in the thought steps.

PROBLEM TABLE 4: GENERALIZATION TASK

Aim:	discovery of an attribute (looking for regularity)
Question:	what have these objects in common
Essential attribute:	flying in the air
Problem table:	* table with a helicopter, butterfly, kite * table 1 and table 3
Instruction:	
Recognising problem type	*Again we are going to form groups, but we are not using blocks this time.* *Look at this table. First we name the objects on the table. This is a ...* (point to all the objects and give them a name).
Question	*These three objects fit together. They look more or less like each other. We should find out ...* why they fit together.
Resolve	*So, what do we have to do here? Yes we have to look for similarities. Who has a suggestion? Right, the first suggestion is: All the objects can fly.* Turn to the evaluation part after each suggestion.
Evaluation	*Okay we have a suggestion. Who knows what to do next. Right we have to check this suggestion. Show how you can do that. Right, I repeat what you did: the kite flies in the air, the helicopter flies in the air and the butterfly flies in the air* (point to every object seperately now). *So, it is okay now, they belong together in one group. Who knows another similarity* (tail)? If the answer in not correct: *No, the objects are not similar with respect to this.* Turn back to the 'resolve part'.
Reflection	Describing in general terms: the concept attribute *Now I want to tell you something. We saw that all these objects* (point) *are a little bit the same: they all fly in the air. We say: these objects have the same attribute. The attribute of the kite is that it flies in the air. The attribute of the butterfly is that it flies in the air. The attribute of the helicopter is that it flies in the air. An attribute tells something about an object. For instance, whether it can fly, or what colour is has.* Compare with blocks of table 1 and 3 *In case of the blocks, we looked at the colour. So we grouped them by colour* (point out). *Who knows another attribute of the block? Very good, the form of the blocks.* Description of the solution process *Now I repeat once more what we actually did in our heads.* *How did we discover the attribute 'flying in the air'? Right, first we searched for attributes of every object seperately* (point to the objects). *Then we compared the discovered attributes. We found out that the objects have (at least) one attribute in common: they can ...* fly (have a tail). Description of the problem type *Objects with the same attribute fit together in one group. Objects that do not share the same attribite do not fit together in one group. Does a car fit in this group? No, a car doesn't have the same attribute: it can not fly. So a car does not belong to this group*

FIG. 4. Part of teacher's manual of the indirect teaching method with respect to the inductive reasoning problem generalization (see Fig. 2).

The lessons are conducted in dialogue form with the whole group. The students sit in a semicircle so that they have a good view of each other, of the teacher and of the problem plate which is shown in front of the class. Crucial to the dialogue is the teacher's ability to incite the students to reason logically by asking questions to a solution of the problem.

The purpose of Study 1 was to carry out a formative evaluation of the implementation of the programme in the third grade of primary schools. Study 2 was a summative evaluation. The following issues were investigated: a) Can the programme be implemented in the classroom? (Study 1); and b)

Does the inductive reasoning skill of students who follow the programme improve to a greater extent than the inductive reasoning skill of students who do not follow the programme (Studies 1 and 2)?

METHOD

Design

Study 1 used two existing classes of third grade students as experimental groups, and one existing class control group. In experimental group 1 (n = 17; mean age 84.0 months), the programme was applied according the direct teaching method; and in experimental group 2 (n = 13; mean age 84.6 months), the programme was applied according to the indirect teaching method. The control group (n = 9; mean age 85.6 months) participated in the regular third grade activities. There were no significant age differences between the three groups. All groups took the same pretests, posttests and follow-up tests. The interval between pretests and posttests was 10 weeks (period in which lessons were followed), and the interval between the posttests and the follow-up tests was 3½ months.

Study 2 was performed in 18 existing classes of third grade students: six were used as experimental groups, and twelve were used as control groups. In the experimental groups (n = 99 in total; mean age 85.5 months), an adapted version of Inductive Reasoning 1 was applied, meaning that the direct teaching method was used. The instruction of this programme, however, was slightly changed on the basis of the results of Study 1 to avoid some possible misunderstandings due to the instructions. The control groups (n = 232 in total; mean age 88.7 months) continued their normal classroom activities. There were no significant age differences between groups. All groups took the same pretests and posttests. Durability of the learning effects was evaluated by administering the follow-up test three months after completion of the programme.

Population and Sample

Study 1 used third-grade students from regular primary education, living in a home situation of social-economic backwardness (students 'at risk' or Chapter 1 Children in US terminology). In the Netherlands, social-economic backwardness is quantified in a 'student score'. This score is based on level of education, social-economic status, and ethnic background of the students' parents. The score is 1.25 for Dutch working-class families; 1.40 for ship-students in a boarding school or a foster home; and 1.90 for students of families with at least one parent not having the Dutch nationality (and with very limited educational and professional level) (Sijtstra, 1992). We first determined a pool of schools of which at least 95% of the students had a student score of 1.90. From this pool we selected schools on the basis of the following criteria: school were (1) public schools; (2) not reform schools; and (3) not involved in fusions or other important organizational changes. Study 2 used the same population as Study 1, and the same criteria for drawing a sample.

Instruments

Inductive reasoning (Study 1) was determined using Raven's Standard Progressive Matrices (SPM; Raven, 1958). In this study, the SPM was used as a far-transfer test. The SPM tasks are meaningless and abstract geometrical, which is in contrast with the training tasks, which are meaningful problems taken from children's everyday life and from school-type subject matters. The SPM consists of five sets of twelve matrix tasks with figurative material. Each task appeals to the reasoning processes of 'apprehension of experience', 'education of relations', and 'education of correlates' (Spearman, 1923/1973; Sternberg, 1977). The SPM tasks show strong similarities with the tasks of cross-classification and system-construction of Inductive Reasoning 1.

The analyses of the learning effects used the SPM total scores and the five types of tasks within this test. Bereiter and Scardamalia (1979) calculated for each of the first 48 tasks of the test the number of schemes which must be simultaneously activated for correctly solving the tasks. This quantification of task characteristics was expressed by means of ‘Mental Demand’. Together with ‘Mental Capacity’, which indicates the number of schemes which someone can activate simultaneously, this forms the basis for the ‘Mental Construct’ concept of Pascual-Leone (1970). The Mental Demand of the first 48 Raven SPM tasks was classified in five categories, denoted M1 to M5. Items in category M1 can be solved perceptually. M2, M3, M4, and M5 items are progressively more complex, and are increasingly better measures of inductive reasoning.

Content analyses of the number and the kind of essential attributes or relations of the tasks used in Inductive Reasoning I indicated Mental Demand ranging from M1 (introduction of the simple class of tasks) to M4 (the most complex tasks at the end of the programme). We assumed that the learning effects manifest themselves in the M2, M3, and M4 tasks because these tasks showed strong similarities with the classes of tasks identified as cross-classification and system-construction in Inductive Reasoning 1 (Hamers, De Koning, & Pennings, 1995).

In Study 2, inductive reasoning was determined using the Programme Related Test (PRT; De Koning, Hamers, & Sijtsma, 1996; De Koning, Hamers, & Sijtsma, in preparation). The items of this test were based on Inductive Reasoning 1. The items represented the six item classes (Fig. 1), and referred to the inductive reasoning processes. The PRT is considered to be a near-transfer test (De Koning et al., 1996).

Procedure

The design of Study 1 offered the possibility to determine the differential learning effects of both experimental groups and the control group. However, because of the confounding of group with teacher the complete and correct application of the programme by the teachers had to be controlled. Therefore, all lessons were observed by an independent observer and, moreover, were evaluated using a teachers’ questionnaire (De Koning & Hamers, 1994b). After each lesson, teachers answered questions concerning the preparation of the lesson, the duration of the lesson, the clarity and the practical usefulness of the instructions, and the reactions of the students. The observer (a trained research assisstent) answered questions (De Koning & Hamers, 1994a) concerning whether or not the teacher followed the instruction as prescribed by the teacher’s manual. The questions in the observation list and the questionnaire overlapped, and different interpretations concerning the instructions could thus become evident. The problem of confounding of group and school was solved by sampling similar schools having students with similar relevant characteristics (see section on Population and Sample).

Prior to answering the research questions, it was first determined whether the research group showed any retardation in inductive reasoning ability. The raw scores on the pretested SPM were transformed into percentage scores (Raven, Court, & Raven, 1992). Based on these scores, three level groups were defined. Forty-one percent of the students were assigned to the low scoring group (score below P_{25}); forty-nine percent to the average scoring group (score between P_{25} and P_{75}); and the remaining ten percent to the high scoring group (score above P_{75}). As the distribution of students clearly disagreed with the standard distribution, it was concluded that the research group was retarded in inductive reasoning.

RESULTS OF STUDY 1

Implementation of the Programme

The teachers’ questionnaire and observation list showed that the teachers reacted positively to the clarity and practical usefulness of the instructions.

However, the teacher in the group using the direct teaching method, could not easily apply the systematic, analytical solution procedure to the first few easy tasks because the solution of these tasks was sometimes obvious without instruction. The teacher of the indirect teaching group, experienced difficulty in applying the instruction method to the difficult tasks. Because the direct teaching method sometimes took much time, the teacher could not always introduce the reflection part of the instruction.

The intended role change in the application phase was not brought about completely in either of the two groups because some of the students were insufficiently competent in the Dutch language. Moreover, while ensuring that the four steps for solution were correctly taken, the teachers were not always consistent in naming the steps. This may have led to the students being unaware of the meaning of the four steps, which in turn led to inability to go ahead with the role swap. Finally, because the exchange of roles was not an integrated part of daily educational practice, this may have had a negative effect on bringing about the role swap.

To summarize, for the teachers involved in this project, the programme, its content, and its didactical requirements were new and were not easy to deal with. In general, however, the observation list and the questionnaire showed that the teachers were able to follow the prescribed instructions.

Learning Effects of the Programme

Does the inductive reasoning skill of the experimental groups show greater improvement than that of the control group? To control for absence of randomization, the pretest scores of the SPM were included as a covariate in the analyses of variance.[1] To eliminate systematic differences between the groups in retarded situations, two covariates were added: number of years that students had been living in the Netherlands, and number of years that they followed Dutch education. Tables 1 and 2 summarize the results. Table 1 shows the means and standard deviations of the pretested, the posttested, and the follow-up tested SPM total scores and M2, M3 and M4 scores.

The combined experimental groups differed significantly (5% level), from the control group on the SPM posttest total score and the SPM follow-up test total score (Table 2). The experimental groups did not differ on either of these tests. The covariates each had a significant influence on the posttest total score and follow-up test total score. There were no significant interaction effects or main effects due to sex.

The combined experimental groups and the control group differed on the M2, M3, M4 posttests and follow-up tests. Kolmogorov-Smirnov tests on the posttest score and follow-up test score of the M2, M3, and M4 tasks showed that scores were normally distributed in all cases except the M4

[1] Assumptions of homoscedasticity, uniform regressions in the case of covariates, and the normal distributions of the scores, are met unless otherwise stated in the text

TABLE 1
Means and Standard Deviations of SPM

	Pretest		Posttest		Follow-up test	
	M	*SD*	*M*	*SD*	*M*	*SD*
Exp. gr. (n = 30)						
Total	15.5	4.8	26.3	7.3	29.5	9.2
M2	4.5	2.7	8.7	2.9	9.8	3.3
M3	0.6	1.1	4.2	2.9	5.4	3.6
M4	0.4	0.7	1.5	1.6	2.1	1.9
Contr. gr. (n = 9)						
Total	17.2	6.4	23.0	5.6	25.1	5.5
M2	3.9	1.7	7.2	2.7	8.4	2.1
M3	2.6	2.7	3.3	2.4	4.2	2.9
M4	0.8	1.1	0.8	0.8	1.1	1.3

posttest scores. On the posttest, the experimental groups differed significantly from the control group on each of the Mental Demand types M2, M3, and M4. The follow-up test showed comparable results. The experimental groups did not differ on the SPM posttest, but they differed on the SPM follow-up test. The group using direct teaching had higher scores on M3 and M4 than the group using indirect teaching.

Depending on sample size, significant differences in combination with small effect sizes can cast doubt on confirmation of the hypotheses (Hager, in preparation). Table 3 shows the effect sizes on the SPM total score and

TABLE 2
Univariate Covariance Analyses with Factors of Group and Sex, and SPM Scores on the Posttests and Follow-up Tests as Dependent Variable

	Exp. Gr.–Contr. Gr. (*N* = 30) (*N* = 9)				Exp. Gr 1[a]–Exp. Gr. 2[a] (*N* = 17) (*N* = 13)			
	Posttest		Follow-up test		Posttest		Follow-up test	
SPM Raven	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
Total 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: Experimental group, direct teaching method. Exp. Gr. 2: Experimental group, indirect instruction method

on the Mental Demand types, M2, M3, and M4. The largest effect sizes[2] in the experimental groups were observed on the M3 and M4 tasks. The effects on the M2 tasks were small. The effects of the follow-up tests and the posttest were almost equal. Because of the significant differences between the experimental groups on the M3 and M4 tasks, the effect sizes were also calculated for the separate groups. The effect sizes of the group using the direct teaching method were in all cases slightly higher than those in the group using the indirect teaching method.

Combining the results for both experimental groups, it can be concluded tentatively that, controling for the confounding of school, school group, and teacher, the direct teaching method brought about better learning effects than the indirect method. Therefore, in Study 2 we used a modestly adapted version of the direct teaching method.

RESULTS OF STUDY 2

The question was whether the inductive reasoning skill of the experimental group showed greater improvement than that of the non-experimental group. In Table 4 the means and standard deviations of the pretest scores and the follow-up test scores on the Programme Related Test are presented. The Scheffé test with factor group and dependent variable total score on the pretest, showed no significant differences between pairs of experimental subgroups, and pairs of control subgroups ($p = .05$). Thus, the six experimental groups were combined, as were the twelve control groups. The results of univariate analysis of variance with factor group and dependent variable total score on the pretest and follow-up test were $F = .81$; $p = .27$, and $F = 27.04$; $p = .00$, respectively. Thus, the follow-up test scores of the experimental groups are significantly higher than the follow-up test scores of the control group. There were no significant interaction effects and main effects due to sex. The corrected effect size on the total scores of the experimental groups was 1.04.

CONCLUSIONS AND DISCUSSION

Inductive reasoning allows us to construct an orderly world by introducing structure. It is a daily activity of both adults and children. A person uses inductive reasoning to construct a coherent pool of knowledge which can be easily used and extended. An orderly base of knowledge is an important requirement for learning. Klauer’s (1989) description of inductive reasoning (see Figure 1) should be seen as a ‘definitional model’ (Klauer & Phye, 1995). This means, that Klauer’s model is presciptive instead of descriptive: It does not explain how subjects solve inductive reasoning problems but

[2] We adopted Cohen’s (1988) convention for describing the magnitude of effect size: ES = 0.20–.49 = small effect; ES = 0.50–.79 = moderate effect; ES = ≥ 0.80 = large effect.

TABLE 3
Effect Sizes (d) on SPM Total Scores and Scores on the Different M Types

	Exp. Groups 1 + 2 (N = 30)		Exp. Group 1[a] (N = 17)		Exp. Group 2[a] (N = 13)	
SPM Raven	Posttest–Pretest	Follow-up test-Pretest	Posttest–Pretest	Follow-up test-Pretest	Posttest–Pretest	Follow-up test-Pretest
Total 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. Gr. 1: Experimental group, direct teaching method. Exp. Gr. 2: Experimental group, indirect teaching method.
* $p \leq .10$. $d = (M_{EG} - M_{CG})/S_p$.
** $p \leq .05$.

TABLE 4
Means and Standard Deviations of the Programme-Related Test

	Pretest		Follow-up test	
	M	*SD*	*M*	*SD*
Exp. group (n = 99)	23.4	6.9	31.8	5.7
Contr. group (n = 232)	24.1	6.5	26.9	5.3

rather specifies a strategy how to solve such problems effectively. Thus, the model is useful for the construction of programmes because it specifies those thinking skills typical of inductive reasoning. Furthermore, it shows the important role of inductive reasoning for cognitive activities that could be taught in the classroom.

The questions put forward were: Is it possible to implement programmes in the natural setting of the classroom? And is there an indication for a differential effect between the instruction methods of the two experimental groups? The answer to the first question was affirmative. Our results showed that teachers were able to deal with the main didactical requirements of the programme. However, do we need such 'across-the-curriculum' programmes in schools? The answer to this question determines to a great extent how thinking will be taught. If thinking is taught in an 'across-the-curriculum' course, objectives for thinking skills and strategies will be the basis of the programme. If thinking is taught in the context of a school subject, content objectives will be the basis. There is a considerable debate as to which context is more effective for teaching at-risk students (Resnick, 1987). Proponents of the first approach argue that low-achieving students may experience overload if they have to learn both content and skills simultaneously. For instance, Feuerstein (1980) developed content-free programmes using geometric shapes and pictures. Most other programmes for teaching thinking, however, use a combination of content-free and daily life formats (e.g., Klauer, 1989). Proponents of the second approach, argue that programmes should be content-related because a substantial part of skills and strategies is content-specific and these skills and strategies do not transfer easily to other areas (Resnick, 1987). We agree with Presseisen's (1988) compromise, that content-related programmes, in which skills are learned as a means to learning the content of solving problems, are generally preferable, unless students experience great difficulty with the content. In this event, an 'across-the-curriculum' course might be the best choice, provided that transfer is built into the programme and that the substance of the programme is well coordinated with the school content courses.

The answer to the second research question was also affirmative. The formative and the summative studies showed that the programme Inductive

Reasoning I produced significantly higher scores on the Raven SPM and on the Programme Related Test (PRT) directly after following the programme and also more than three months later, than in the control group. Moreover, given that Klauer's training used small groups, the effect sizes found here were convincing. A possible explanation is that the teachers were able to successfully implement the programme using a detailed instruction manual (De Koning & Hamers, 1995). Furthermore, the effects on the follow-up test in most cases equalled those of the posttest. Thus, a certain amount of durability of the results could be established. Our positive results in combination with the many experiments of Klauer (1989) give rise to optimism.

This study revealed tentative evidence that instruction according to the direct teaching method leads to good learning results. Gal'perin (1980 demonstrated the effectiviness of the direct teaching method by means of experimental research and by means of a theoretical analysis. The concept of a 'complete orientation basis' occupies a central position in this analysis. This concept entails the totality of knowledge and of existing information on which an individual bases his/her execution of a task. As an explanation for the difference in learning effects between the two methods found in our study, we accepted the need of backwards students for having a complete orientation basis, at least in the initial phase of the solution process. As a consequence, in future research a mixture of a direct teaching method and an indirect teaching method may be constructed, which depends on characteristics of the students and on the task difficulty.

Ninety-five percent of the students in this study belonged to Turkish and Moroccan minority groups. Their parents came to the Netherlands as so-called 'guest workers'. The label 'ethnic minority group children' is used to designate children belonging to an ethnic group whose values, customs, thought patterns, or language differ from the dominant culture. The differences between the minority groups and the majority group can shape the attitude of minority group children to the general school situation. The policy of the Dutch government is directed towards integration of these groups into Dutch society. The government provides 'minority' schools with extra financial facilities for remedial teaching and other compensating activities. Nevertheless, ethnic minorities suffer from forms of social-economic deprivation. The resulting relative poverty affects rates of learning, i.e., the acquisition of thinking skills (e.g., general problem solving skills, inductive reasoning), and school success (Van Langen & Jungbluth, 1990). Our studies were intended to stimulate these thinking skills.

In the programme and in the tests the transfer issue was seriously addressed. The positive results on the posttests and follow-up tests indicated that transfer had been successful. A next step in future research is that we should stimulate and help the teachers to enhance transfer not only to tests but also to other domains of the school curriculum, and to problems encountered in out-of-school settings. These efforts should provide a clearer picture of the conditions under which inductive reasoning acquired in these programmes, will effectively transfer to other contexts. In new research (De Koning & Hamers, in preparation), it will be investigated whether improvement of inductive reasoning is also evident in problem solving in the school subject of comprehensive reading.

REFERENCES

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.

Costa, A. L. (Ed.). (1991). *Developing minds: Programs for teaching thinking.* Alexandria, VA: Association for Supervision and Curriculum Development.

Curtis, R. V. (1988). When is a science analogy like a social studies analogy? A comparison of text analogies across two disciplines. *Instructional Science,* **13,** 99–117.

De Bono, E. (1983). The cognitive research trust (CoRT) thinking program. In Maxwell, W. (Ed.), *Thinking. The expanding frontier.* Philadelphia: The Franklin Institute Press.

De Koning, E., & Hamers, J. H. M. (1994a). *Observation list teachers' instruction.* Utrecht: ISOR Utrecht University.

De Koning, E., & Hamers, J. H. M. (1994b). *Questionnaire for teachers' instruction styles.* Utrecht: ISOR Utrecht University.

De Koning, E., & Hamers, J. H. M. (1994c). *Student questionnaire.* Utrecht: ISOR Utrecht University.

De Koning, E., & Hamers, J. H. M. (1995). *Inductief redeneren 1 [Inductive reasoning I].* Utrecht: Universiteit Utrecht ISOR.

De Koning, E., & Hamers, J. H. M. *How to infuse thinking skills in comprehensive reasoning?* Manuscript in preparation.

De Koning, E., Hamers, J. H. M., & Sijtsma, K. (1996). *Test for inductive reasoning. Test and manual.* Den Haag: HCO.

De Koning, E., Hamers, J. H. M., & Sijtsma, K. *Designing a test for inductive reasoning: integrative use of several data analyses methods.* Manuscript in preparation.

Donaldson, M. (1978). *Childrens' minds.* London: Fontana.

Feuerstein, R. (1980). *Instrumental Enrichment. An intervention program for cognitive modifiability.* Baltimore: University Park Press.

Gal'perin, P. J. (1980). *Zu Grundfragen der Psychologie [Main questions in psychology].* Berlin: Volk und Wissen.

Guthke, J., & Wiedl, K. H. (1995). *Dynamisches Testen [Dynamic Testing].* Göttingen: Hogrefe.

Halpern, D. F. (Ed.). (1992). *Enhancing thinking skills in the sciences and mathematics.* Hillsdale, NJ.: Erlbaum.

Hamers, J. H. M., De Koning, E., & Pennings, A. H. (1995). *Stimulering van denkvaardighe-*

den van allochtone leerlingen [Stimulating thinking skills of minority group children]. Utrecht: Universiteit Utrecht ISOR.

Hamers, J. H. M., & Overtoom, M. (Eds.). (1997). *European inventory of teaching thinking programmes.* Utrecht: SARDES.

Hamers, J. H. M., Sijtsma, K., & Ruijssenaars, 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 cristallized ability intelligence. *Journal of Educational Psychology,* **57,** 253–270.

Klauer, K. J. (1989). *Denktraining für Kinder 1. Ein Program zur intellectuellen Förderung [Inductive reasoning. A programme for the stimulation of inductive reasoning].* 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.

Nickerson, R. S., Perkins, D. N., & Smith, E. E. (1985). *The teaching of thinking.* Hillsdale, NJ.: Erlbaum.

Pascual-Leone, J. A. (1970). A mathematical model for the transition rule in Piaget's development stages. *Acta Psycholica,* **63,** 301–345.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension monitoring activities. *Cognition and Instruction,* **2,** 117–175.

Presseisen, B. Z. (1988). *At-risk students and thinking: Perspectives from research.* Washington, DC: NEA/RBS.

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.

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. (1985). *Beyond IQ. A triarchic theory of human intelligence.* New York: Cambridge University Press.

Sijtstra, J. (1992). *Balans van het taalonderwijs halverwege de basisschool [Evaluation of language education half-way primary school].* Arnhem: CITO.

Van Langen, A., & Jungbluth, P. (1990). *Onderwijskansen van migranten [Educational opportunities of immigrant students].* Lisse: Swets & Zeitlinger.

Van Parreren, C. F. (1981). *Onderwijsproceskunde [Educational processes].* Groningen: Wolters-Noordhoff.