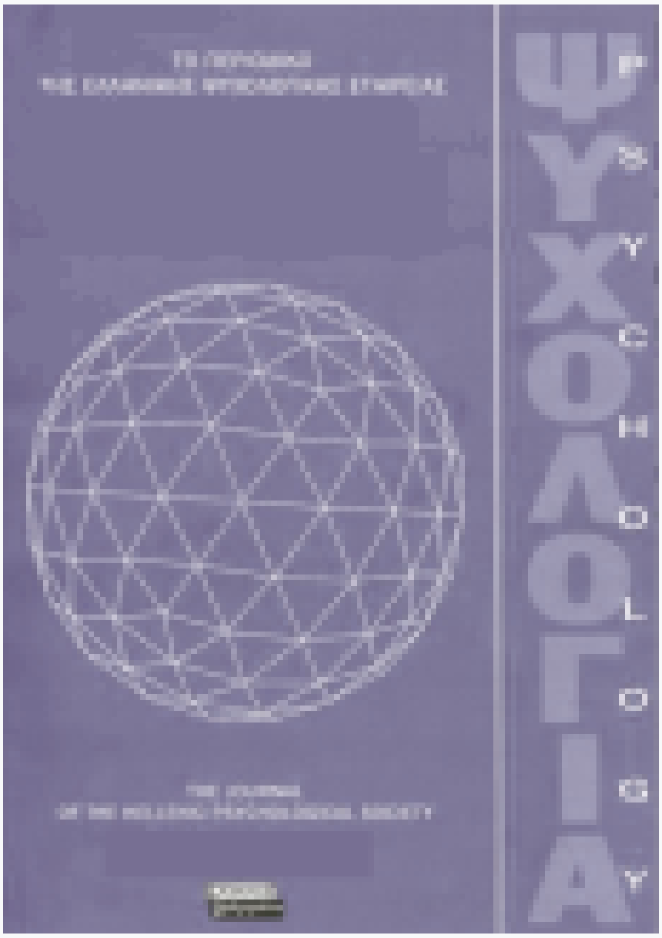
Psychology: the Journal of the Hellenic Psychological Society

Vol 6, No 3 (1999)



Teaching thinking: The role of general and domain-specific abilities in cognitive change

Anastasia Efklides, Maria Papadaki, Georgia Papantoniou, Martha Koutsioumba, Grigoris Kiosseoglou

doi: [10.12681/psy_hps.24281](https://doi.org/10.12681/psy_hps.24281)

To cite this article:

Efklides, A., Papadaki, M., Papantoniou, G., Koutsioumba, M., & Kiosseoglou, G. (2020). Teaching thinking: The role of general and domain-specific abilities in cognitive change. *Psychology: The Journal of the Hellenic Psychological Society*, *6*(3), 342–364. https://doi.org/10.12681/psy_hps.24281

Teaching thinking: The role of general and domain-specific abilities in cognitive change

ANASTASIA EFKLIDES
MARIA PAPADAKI
GEORGIA PAPANTONIOY
MARTHA KOYTSIOUMBA
GRIGORIS KIOSSEOGLOU
Aristotle University of Thessaloniki, Greece

ABSTRACT This study aimed at investigating, first, the role of general ability *(g)* and domain-specific abilities in cognitive change, and second, the possible interaction of training method with narrow abilities such as verbal/semantic, visual-spatial, and numeric fluency. The domain-specific abilities involved were quantitative-relational (QR) and causal-experimental (CE). The sample comprised 1127 students of 12, 14, 16 and 20 years of age of both genders. All participants were tested with a battery of 7 tasks (two verbal, two numeric, and three visual-spatial) tapping *g*. Four tasks of proportional reasoning were addressed to the QR ability and four tasks involving experimentation addressed CE ability. The QR and CE tasks were administered as pre- and posttest. Participants received three forms of training on the QR or CE ability, namely algorithmic, metacognitive, and computer-assisted. Structural modeling analysis showed that both *g* and domain-specific abilities are involved in cognitive change. It was also found that the C-A training made use of verbal/semantic fluency.

Key words: Cognitive intervention, domain specificity, general intelligence.

The idea that thinking can be taught lies at the core of any educational enterprise. Even if this idea were implicit and not stated in the goals pursued by the various educational systems in the past, it underpinned the efforts of educators and shaped the selection of the abilities and skills to be trained at school. Furthermore, the intuitive conceptualization of the structure and functioning of thought guided the selection of the school subjects to be taught and the teaching methodology to be used.

In psychology, the teaching of thinking became a prominent issue in the last three decades, since researchers started to test Piaget's views related to the developmental constraints of thinking, that is, the stage-like development of thought. The first question was if development of thinking can be accelerated through direct intervention on thought structures. The second question was if cognitive acceleration and thought restructuring is better achieved through the exercise of cognitive conflict rather than other teaching methods.

However, there was another reason that

Note: This study was supported by the Greek General Secretariat of Research and Technology through a grant to the first author.

Address: Anastasia Efklides, School of Psychology, Aristotle University, 540 06 Thessaloniki, Greece. Tel.: *30-31-997374, Fax: *30-31-997384, E-mail: efklides@psy.auth.gr

brought the piagetian theory into the foreground of educational work. This reason had to do with the perceived relevance of thought structures, as described in the piagetian theory, with school subjects and particularly mathematics and science. This made piagetian theory the best candidate for the understanding of thinking in complex domains, such as science. (For a review of training efforts on piagetian concepts see Brainerd, 1983; Goosens, 1992.)

As piagetian theory got under scrutiny, a principle tenet of it, namely, the unified nature of thought structures, or generality of thinking, was challenged. This was so because research failed to reveal a single omnipotent thought structure or transfer from the abilities (or concepts) trained to all the other abilities supposedly controlled by the same structure. This led to neo-piagetian theories which claimed that thought structures or abilities (or skills) are organized in smaller units or systems, which center around specific domains (Case, 1985; Demetriou & Efklides, 1987, 1994; Fischer, 1980; Fischer & Farrar, 1988). This kind of research led to the realisation that thinking is a much more complex phenomenon than assumed, and the piagetian theory is neither the only nor the best account of it.

Indeed, one of the main contributions of cognitive developmental research in the 1980s was the introduction of the concept of *domain specificity* (Carey, 1985; Chi, Feltovich, & Glaser, 1981; Demetriou & Efklides, 1981, 1987; Keil, 1984). Domain specificity implies that knowledge is acquired and develops within specific, content-rich areas and is influenced by factors operating in these areas rather than general ones. In its strong version, this theoretical view allows no room for a whole structure, dictating the functioning of thought, or for general abilities and skills operating across domains.

On the other hand, it is well documented that intelligence as measured by intelligence tests, is correlated positively with academic performance (for a review see Neisser, Boodoo, Bouchard, Boykin, Brody, Ceci, Halpern, Loehlin, Perloff, Sternberg, Urbina, 1996). The question is why this is so, and if it is general ability per se that pervades and shapes performance or if there is specificity in abilities analogous to that found in thought structures. The truth of the matter is that modern psychometric research has shown the existence of abilities of various levels of generality, i.e., broad and narrow, organized hierarchically. At the top of the hierarchy there is a general factor *(g)*, presumably tapping general intelligence (Carroll, 1993; Gustafsson, 1984, 1988). Therefore, although the concept of intelligence, as general ability, in the piagetian sense has been refuted, it is still strong in the psychometric tradition and we need to determine its role in cognitive development.

These theoretical developments have obvious bearing on the teaching of thinking, because they made explicit the question: What do we teach: general thinking ability or domain-specific abilities? And if we teach domain-specific, how do we determine which abilities should be trained and which method should we use for cognitive intervention? Finally, does training of domain-specific abilities mean that general ability is not part of the mechanism of cognitive change?

The phenomenal simplicity of the question general vs. domain-specific was made clear when researchers set out to define the two terms and operationalize them. A wealth of approaches resulted (for reviews see Baron & Sternberg, 1987; Hamers & Overtoom, 1997; Kuhn, 1990; Nickerson, Perkins, & Smith, 1985) which made clear that there is no consensus in the field. In essence, selection of the abilities to be trained is guided by the theoretical approach one adopts (Kuhn, Amsel, & O' Loughlin, 1988) or by basic school subjects, such as language, physics, and mathematics. This situation has the drawback that there is no comparative or integrative work which could tell us which abilities are really necessary for the development of thought in its various aspects and which is the mechanism of cognitive change.

Our research was an attempt in this direction. Following the theory of Experiential Structuralism developed by Demetriou and Efklides (Demetriou & Efklides, 1987, 1994) we defined domains in terms of basic categories of mind, such as quality, quantity, causality, space and semantic representation of states and events. These domains are processed by middle-level abilities (called Specialized Structural Systems, SSSs) which function in between general ability, on the one hand, and task-specific abilities on the other. The SSSs identified by the theory are the Qualitative-Analytic, the Quantitative-Relational, the Causal-Experimental, the Spatial-Imaginal, and the Verbal-Propositional. (For an overview of the theory see Efklides, 1999.)

Following modern psychometric research we defined general ability as the ability that underlies performance on a variety of tasks, representing different narrow abilities. According to Gustafsson (1984, 1988) general intelligence *(g)* is perfectly correlated with inductive ability; but such a definition of *g* may lead one to overlook other aspects of it, namely the verbal and visualization aspects of it.

Our research questions were:

1. Is cognitive change limited within knowledge domains or does it make use of *g*?

2. Furthermore what is the role of other, middle-level abilities identified in the psychometric tradition such as fluency in the processing of verbal, and visual-spatial material? Are these middle-level abilities involved in the process of cognitive change within domains?

3. Do *g* and the above mentioned middle-level abilities (domain-specific or g-related) interact with the instructional method employed?

In order to answer the above questions we performed a study in which we included measures of two domain-specific abilities, namely the quantitative-relational (QR) and the causal-experimental (CE); measures of verbal, visual-spatial and numeric abilities from which *g* could be inferred; this numeric set of tasks was included because QR abilities presuppose numeric fluency rather than verbal or visual-spatial fluency. Finally, three forms of training were used, namely algorithmic, metacognitive, and computer assisted (C-A).

The *algorithmic training* involved a step by step description of the solution process, aiming at providing participants with the procedures needed. In essence it aimed to provide participants with the means with which they could "think" (i.e., solve the problem given). It required no reflection on the thinking process or about the general strategies that could be used in the present case.

A training procedure that requires thinking *about* thinking rather than thiking *with* it (Kuhn et al., 1988) is metacognitive in nature, because it requires awareness of the thinking process. This kind of training does not involve awareness of the monitoring process, to which the term "metacognitive" usually refers to. In our case, *metacognitive training* involved description of the general strategy to be used, and the participants had to accomodate it to the specifics of the problem to be solved.

Finally, the *computer-assisted* (C-A) training involved a step by step procedure accommodated for computer use. It presented a number of questions, each of them corresponding to each of the steps of the algorithmic training. The difference from algorithmic training was that the person had a multiple-choice format from which to select the correct answer. There was also immediate feedback on every selection in the form of right/wrong. A more extensive description of the correct answer in the form of the metacognitive training was given upon selecting it and before going on to the next question. This form of training engaged participants in thinking but did not provide a readymade procedure, as the algorithmic training did. Participants had to infer or construct the problem-solving procedure via their selection of the alternatives of each multiple-choice question. The informative feedback at the end of the selection process was meant to facilitate this constructive process and the awareness of the solution process, although participants were not required to apply it as was the case of metacognitive training.

Hypotheses

The hypotheses tested were the following:

1) If cognitive change is limited within domains, then there will be no use of *g* in the posttest performance on QR and CE tasks. 2) The same regards the verbal/semantic, visual-spatial, and numeric abilities. 3) If cognitive change is limited within each domain, then there will be no transfer of training from QR ability to CE ability and vice versa. 4) However, if *g* is manifested in novel and difficult situations, where the person has no immediately available response, then cognitive change will make use of both *g* and domain-specific abilities. This will be obvious if the instructional method poses demands on the person's *g*. As was made clear from the description of the instructional methods used in this study, the three training forms differed in terms of processing demands; the assumption was that algorithmic training was the least demanding; metacognitive training was next, and most demanding of all was the C-A. Therefore, algorithmic training would not require the use of general intelligence whereas the other two forms of training, and particularly C-A, would do. This implies that in the case of algorithmic training, cognitive change would rely exclusively on domain-specific processes whereas in the case of metacognitive and C-A training it would involve both general intelligence and domain-specific abilities.

5) Finally, the *effectiveness of the training method* would be inverse to their cognitive demands, the algorithmic being more effective than the metacognitive and this, in turn, more effective than C-A.

Method

Design

In order to test the above hypotheses, an intervention study was designed. It involved two experimental groups and one control group. The first experimental group received training on the quantitative-relational (QR) SSS (Quantitative-Relational Treatment Group, QRTG). The targeted ability was proportional reasoning. The second experimental group received training on causal-experimental (CE) SSS (Causal-Experimental Treatment Group, CETG). The targeted ability was experimentation. The control group (Control Treatment Group, CTG) received no training at all. The idea guiding the selection of targeted abilities was that they develop during the age period covered in the study, that is, adolescence.

All participants were tested before and after training with the same battery of tasks. The battery consisted of four QR and four CE tasks. At the pretest all participants were also required to solve a set of tasks addressed to General Intelligence *(g)*, and specifically to three aspects of it: verbal/semantic fluency, visual-spatial fluency, and numeric fluency.

Participants

The sample comprised 1127 students of 12, 14, 16, and 20 years of age. Specifically, there were 356, 413, 314 students of 7th, 9th, and 11th grade, respectively, and 44 university students (see Table 1). The small number of university students was due to the fact that, despite the large number of them tested at the pretest (120), only a limited number of them had performance low enough to be selected for training. Both genders were about equally represented. All participants came from low and upper middle class families.

Tasks

There were three sets of tasks: Quantitative-Relational tasks, Causal-Experimental tasks, and General Intelligence task.

The QR and CE sets of tasks were constructed so that they had the same structure, and differed only in terms of the ability required for their processing. They were first used in Efklides, Demetriou, and Gustafsson (1992). The structure of the tasks resembled Fischer's (1980; Fischer & Farrar, 1988) hierarchy of skills levels. Thus, although all four tasks in a set tapped the same ability, they differed in structural complexity

Table 1
Distribution of subjects according to age, treatment group, training form, and level of training

Age	Training form	Level I	Level II	N
		QRTG		
12	Alg.	37	17	142
	Met.	40	13	
	C-A	22	13	
14	Alg.	38	22	163
	Met.	36	24	
	C-A	24	19	
16	Alg.	18	20	114
	Met.	19	22	
	C-A	15	20	
Univ. students	Alg.		5	18
	Met.		5	
	C-A		8	
		CETG		
12	Alg.	41	16	149
	Met.	42	16	
	C-A	21	13	
14	Alg.	36	22	152
	Met.	39	17	
	C-A	20	18	
16	Alg.	21	16	111
	Met.	22	17	
	C-A	14	21	
Univ. students	Alg.		5	14
	Met.		5	
	C-A		4	
		CTG		
12				65
14				98
16				89
Univ. students				12

and respective difficulty. Each of the tasks in a set corresponded to one of the four developmental levels of the tier of abstract thought, namely, the level of single abstract sets, and the levels of abstract mappings, abstract systems, and systems of abstract systems. The tier of abstract thought is acquired in adolescence from 12 years onwards. Each level is acquired in approximately two years of time after the preceding one. The construction of the tasks allowed the identification of the cognitive level of the person in each ability, depending on the most difficult of the four tasks he/she had successfully solved. Furthermore, change of cognitive level rather than simple quantitative increase of performance scores could be used as criterion for the success of the intervention.

The *g* tasks represented the verbal, the imaginal, and the numeric symbolic systems. All tests, except the Number Series test were selected from the Kit of Factor-referenced Cognitive Tests (Ekstrom, French, & Harman, 1976). They were all time constrained tests.

Quantitative-Relational tasks. Four problems involving proportional relationships were addressed to the quantitative-relational ability (QR1-QR4). The scoring of the QR tasks was 0: for no or incorrect answer; 1: for partially correct answer; 2: fully correct answer.

QR1: Students were presented with a two times two table showing a relationship between watering frequency (twice and four times/month) and yield (2 and 6 kgs/hectare for plant A and 3 and 6 kgs/hectare for plant B). The task was to select from a number of alternatives which plant is more affected by watering and to explain why (i.e., produce the calculations necessary to justify one's choice). In this task, two variations had to be co-ordinated into a single set that forms an abstraction.

QR2: Two tables like the one in QR1 (i.e., a double table) were presented, showing the effects of watering on plants A and B in two areas I and II. Thus, in this task two single sets/abstractions had to be combined.

QR3: Two double tables were presented showing the effects of watering on plants A and B in areas I and II, when fungi are not present. Thus four single (or two double) sets of data, representing a system of abstractions, had to be combined to solve the task.

QR4: Four double tables were presented showing the effects of watering on plants A and B in areas I and II, when fungi are present, with or without use of fungicide, and when fungi are not present, with or without use of fertilizer. Thus, four double or eight single sets of data had to be combined to solve the task. This task represents a system of abstract systems.

Causal-Experimental tasks. Four problems were constructed involving the design of experiments in order to test hypotheses (causal-experimental ability, CE1-CE4). As stated above, these tasks were structurally equivalent to the QR tasks in the sense that they also tapped the four skill levels of the tier of abstract thought. The scoring of the CE tasks was on a 3-point scale ranging form 0: no answer or wrong answer; 1: partially correct answer; 2: fully correct answer.

CE1: A simple hypothesis was given ("the increase in watering frequency increases the productivity of plants") and the student was asked to use plants A and/or B and two watering frequencies (twice a month or four times a month) to design an experiment to test the hypothesis (single abstraction). A table was presented in which the student had to fill in the appropriate plant and watering, following the principle of "all the other things being equal..."

CE2: A hypothesis was given about the interaction between two factors ("watering increases the productivity of plant A, but it does not affect the productivity of plant B"). An experiment, integrating two single ones, had to be designed to test the above hypothesis (abstraction mapping).

CE3: In this task, the experiment to be designed had to test two interaction hypotheses, regarding the effects of watering on A in areas I and II and on B in areas I and II (abstract system). Thus a three-way design (plant X area X watering) had to be proposed.

CE4: In this task, yet another factor, fertilization, had to be taken into account. The solution of the task required a four-way experiment (plant X area X fertilizer X watering). Such a design captures the interaction of two abstract systems, therefore it is a system of systems.

General Intelligence tasks. Two of the tests addressed language-related abilities. They tapped *verbal/semantic fluency*. The first was the *Synonyms* test (SYN) and involved 10 items. The task was to produce as many synonyms as possible to each of the words presented. As synonyms were accepted words which had a relevant meaning and could be used in place of the word given in various contexts. The second semantic fluency task was the *Opposites* test (OP), which involved 10 items and required production of words with opposite meaning to that of the words given.

There were two sets addressed to numeric abilities. They tapped *arithmetic operations fluency*. These were: the *Number Series* test (NS) and the *Number Facility* (NF) test. The NS test contained 20 items in which a series of five or six numbers was given, and the task was to add two more numbers to the series (Gustafsson, Lindstrom, & Bjorck-Akesson, 1981). The NF test involved a large number of additions, substractions, multiplications, and divisions.

Finally, there were three tasks addressing the visual-spatial abilities. They measured *figural fluency, figural flexibility,* and *visualization*. The *Symbols* test (SYM) is a figural fluency task. It involved 5 items, which gave a word or a phrase and required the student to draw up to five different symbols to stand for it. The *Toothpicks* test (TP) is a figural flexibility task. It also involved 5 items, which tapped spatial arrangements of a set of toothpicks. The student was asked to present up to five different arrangements according to sets of specified rules. Finally, the *Paper Folding* test (PF) is a visualization task. It involved 10 items which required mental folding and unfolding of pieces of papér.

Training

There were three forms of training: the algorithmic, the metacognitive, and the computer-assisted.

The *algorithmic training* consisted of three parts: an introduction, explaining that the student had made a mistake at the previous testing and now he/she would be given instructions how to solve it. At the second part, a problem similar to the ones of the pretest, but applying to a different situation, was presented; the solution was then given in a step by step fashion. At the third part, the student was presented with a new problem (similar to the previous one) and was asked to solve it. Once the problem had been solved, feedback was provided. The feedback consisted in the detailed solution of the problem. The students were asked to study it and correct their mistakes. When the students finished this procedure, they were given the posttest.

For example, the QR training had the following form: In order to find out which plant, A or B, had the more productivity change, you need to divide the productivity of the plant when it is watered 4 times with the productivity of the plant when it is watered 2 times. That is, Plant A : 16:4 = 4, i.e., four times increase; Plant B : 20:5 = 4, i.e., four times increase. Therefore, A = B in terms of productivity change.

The respective algorithmic CE training had the following form: In order to test the hypothesis about the effect of light on the productivity of plants A and B you need to make the following experiment:

A. Plant	Light	B. Plant	Light
1. A	Dark	1. B	Dark
2. A	Light	2. B	Light

The *metacognitive training* was verbal in nature and focused on the general process (or strategy) rather than on the details of the problem-solving procedures. For example, the QR metacognitive training had the following form: In order to find out which plant had the more productivity change, you must compute the rate of productivity change for each plant and then compare the two outcomes. The respective CE metacognitive training had the following form: In order to test the hypothesis about the effect of light on the productivity of plants A and B, you need to make an experiment in which you keep all the other factors the same and vary only the light. After the solution of the new problem, feedback was also provided, including the detailed solution and the verbal explanation which was according to the instructions given.

The *computer-assisted* (C-A) training made use of the algorithmic presentation, accomodated for computer use. The computer application we used involved presentation of the problem and a number of questions, each of them tapping part of the solution of the problem. Only one question at a time was presented. Three or four alternative answers were provided to each question, and the student had to select the one he/she thought was the correct one. There was immediate feedback on every selection made in the form of: Right-Wrong. The selection procedure was terminated only when the student made the correct choice. At this point, the feedback was more extensive and included the principle on which the correct answer was based. In this way we wanted to make sure that even in case the correct selection was random or for wrong reasons, the student would be informed about the principles underlying the correct answer.

There were two levels of training: *Level I* training was administered to students who had not solved correctly the level 2 task, i.e. the QR2 or CE2 task for the respective treatment group. That is, it was given to students scoring 0, 1, 2 on the QR1 task and 0 or 1 on the QR2 task. *Level II* training was administered to subjects who had solved correctly the level 2 task (that is, to subjects scoring 2 on QR1 and QR2 tasks). The training tasks were similar in structure to the initial tasks but differed in content from the respective level 2 and level 4 tasks of the pretest. Students scoring 2 on QR4 task were not trained as they had achieved the highest level of thinking captured by the tasks. Thus, students were trained either one or two levels above their own.

Procedure

All students were tested before and after the training period with the QR and CE tasks. All testing was carried out in groups in the students' regular classrooms. The pretest session lasted approximately two school hours, and comprised the QR and CE tasks, and the *g* tasks. The training session was held about two weeks later, followed by administration of the QR and CE tasks as posttests. The training session lasted approximately half an hour. The training leaflets were personally addressed to each student according to their assignment to the experimental groups and the level of training.

The control treatment group received no training at all; students were instructed that they would be given no training and that at the posttest they should do their best to try to attend to the details of the tasks now that they were familiar with the requirements. It was particularly stressed that they must try to improve their performance.

Control group and experimental group students were tested in the same classroom. No time limit was imposed at any of the phases of the experiment.

Results

The mechanism of change

In order to test the hypotheses regarding cognitive change, the data were analysed with confirmatory factor analysis using the EQS statistical program (Bentler, 1993). For an overview of the structural analyses see Efklides (1999). The model tested involved variables for the pretest factors, posttest factors, and the *g* factor (in terms of fluency).

The *g* factor was represented as latent factor of three variables. Specifically, the first variable, the Verbal/semantic, was formed as the mean score of performance on the two verbal tasks. The second variable, the Visual-spatial, was formed as the mean score of performance on the three visual-spatial tasks, and the third variable, the Numeric, was formed as the mean score of performance on the two numeric tasks.

The pretest QR performance was represented by two variables; the first variable (PreQR1) was formed by the mean performance score on the QR1 and QR3 tasks. The second variable (PreQR2) was formed by the mean performance score on the QR2 and QR4 tasks. The same rationale guided the formation of posttest variables, namely the PostQR1 and PostQR2. The two PreQR variables loaded the latent PreQR factor; the two PostQR variables loaded the latent PostQR factor.

The pretest and posttest CE variables were formed similarly to the QR variables (PreCE1, PreCE2, PostCE1, and PostCE2). But there was one more pretest and posttest CE variable (PreCE3 and PostCE3, respectively), which represented the mean score of two items which tapped students' explanation of the experiments they proposed in the CE2 and CE3 task. This item was metacognitive in nature. The three PreCE and the three PostCE variables loaded the PreCE and PostCE latent factors, respectively.

The analysis proceeded as follows: First we confirmed the existence of the latent factors corresponding to the abilities presumed, namely the PreQR and PreCE factors, the PostQR and PostCE factors and the *g* factor. In a second step we introduced a path model, testing the effects of *g* (fluency) on the pretest factors (i.e., PreQR and PreCE). The effect of *g* on the PreCE factor was non-significant in the CETG but in the other two treatment groups was significant. Then we tested the possible effects of the pretest factors on the posttest factors (i.e., PostQR and PostCE). The effects of PreQR and PreCE on PostQR and PostCE factors, respectively, were confirmed. There was also an effect of QR factors on CE factors both in the pretest and the posttest. This effect was particularly strong in the CETG (namely, the PreQR on the PreCE factor), which suggests that the effect of *g* was replaced by the influence of the PreQR factor. The effects of PreQR and PostQR on CE factors were confirmed in all three treatment groups.

However, when the effect of the *g* factor on posttest factors was tested, no single model could be verified in all treatment groups. It was found that the *g* factor was related only to the factor corresponding to the ability trained; that is, in the QRTG it was related to the PostQR factor (see Figure 1a), in the CETG it was related to the PostCE factor (see Figure 1b), and in the CTG it was related to both the PostQR and PostCE factors (see Figure 1c). This finding implies that students mobilized both their general ability *(g)* and domain-specific ability in order to respond to the training provided. The Control Group, which had no training at all, relied on general ability for the solution of both the QR and CE posttest tasks.

The fit indices of the three final models were as follows:

For the QRTG: $x^2(58)=67.809$, $p=.18$; Bentler-Bonett Normed Fit Index (NFI)=.972, Bentler-Bonett NonNormed Fit Index (NNFI)=.994, Comparative Fit Index (CFI)=.996. *For the CETG:* $x^2(55)=63.288$, $p=.21$; NFI=.971, NNFI=.994, CFI=.996. *For the CTG:* $x^2(50)=59.796$, $p=.16$; NFI=.958, NNFI=.989, CFI=.993.

These results show, first, that Hypothesis 1 was not confirmed, because *g* was found to be part of the mechanism of cognitive change. What is important to note, however, is the allignment of *g* to the ability trained; this implies that both domain-specific and general abilities corroborate to produce cognitive change. Second, there was transfer (or effect) from QR ability to CE, but not vice versa. This finding is contrary to Hypothesis 3, although it should be further investigated. The paths from QR to CE factors were stronger in the CETG, which means that in this group students used their QR ability in order to solve the CE tasks. However, the correlation between PostQR and PostCE (.405) was weaker than the correlation between the PreQR-PreCE factors (.627), which indicates that the training of the CE ability led to a relative independence from the QR ability. This finding suggests that solving a problem may make use of processes/abilities that are more appropriate for a different domain

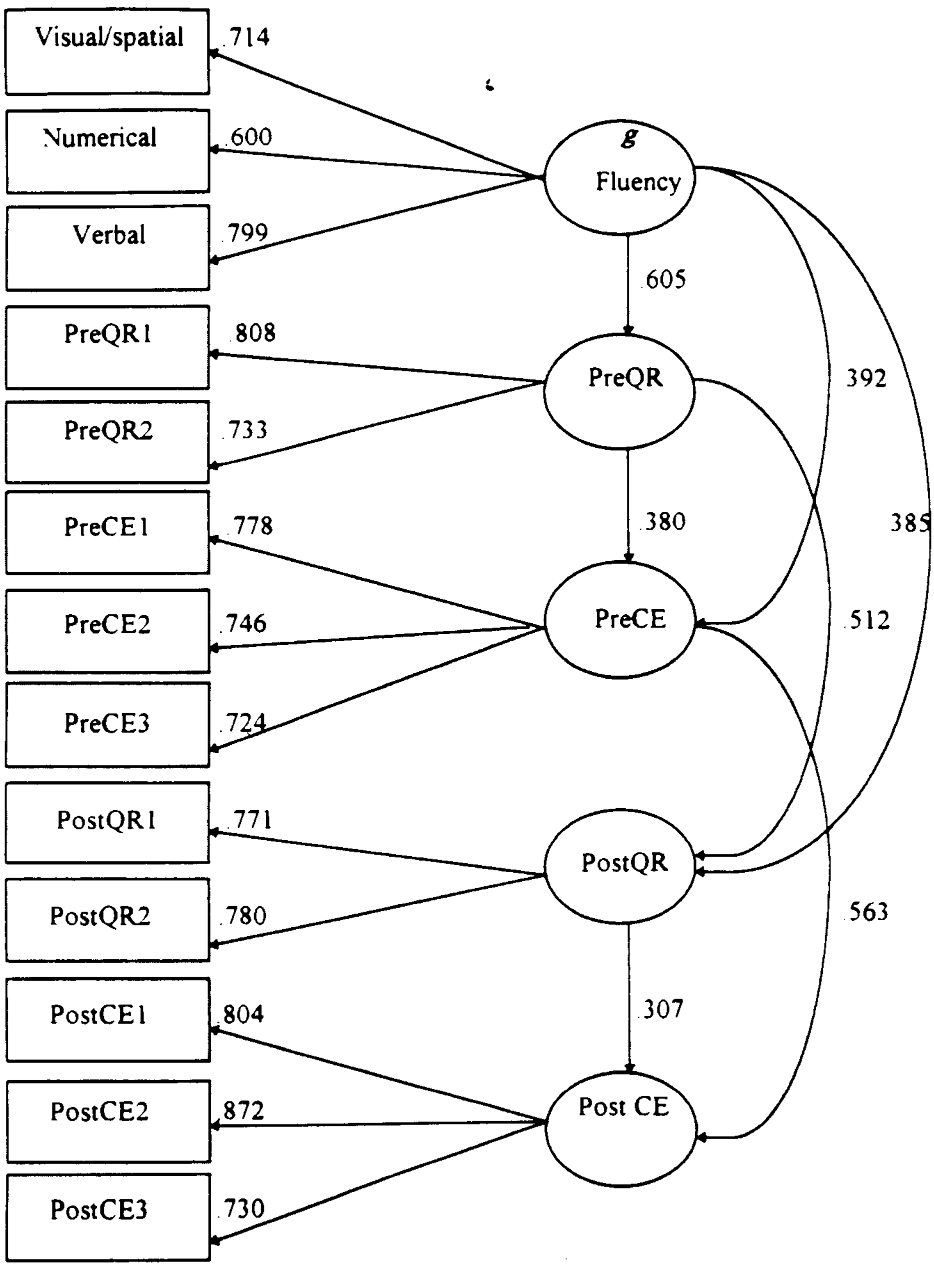


Figure 1a
The structural model best fitting the data of the QRTG

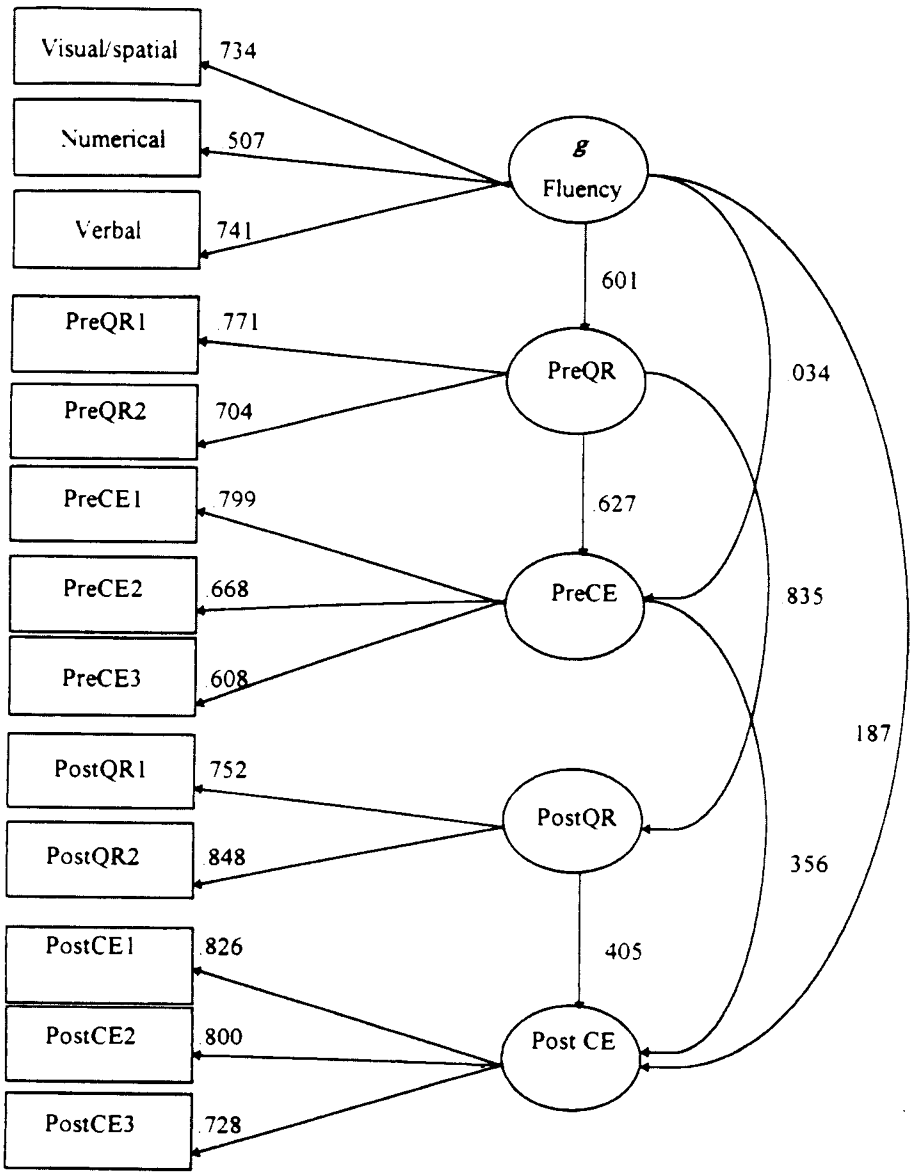


Figure 1b
The structural model best fitting the data of the CETG

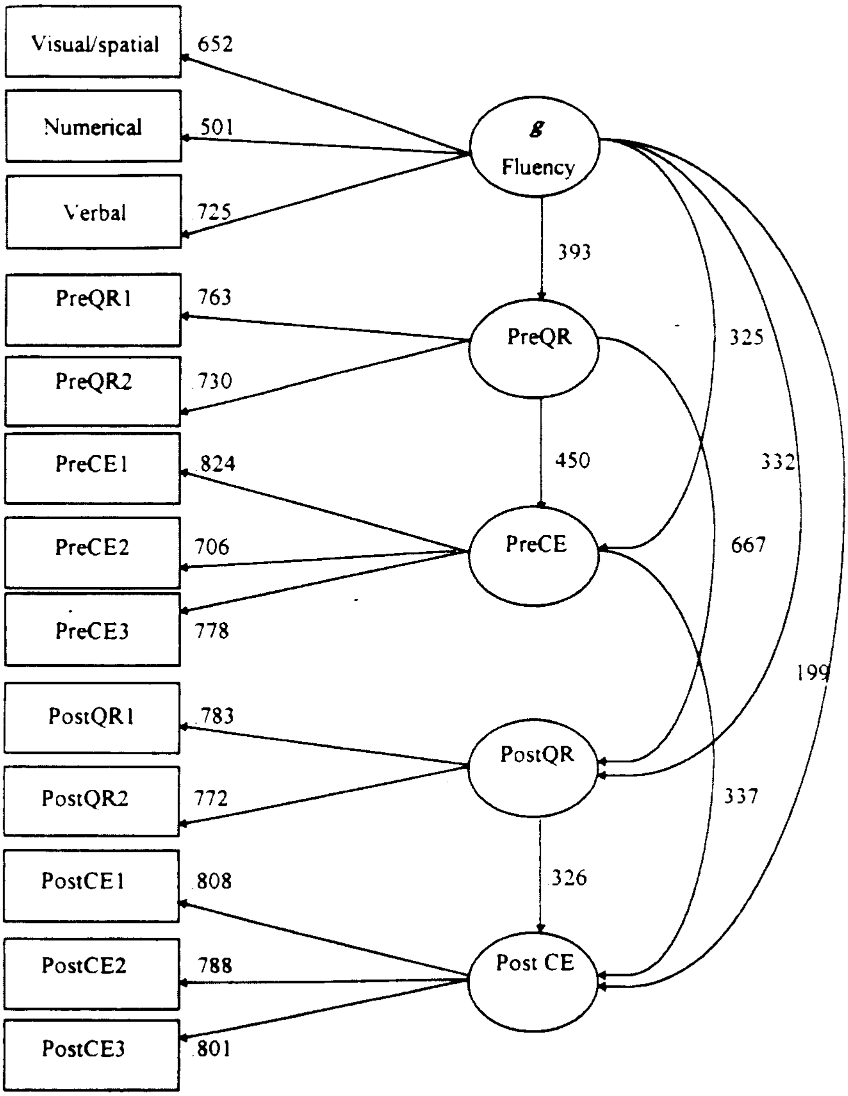


Figure 1c
The structural model best fitting the data of the CTG

of knowledge. However, training leads to differentiation and specialization of thought structures. In this endeavour *g* makes a contribution, particularly for the QR ability.

Having established that *g* forms part of the mechanism of cognitive change, we went on to testing Hypothesis 2, which regarded the sheer influence of verbal/ semantic, visual-spatial, and numeric abilities on the posttest factors. To do this we tested the previous models, only that we omitted the second-order *g* factor and instead used the first-order verbal/semantic, visual-spatial, and numeric factors as predictors of pretest and posttest performance on QR and CE tasks. These models were not confirmed. Therefore it was *g* rather than each of the lower-order abilities per se that intervenes and influences performance.

Finally, the same structural modeling approach was used in order to test Hypothesis 4. This hypothesis regarded the interaction of training method (i.e., training form) with *g*. Specifically, if the form of training applied (algorithmic, metacognitive, and computer-assisted) differentiated the effects of *g* on posttest factors. In this sake we tested the already identified model for each treatment group in the three sub-groups of training form within each treatment group.

In the case of *algorithmic training* it was found that in the QRTG the effect of *g* on the PostQR was retained whereas in the CETG it was not (it was non-significant). The fit indices of the models were: *For the QRTG:* $x^2(58)=73.058$, $p=.09$; NFI=.933, NNFI=.980, CFI=.985. *For the CETG:* $x^2(58)=70.717$, $p=.12$; NFI=.930, NNFI=.982, CFI=.986.

Metacognitive training in the QRTG relied on *g* for posttest performance. However, in the *CETG* the effect of g on the PostCE factor was marginally significant [$x^2(1)=3.76$, $p=.052$]. The fit indices of the models were: *For the QRTG:* $x^2(58)=59.007$ $p=.44$; NFI=.931, NNFI=.998, CFI=.999. *For the CETG:* $x^2(58)=74.572$, $p=.07$; NFI=.913, NNFI=.971, CFI=.979.

Finally, *C-A training* made use of *g* in both the QRTG and CETG. The fit indices of the models were: *For the QRTG:* $x^2(58)=72.409$, $p=0\ .09$; NFI=.893, NNFI=.967, CFI=.976. *For the CETG:* $x^2(55)=54.146$, $p=0\ .51$; NFI=.882, NNFI=1.003, CFI=1.000.

In order to further test the effect of the lower order factors (namely the verbal/semantic, visual-spatial, and numeric) on posttest performance and their possible interaction with the form of training, the previous models were now modified so as to include paths relating the symbolic variables with the rest of the variables, or covariances between the residuals of the symbolic variables and those of the latent factors. In the main, these effects were small and inconsistent across the various subgroups. The general trends that emerged, however, were: first, PreCE performance was related to the numeric fluency variable. This finding is in accordance with the PreQR-PreCE relationship, and it means that students who were fluent with arithmetic operations, tended to use this ability when they processed the CE tasks. That is, they focused on the numerical aspect of the tasks.

Second, in algorithmic training both PreQR and PostQR performance was found to be related to numeric fluency, as expected. Third, there was an unexpected finding in the case of computer-assisted-training-groups. It was found that in the QRTG, PostQR and PostCE performance was related to the verbal/semantic fluency variable. This direct effect of the verbal/semantic variable on performance variables rather than on the respective QR and CE factors, weakened the effect of the *g* factor on the PostQR factor which in this case was non significant. In the CETG, the verbal/semantic fluency variable influenced only the PostCE variables; the PostQR variables were related to the numeric fluency variable. Again the role of the *g* factor on the PostCE factor was weakened. In metacognitive training there was no effect of the symbolic variables on posttest performance in either treatment group.

These findings imply that the students in the computer-assisted training used their verbal/semantic ability to handle the information provided. Furthermore, although this verbal "integration" was used in relation to the ability trained (QR or CE in the respective treatment groups), it was generalized from QR to CE tasks in the case of QRTG but not from CE to QR tasks in the CETG. This indicates that verbal processing was more appropriate for the CE tasks, and it did not transfer to QR tasks (in the CETG), where numeric fluency is most appropriate. Of course, verbal integration or verbal strategies may be used in the processing of quantitative (mathematical) tasks if the method of training demands it (see also Kaizer & Kranzler, 1995); however, it was not necessary for the processing of QR tasks, and that is why it was not generalized from CE to QR tasks in the CETG.

Therefore, lower order *g*-related symbolic fluency factors play a role in the training of abilities and interact with both the ability trained and/or the training method used, as suggested by Hypotheses 2 and 4.

Effectiveness of training

Structural modeling analyses revealed the mechanism of cognitive change, but did not show details regarding the effectiveness of training. To investigate the effects of training, in general, and training form, in particular, a series of ANOVAs were performed. First, the 3[Treatment group (QRTG, CETG, CTG)] X 2[Testing (pre- and posttest)] X 2[Ability (QR and CE)] X 4(Task) MANOVA with the last three factors as within subjects factors showed a significant interaction of the four factors: $F(6, 3285)=4.57$, $p<.0001$ which suggests that the three treatment groups were differentiated in their performance as a function of testing (pre- and posttest), the ability on which they had been trained, and the task. For a detailed presentation of the results and the effect sizes see Efklides (in 1999).

In order to further scrutinize our data and determine the effects of form of training, and whether it interacted with the level of training (that is, level I or II) and with person variables such as age (or gender) a 3(Age) X 3(Form of training) X 2(Level of training) X 2(Testing) X 8(Task) MANOVA was performed within each treatment group.

In the QRTG, level of training was found to interact with age, form of training, and task, $F(12,1203)=1.86$, $p=.036$ and with age, testing (pre-, posttest) and task, $F(6,1203)=2.96$, $p=.007$. The interaction of level with testing and task was highly significant, $F(3,1203)=73.59$, $p=.000$. In the CETG, level of training interacted with testing and task, $F(3,1182)=41.65$, $p=.000$ but it did not interact with age or form of training.

Since the analyses indicated that level of training was a significant factor in the determination of performance, the MANOVAs regarding the effect of the form of training were applied separately, firstly, on the group of subjects who received Level I training and, second, on the group of subjects who received Level II training. Furthermore, in order to be able to identify the exact effect of training form and the possible transfer from the trained to the non-trained ability, we selected (within each level of training) those subjects who scored similarly in the two SSSs at the pretest criterion tasks, i.e., the QR2 and CE2 tasks. This was deemed necessary, because subjects had been appointed to the various treatment groups according to their performance on the SSS trained, regardless of their performance on the other SSS. These data are given in Tables 2 and 3.

As can be seen in Table 2, training QR and CE abilities in the respective treatment groups led to significant gains in the level 1 and 2 tasks, namely QR1, QR2, and CE1, CE2 tasks; the effect was stronger with level 2 than level 1 tasks. This means that the training provided transferred more readily to the task similar to the one used in the training. Improvement of performance on the respective level 1 task (QR1 or CE1) indicates a transfer of training to lower level tasks. It should also be noted that training led to relative improvement of QR3 and CE3 tasks, but the overall performance of students who received Level I training on these tasks was very low.

Table 2
Mean performance of subjects who received Level I training as a function of ability, testing, and task (Subjects were matched for their pretest cognitive level in both abilities)

		Pretest tasks				Posttest tasks			
Treatment Group	N	1	2	3	4	1	2	3	4
		Ability QR							
QRTG	210	.838	.367	.286	.219	1.271	1.067	.400	.262
		(.871)	(.483)	(.453)	(.415)	(.874)	(.839)	(.605)	(.492)
CETG	204	.809	.382	.337	.176	.985	.755	.348	.245
		(.847)	(.487)	(.473)	(.382)	(.890)	(.797)	(.594)	(.475)
CTG	99	.889	.333	.323	.061	.859	.697	.404	.222
		(.868)	(.474)	(.470)	(.240)	(.869)	(.788)	(.588)	(.418)
		Ability CE							
QRTG	210	.814	.576	.157	.033	1.090	.976	.786	.190
		(.776)	(.495)	(.365)	(.180)	(.822)	(.728)	(.767)	(.491)
CETG	204	.917	.637	.186	.049	1.069	1.118	.490	.137
		(.811)	(.482)	(.390)	(.216)	(.803)	(.733)	(.669)	(.410)
CTG	99	.778	.556	.293	.111	.909	.929	.424	.192
		(.790)	(.499)	(.457)	(.303)	(.905)	(.786)	(.656)	(.444)

Inspection of Table 3 confirms the above findings in the Level II training groups, namely the immediate transfer to the QR4 and CE4 tasks. In the CETG there was also a significant improvement of performance of CE3 task; there was no similar transfer in the QRTG as regards the QR3 performance. It should be pointed out that training level 4 tasks led to relative decrease of performance on level 1 (QR1/CE1) and 2 (QR2/CE2) tasks, which may be due to lack of interest in them, once attention was directed to highly complex tasks.

Finally, as regards transfer of training from one ability to the other it seems that training the QR SSS improved performance on CE tasks; this is indicative of transfer from QR to CE SSS but not the other way round, although ÇETG subjects also improved to QR SSS more than the CTG subjects.

As regards the effect of form of training, it was found that in Level I training, the main effect of form of training was significant, $F(2,408)=3.42$, $p=.034$. There was also a form of training by gender, testing, and task interaction, $F(14,2814)=1.80$, $p=.034$. Form of training did not interact with age. For Level II training, there was no main effect or interaction of form of training. Tables 4 and 5 present the data regarding the form of training.

As shown in Figures 2a, 2b, 2c and 2d, in Level I algorithmic training was about equally effective with metacognitive training whereas C-A training was less effective. In Level II, C-A training continued to be less effective than either the algorithmic or the metacognitive.

It can be concluded, then, that Hypothesis 5 was verified, since the less demanding training forms led to more cognitive change than C-A training.

Table 3
Mean performance of subjects who received level II training as a function of ability, testing, and task (Students were matched for their pretest cognitive level in both abilities)

		Pretest tasks				Posttest tasks			
Treatment Group	N	1	2	3	4	1	2	3	4
					Ability QR				
QRTG	71	1.944	1.887	1.000	.254	1.887	1.732	1.127	.930
		(.232)	(.318)	(.862)	(.438)	(.318)	(.585)	(.887)	(.900)
CETG	75	1.840	1.933	.933	.493	1.760	1.800	.893	.733
		(.466)	(.342)	(.859)	(.760)	(.566)	(.403)	(.879)	(.723)
CTG	44	1.864	1.932	1.023	.386	1.705	1.477	1.273	.568
		(.409)	(.334)	(.792)	(.493)	(.701)	(.762)	(.817)	(.818)
					Ability CE				
QRTG	71	1.648	1.704	1.437	.732	1.775	1.704	1.310	1.070
		(.612)	(.571)	(.788)	(.910)	(.566)	(.545)	(.821)	(.931)
CETG	75	1.853	1.840	1.067	.213	1.813	1.653	1.507	1.053
		(.392)	(.404)	(.811)	(.412)	(.512)	(.604)	(.724)	(.884)
CTG	44	1.727	1.636	1.136	.273	1.523	1.500	1.159	.727
		(.451)	(.613)	(.878)	(.451)	(.698)	(.591)	(.805)	(.788)

Discussion

Our research aimed to study the role of general and domain-specific abilities in cognitive change. The results showed that the mechanism of cognitive change involves both general intelligence *(g)* and domain-specific abilities. What is more interesting, however, is the finding that *g* interacted with the instructional method used to induce cognitive change. Algorithmic training made the least demands on *g* whereas computer-assisted training the most. Yet, of three aspects of *g* which were used in this study, namely the narrow factors that correspond to verbal/semantic, visual-spatial, and numeric fluency, algorithmic training in the QRTG made use of the numeric ability whereas in the CETG of the verbal/semantic. Metacognitive training did not rely on any of the three narrow abilities, and C-A training relied exclusively on the verbal/semantic ability. These findings need further clarification if we are to answer questions pertaining to the mechanism of cognitive change and transfer of training.

As regards the mechanism of cognitive change, it was found that *g* was involved in both pretest and posttest performance. When training was provided, *g* was focused on the ability trained rather than the one non-trained. Posttest performance also depended on the respective domain-specific ability although, quite unexpectedly, CE performance also depended on the QR ability. In the case of QR ability, cognitive change was related only to the QR SSS and not to the CE SSS. These findings are very important because they show that the mechanism of cognitive change is not necessarily contained in single domains. It may

Table 4
Mean performance as a function of treatment group, form of training, ability, testing, and task of subjects who perceived level I training
(Students were matched as to their pretest cognitive level in the two abilities)

			Pretest				Posttest			
Treatment Group	Training form	N	QR1	QR2	QR3	QR4	QR1	QR2	QR3	QR4
			Ability QR							
QRTG	Alg	79	.810	.342	.354	.215	1.266	1.076	.443	.304
			(.878)	(.477)	(.481)	(.414)	(.873)	(.888)	(.635)	(.540)
	Met	82	.780	.402	.268	.244	1.317	1.061	.390	.232
			(.875)	(.493)	(.446)	(.432)	(.859)	(.822)	(.643)	(.479)
	C-A	49	.980	.347	.204	.184	1.204	1.061	.347	.245
			(.854)	(.481)	(.407)	(.391)	(.912)	(.801)	(.481)	(.434)
CETG	Alg	79	.722	.329	.380	.152	1.013	.709	.278	.177
			(.816)	(.473)	(.488)	(.361)	(.899)	(.787)	(.530)	(.384)
	Met	84	.798	.405	.298	.214	.917	.750	.345	.262
			(.833)	(.494)	(.460)	(.413)	(.881)	(.790)	(.478)	(.469)
	C-A	41	1.000	.439	.317	.146	1.073	.854	.488	.341
			(.922)	(.502)	(.471)	(.358)	(.905)	(.792)	(.711)	(.617)
			Ability CE							
QRTG	Alg	79	.772	.595	.177	.051	1.101	.949	.544	.165
			(.784)	(.494)	(.384)	(.221)	(.841)	(.714)	(.765)	(.436)
	Met	82	.720	.512	.098	.000	1.061	.915	.610	.195
			(.805)	(.503)	(.299)	(.000)	(.791)	(.724)	(.766)	(.531)
	C-A	49	1.041	.653	.224	.061	1.122	1.122	.612	.224
			(.676)	(.481)	(.422)	(.242)	(.857)	(.754)	(.786)	(.511)
CETG	Alg	79	.785	.582	.152	.038	.873	1.101	.506	.089
			(.795)	(.496)	(.361)	(.192)	(.838)	(.744)	(.677)	(.328)
	Met	84	.976	.619	.107	.012	1.143	1.143	.405	.143
			(.791)	(.489)	(.311)	(.109)	(.763)	(.747)	(.604)	(.443)
	C-A	41	1.049	.780	.415	.146	1.293	1.098	.634	.220
			(.865)	(.419)	(.499)	(.358)	(.750)	(.700)	(.767)	(.475)

Table 5
Mean performance as a function of treatment group, form of training, ability, testing, and task of subjects who perceived level II training
(Students were matched as to their pretest cognitive level in the two abilities)

			Pretest				Posttest			
Treatment Group	Training form	N	QR1	QR2	QR3	QR4	QR1	QR2	QR3	QR4
			Ability QR							
QRTG	Alg	16	2.000	1.938	.938	.250	1.875	1.750	.875	1.000
			(.000)	(.250)	(.929)	(.447)	(.342)	(.577)	(.885)	(.966)
	Met	15	2.000	1.867	.600	.267	2.000	1.733	1.000	.600
			(.000)	(.352)	(.910)	(.458)	(.000)	(.594)	(.845)	(.828)
	C-A	18	1.944	1.833	.944	.222	1.722	1.500	.778	.500
			(.236)	(.383)	(.873)	(.428)	(.461)	(.786)	(.878)	(.786)
CETG	Alg	27	1.926	2.000	.815	.111	1.741	1.741	.704	.667
			(.267)	(.000)	(.786)	(.320)	(.594)	(.447)	(.823)	(.679)
	Met	17	1.765	2.000	.765	.235	1.882	1.824	.706	.529
			(.664)	(.000)	(.831)	(.437)	(.332)	(.393)	(.849)	(.624)
	C-A	17	1.765	1.882	.765	.294	1.529	1.706	1.000	.941
			(.562)	(.485)	(.903)	(.470)	(.800)	(.470)	(.935)	(.827)
			Ability CE							
QRTG	Alg	16	1.375	1.875	1.000	.250	1.875	1.687	1.125	.688
			(.806)	(.342)	(.894)	(.447)	(.342)	(.479)	(.957)	(.793)
	Met	15	1.533	1.467	1.333	.000	1.333	1.400	1.067	.733
			(.640)	(.743)	(.816)	(.000)	(.900)	(.737)	(.799)	(.884)
	C-At	18	1.611	1.611	1.278	.222	1.833	1.722	1.111	.833
			(.608)	(.608)	(.826)	(.428)	(.514)	(.575)	(.832)	(.985)
CETG	Alg	27	1.926	1.778	.963	.148	1.926	1.630	1.630	1.074
			(.267)	(.506)	(.759)	(.362)	(.267)	(.629)	(.492)	(.874)
	Met	17	1.824	1.941	.824	.176	1.765	1.529	1.529	.882
			(.393)	(.243)	(.883)	(.393)	(.562)	(.624)	(.717)	(.857)
	C-At	17	1.824	1.882	1.176	.176	1.706	1.706	1.353	1.059
			(.529)	(.332)	(.809)	(.393)	(.686)	(.686)	(.931)	(.899)

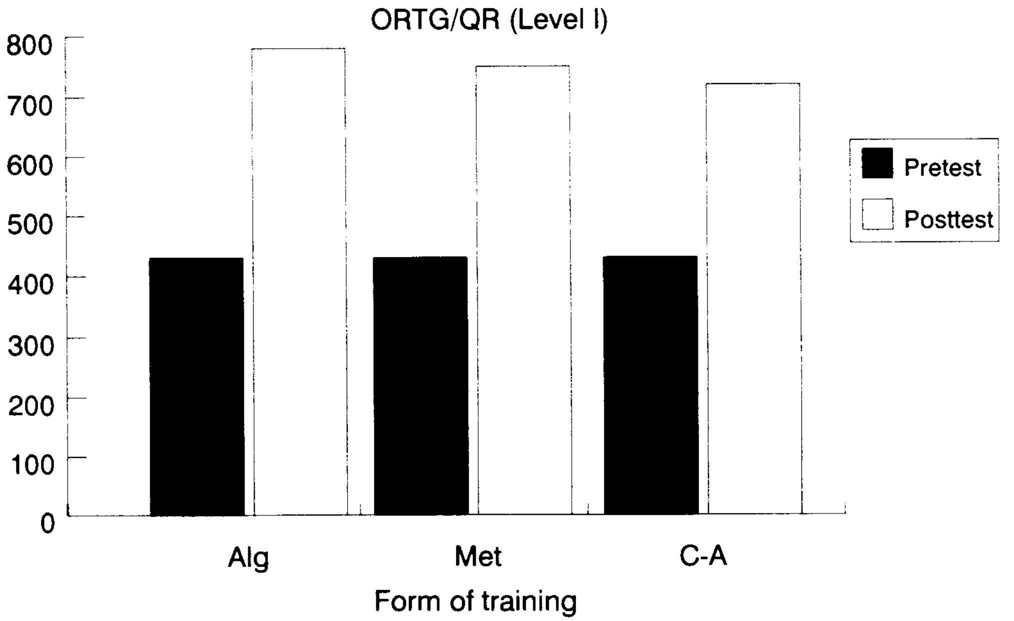


Figure 2a
Mean pre-and posttest performance as a function of form of training for level I and Level II training

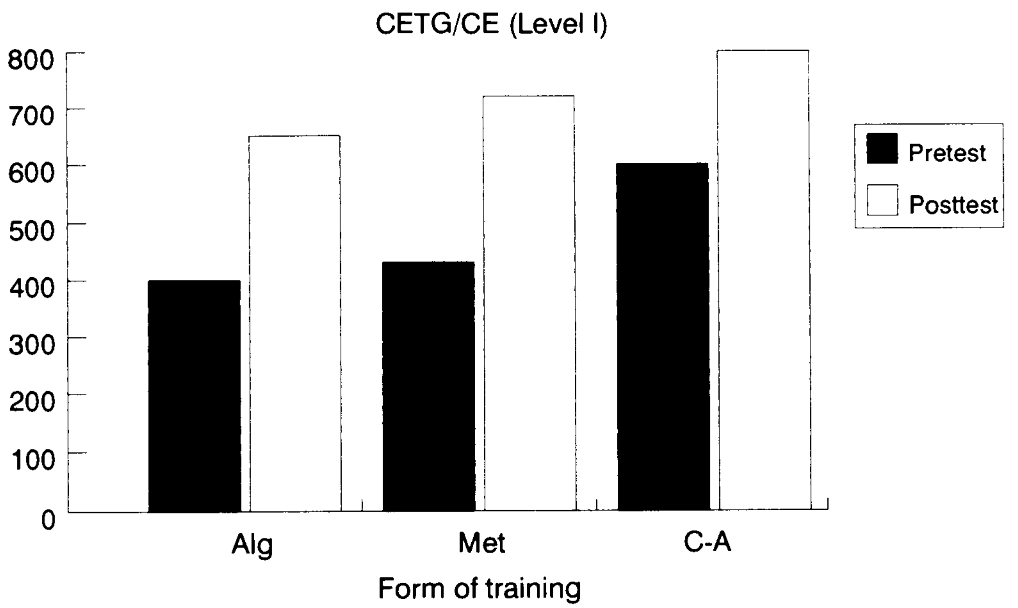


Figure 2b
Mean pre-and posttest performance as a function of form of training for Level I and Level II training

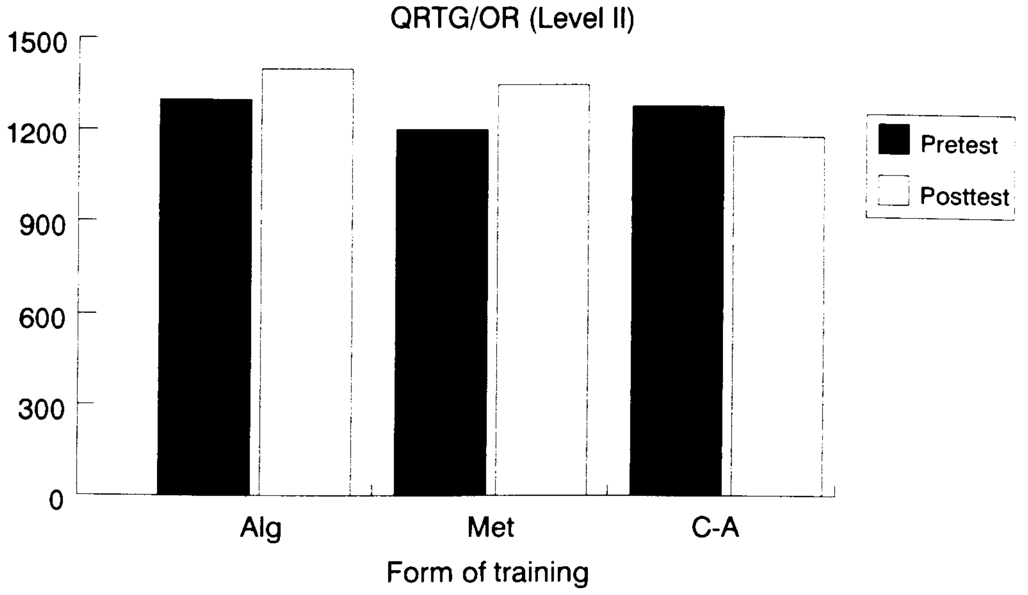


Figure 2c
Mean pre-and posttest performance as a function of form of training for level I and Level II training

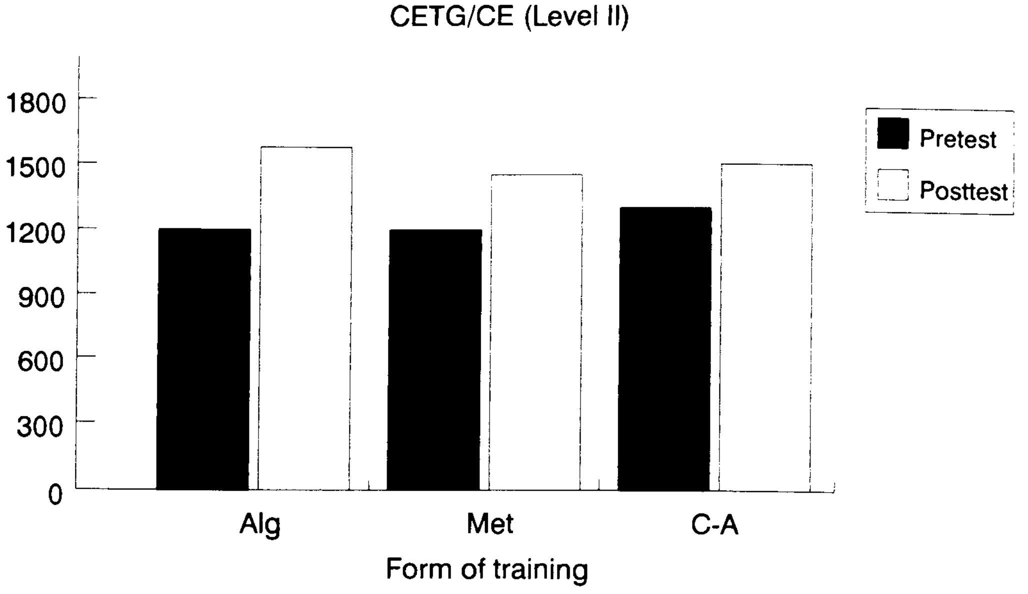


Figure 2d
Mean pre-and posttest performance as a function of form of training for level I and Level II training

and it may not.

However, the mechanism of cognitive change cannot be fully understood unless we introduce into the picture the middle-level, narrow abilities related to *g*, namely the symbolic fluency factors (see also Holmberg & Gustafsson, 1993). What is worth noting is that cognitive change also involved the numeric or verbal/semantic fluency, depending on the training method used.

As stated above C-A training relied on verbal/semantic fluency both in the QRTG and CETG. This implies that change of QR ability in C-A training was mediated by verbal processes rather than numeric. Therefore even QR ability may make use of non-domain-specific symbolic abilities if the instructional method demands it. It could be argued then when training is provided, *g* is used for the understaing of the training material and its transformation and adaptation to the domain-specific procedures and/or task requirements. This transformation is facilitated if the instructional method makes explicit the procedure to be used and the symbolic system best fit for its application. Algorithmic training succeeded in doing this, whereas metacognitive training worked at the procedural level without use of any particular symbolic system or ability. Computer-assisted training relied on the verbal/semantic symbolic system for the understand and integration of the training material at a cost for the procedural part of thinking in the QRTG.

We come now to the transfer of training issue. Our data showed that training the QR ability transferred to CE ability more than training the CE ability. Control group subjects also improved in their posttest performance but less so than the groups that received training in either ability. This is indicative of transfer of training from the ability trained to the non-trained.

There is one point though that needs to be discussed, and this is why QR training transferred to CE ability more than the other way round. Our data showed that even pretest CE ability depended on the respective QR ability. This probably means that CE ability was not so well formed and students processed CE tasks as numeric ones via QR ability. This reliance on QR ability continued after training, only that in the CETG it was not as strong as in the pretest. Therefore training the CE ability led to its relative independence from the QR ability and to its formation as distinct one with its own procedural and symbolic character. From this point of view training the CE ability worked as inhibitory of the transfer from the QR ability. This effect sounds as "contra-transfer", and evidently reflects another aspect of cognitive functioning. If transfer facilitates integration of cognitive structures, contra-transfer facilitates cognitive differentiation and individuation abilities.

Consequently, training works in two ways: one is to boost thinking both in the domain trained and non-trained and the other is to differentiate thought structures so that they become more tuned to the domain and task at hand. These effects can be enhanced or moderated depending on the training method used. More research is needed in order to clarify this issue as well as the role of individual differences in the ways through which cognitive change is induced.

Finally, the effect of level of thinking should be mentioned. In our study, level of training was determined by the pretest cognitive level of the students and, specifically, if they had achieved (Level II) or not (Level I) the relational level of thinking (that is, thinking that corresponds to the QR2 and CE2 task). Students were trained either one or two levels of thinking above theirs. What was found was that near-transfer, that is to the task corresponding to the training task, was higher than to either lower level or higher level tasks (for Level I training). Still, there was transfer to the other tasks of the same ability, which means that there was transfer within the trained ability and the non-trained ability (far-transfer). This finding has a bearing on instruction, because it suggests that instruction at a relatively higher (more complex) level than students' current one may generalize to lower (less complex) tasks. However, for very advanced students, instruction at a highly complex level may prevent them from paying adequate attention to simple tasks, which are well within their grasp.

In conclusion, this study provided evidence for transfer of training but we should be cautious of the effects of instructional method with regard to the desired results.

References

Bentler, P. M. (1993). EQS: *Structural equations program manual.* 2nd edition. Los Angeles, CA: BMDP Statistcal Software.

Brainerd, C. J. (1983). Varieties of strategy training in piagetian concept learning. In M. Pressley & J. R. Levin (Eds.), *Cognitive strategy research: Educational applications* (pp. 3-27). New York: Springer-Verlag.

Baron, J., & Sternberg, R. (Eds.) (1987). *Teaching thinking skills: Theory and practice.* New York: Freeman.

Carey, S. (1985). Are children fundamentally different kinds of thinkers of causal reasoning? In W. Friedman (Ed.), *The developmental psychology of time* (pp. 209-254). New York: Academic.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analysis studies.* Cambridge, UK: University of Cambridge Press.

Case, R. (1985). *Intellectual development: Birth to adulthood.* New York: Academic.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121-151.

Demetriou, A., & Efklides, A. (1981). The ideal of the whole and the reality of the parts. In J. A. Meacham & N. R. Santilli (Eds.), *Social development in youth: Structure and content* (pp. 20-46). Basel, Switzerland: Karger.

Demetriou, A., & Efklides, A. (1987). Experiential Structuralism and neo-piagetian theories: Toward an integraded model. *International Journal of Psychology, 22,* 679-728.

Demetriou, A., & Efklides, A. (1994). Structure, development, and dynamics of mind: A meta-piagetian theory. In A. Demetriou & A. Efklides (Eds.), *Intelligence, mind, and reasoning* (pp. 75-109). Amsterdam: Elsevier.

Efklides, A. (1999). Training domain-specific abilities: The case of Experiential Structuralism. In J. H. M. Hamers, H. van Luit, & B. Csapo (Eds.), *Thinking skills and teaching thinking* (pp. 105-129). Lisse, The Netherlands: Swets & Zeitlinger.

Efklides, A., Demetriou, A., & Gustafsson, J.-E. (1992). Training, cognitive change and individual differences. In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *Neo-piagetian theories of cognitive development: Implications and applications for education* (pp. 122-143). London: Routledge.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for Kit of Factor-referenced Cognitive Tests.* Princeton, NJ: Educational Testing Service.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 8,* 477-531.

Fischer, K. W., & Farrar, M. J. (1988). Generalization about generalization: How a theory of skill development explains both generality and specificity. In A. Demetriou (Ed.), *The neo-piagetian theories of cognitive development: Toward an integration* (pp. 137-171). Amsterdam: North-Holland.

Goosens, L. (1992). Training scientific reasoning in children and adolescents: A critical commentary and quantitative integration. In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *Neo-piagetian theories of cognitive development: Implications and applications for education* (pp. 160-182). London: Routledge.

Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8,* 179-203.

Gustafsson, J.-E. (1988). Hierarchical models of individual differences in cognitive abilities. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4, pp. 35-71). Hillsdale, NJ: Erlbaum.

Gustafsson, J.-E., Lindstrom, B., & Bjorck-Akesson, E. (1981). *A general model for the organization of cognitive abilities.* Report from the Department of Education, University of Goeteborg, Sweden.

Hamers, J. H. M., & Overtoom, M. (Eds.). (1997). *Inventory of european programmes for teaching thinking.* Utrecht, The Netherlands: Sardes.

Holmberg, L. M., & Gustafsson, J.-E. (1993). Efficiency of program handling as a function of verbal and iconic interfaces and individual differences in ability. Special issue: Swedish research on learning and instruction with computers. *Computer in Human Behavior, 9,* 227-245.

Kaizer, C., & Shore, B. M. (1995). Strategy flexibility in more and less competent students on mathematical word problems. *Creativity Research Journal, 8,* 77-82.

Keil, F. (1984). Mechanisms in cognitive development and the structure of knowledge. In R. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 81-99). New York: Freeman.

Kuhn, D. (Ed.). (1990). *Developmental perspectives on teaching and learning thinking skills. Contributions to Human Development* (Vol. 21). Basel, Switzerland: Karger.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills.* San Diego, CA: Academic Press.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51,* 77-101.

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