Developmental Psychology
1993, Vol. 29, No. 3, 480–497

0012-1649/93/$3.00

Structure and Development of Causal–Experimental Thought: From Early Adolescence to Youth

Andreas Demetriou, Anastasia Efklides, Maria Papadaki, Georgia Papantoniou, and Andreas Economou

This study investigated the structure and development of causal–experimental thought. Ss (N = 260) from 12 to 16 years of age were examined by 3 test batteries. The batteries involved items addressing combinatorial, hypothesis handling, experimentation, and model construction abilities. Confirmatory factor analysis indicated that these abilities, although distinct from each other, do share a common functional core, and they are organized in a higher order causal–experimental specialized structural system. Rasch scaling indicated that these 4 kinds of abilities follow overlapping developmental trajectories. Saltus modeling suggested that development is continuous rather than discontinuous. Individual differences were found in the rate of acquisition but not in the structure of abilities. A model of synergic developmental causality was proposed to account for this pattern of development.

This article presents a study designed to uncover the structure and development of causal–experimental thought. The design of this study was inspired by experiential structuralism, which is our theory of the organization and development of human cognition (Demetriou, in press; Demetriou & Efklides, 1988, 1989; Demetriou, Efklides, & Platsidou, 1993). The study is part of an extended research project that aims to detail the nature and interpatterning of the various systems and subsystems that, according to the theory, constitute the human developing mind. Thus, in the introduction, the general premises of the theory are first summarized. The study on the structure and development of the causal–experimental thought is then presented.

Architecture of Developing Mind

Basic Premises of the Theory

Our theory postulates that cognition develops across three fronts. The first involves a set of *Specialized Structural Systems* (SSSs) that enable the person to represent, mentally manipulate, and understand specific domains of reality and knowledge. Five SSSs have been identified: (a) the *qualitative–analytic,* (b) the *quantitative–relational,* (c) the *causal–experimental,* (d) the *spatial–imaginal,* and (e) the *verbal–propositional.* These SSSs specialize in the representation and processing of categorical, quantitative, causal, spatial–figural, and syntactic–formal structures, respectively.

Each of these SSSs is considered to be a cohesively organized universe of concepts and operations related to a broad domain of reality and knowledge. The formation and preservation of the SSSs are considered to be governed by a set of five organizational principles: namely, the principles of *domain specificity, procedural–computational specificity, symbolic bias, subjective distinctness,* and *developmental variation* of concepts and operations. That is, it is assumed that if several component abilities are concerned with the same reality domain, possess the same formal and procedural properties, tend to be represented through the same symbol system, and are felt or recognized by the thinker himself or herself as being similar, then they will be coordinated to form an SSS (Demetriou, in press; Demetriou & Efklides, 1988). The specific status of the various SSSs may differ for each individual, depending on his or her specific activities, which may connect him or her with some aspects of reality and knowledge more than with others.

The second front is a system of higher order control structures that enable the person to understand and regulate both his or her own cognitive activity and interactions with the environment and with other persons. In our terms, this is the hypercognitive system (see Demetriou et al., 1993, for the advantages of the term *hypercognitive* over the term *metacognitive*). At a macroprocessing level, this system involves three distinct components. First, it includes a model of *cognitive organization and functioning,* which involves representations of different cognitive functions (e.g., memory and attention) and SSSs, as well as prescriptions about the use of these functions in relation to different kinds of problems and SSSs. Second, it involves a model of *intelligence,* which organizes knowledge and beliefs about what is intelligent in a given environment and includes decision rules about how one may act intelligently. Finally, it includes a model of *cognitive self,* which involves the person's

---

Andreas Demetriou, Anastasia Efklides, Maria Papadaki, Georgia Papantoniou, and Andreas Economou, Department of Psychology, School of Philosophy, Aristotelian University of Thessaloniki, Thessaloniki, Greece.

This article was written when Andreas Demetriou was a Visiting Scholar at the School of Education, Stanford University, in June–July, 1991. The studies reported in this article were supported by a grant provided by the Greek General Secretariat of Research and Technology to Andreas Demetriou and Anastasia Efklides (87 EΔ 178).

We especially thank Robbie Case for his comments on an earlier version of this article. The constructive suggestions of four anonymous reviewers were very helpful in our attempt to strengthen the article.

Correspondence concerning this article should be addressed to Andreas Demetriou, Department of Psychology, School of Philosophy, Aristotelian University of Thessaloniki, Thessaloniki 540 06, Greece. Electronic mail may be sent to bczz04@grtheunl.

representations and preferences about himself or herself as a cognitive and intelligent being. At a microprocessing level, the hypercognitive system involves processes that enable the person to activate the appropriate SSS and the right SSS-specific components depending on the requirements of the task at hand.

The third front may be conceived as the system in which information is represented and processed by the thinker to make sense of and, if required, go beyond that information so as to produce solutions to problems. This system involves three basic components: speed of processing, control of processing, and working memory. *Speed of processing* basically refers to the minimum speed required if a given mental act is to be efficiently executed. *Control of processing* refers to a mechanism that functions under the guidance of the task goal, like a filter, permitting only goal-relevant schemes to enter processing space. *Working memory* refers to the maximum number of schemes that the person can process at the same time.

The systems making up the three fronts are considered to be at constant interplay during micro- and macrodevelopment. The processing system is considered a unit of finite potential, which constrains the complexity and quality of the information structures the intellect can represent and process at a given moment in its development. The hypercognitive system constrains the manner in which the processing system and the SSSs are used in response to specific environmental demands. Thus, it shapes the quality of the specific skills and concepts that are constructed within any specific structural system. The SSSs are incessantly shaped out of the interaction with specific domains of the environment, and they provide the context in which the processing system is used, as well as the basis for building ever more complex hypercognitive skills and strategies (see Demetriou, in press; Demetriou & Efklides, 1989; Demetriou et al., 1993).

In fact, the theory postulates that development is possible precisely because of the multistructural and multisystemic nature of mind. Specifically, it assumes that a change in any of the three kinds of systems just described can, under appropriate circumstances, propagate rapidly within the system and initiate a series of changes in any of the other systems. This is so because the systems are functionally tuned to each other. Therefore, a change in any of them is a disturbance factor that puts the dynamic tuning of the whole system in jeopardy. The direction of change is dictated by the system that changes first. That is, this system will tend to pull the other systems in the direction toward which it has already moved (Demetriou et al., 1993). We hoped that the present study would shed light on how synergic developmental causality may operate.

Figure 1 shows how the cognitive system is represented by our theory in terms of the figural conventions of structural modeling. It can be seen that performance on any task is thought to be conjointly determined by forces peculiar to (a) the processing system (represented by the PS factor); (b) the common core that marks a given SSS as distinct from other SSSs (represented by the ${}_{cr}SSS_{qr}$ or the ${}_{cr}SSS_{ce}$ factors; the symbols "cr," "qr," and "ce" stand for the common core of an SSS, the quantitative–relational SSS, and the causal–experimental SSS, respectively, throughout this article); and (c) the component (cm) skills that define subsystems of ability within the context of an SSS (represented by the cm factors). This includes, for example, the combinatorial or the experimentation skills in the causal–experimental SSS or the ability to execute arithmetic operations or grasp proportional relations in the quantitative–relational SSS (represented by the four ability-specific factors). At the same time, the SSS-specific component abilities are intercoordinated under the pressure of the organizational principles. The intercoordination at this level ensures the concerted functioning of the SSSs as such. The second-order ${}_{pr}SSS_{qr}$ or ${}_{pr}SSS_{ce}$ (the symbol "pr" stands for the organizational principles) factors stand for this aspect of the SSS organization. Finally, the SSSs themselves may be intercoordinated for the sake of the attainment of complex cognitive goals requiring the integrated activation of more than one SSS. The organization at this level is thought to occur mainly through the hypercognitive system (represented by the third-order HP factor). Curved two-way arrows connecting two or more factors indicate that the factors may be allowed to correlate instead of being regressed on a common higher order factor.

Space limitations do not allow us to discuss the methodological, statistical, and technical problems that one would have to solve to be able to evaluate the general representation of developing mind proposed here. We refer the interested reader to Demetriou et al. (1993; see also Connell & Tanaka, 1987). However, note that this representation can only be tested in a stepwise fashion. Specifically, the models that assume that the general factor is of the first order should be tested separately from models that assume that the general factor is of the second order (see the models shown in Figure 3). We need to point out, however, that the different models are complementary to each other rather than alternative representations of the same reality. That is, the different kinds of factors account for different portions of the variance the investigator is dealing with. The differently shaded areas of the squares in Figure 1 (in the conventional language of structural modeling, squares stand for observed variables and circles stand for latent variables or factors) are intended as representations for the different portions of variance accounted for by the different kinds of factors. Thus, a first-order general factor captures what is common among a number of observed variables, whereas a second-order general factor captures what is common among a number of latent variables or factors. The correlations between a number of factors provide information about their pairwise relations. These relations are disguised in models in which the factors are regressed to a higher order common factor.

Relations With Other Theories

This theory bears similarities to other developmental, cognitive, and psychometric theories. In fact, this theory aspires to function as a frame for unifying a number of other theories. Thus, we discuss briefly the relations between this theory and some other theories. The aim is to highlight how this theory tries to advance our understanding of cognitive organization and change by drawing on the strong points of earlier work.

The debt to Jean Piaget deserves first mention. In fact, this theory originated as an attempt to cope with the anomalies of Piagetian theory, most notably, inconsistent performance over different tasks presumed to represent the same operational structure (Demetriou & Efklides, 1988; Flavell & Wohlwill, 1969). Thus, the theory is still concerned with the development

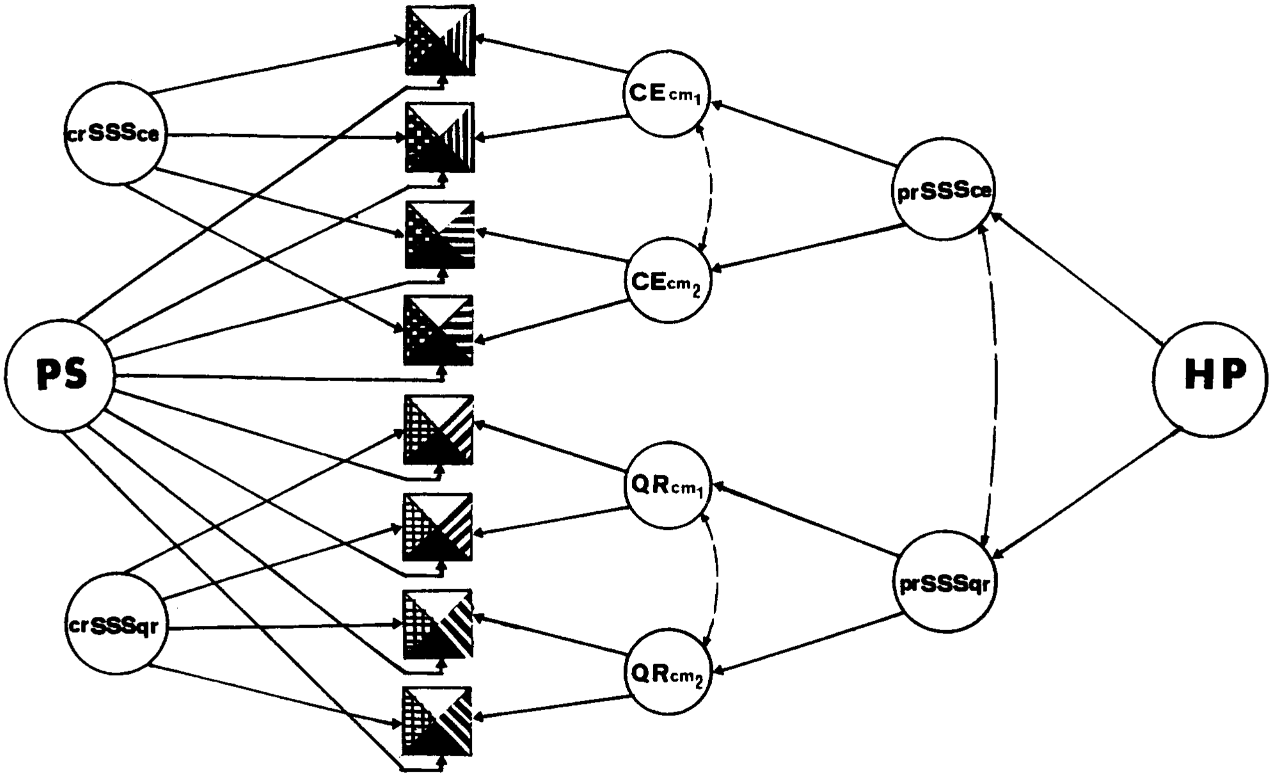

Figure 1. The general structure of mind as represented by the theory. (This representation involves the minimum number of components possible. At a full scale, it would involve all five Specialized Structural Systems (SSSs) and all component abilities involved in each SSS. PS = the processing system; HP = the hypercognitive system; cr = common core of an SSS; ce and CE = causal-experimental SSS; qr and QR = quantitative-relational SSS; cm = component; pr = organizational principles.)

of the same categories of thought (i.e., categorical, causal, quantitative, propositional, and spatial thought) that were studied by Piaget for over half a century. However, there is a great difference between this theory and Piagetian theory. Understanding the organization and change of different domains of cognitive activity is in the foreground of our research agenda. For Piaget, at best, it was in the background. At worse, it was noise hindering him in his attempt to understand the epistemic subject (see Piaget, 1973). Thus, a basic assumption of this theory is that pinpointing individual differences in cognitive functioning and change may reveal the operation of general cognitive mechanisms. For Piaget (1970), individual differences had nothing to tell about underlying cognitive structures. In as far as this assumption is concerned, this theory is closer to psychometric models of intelligence, in particular, the hierarchical models such as that of Vernon (1961). However, these models underestimate the developmental dimension of intellect.

The theory is closer to modern neo-Piagetian theories, especially those of Case (1985, 1992), Fischer (1980; Fischer & Farrar, 1988), and Pascual-Leone (1970, 1988). It shares with them two basic assumptions: first, that cognitive structures may be better understood by an analysis of their information-processing and semantic rather than their logical characteristics; second, that expanding processing capacity is a cause of changes in the complexity and the quality of cognitive structures the person can construct. This assumption, originally suggested by McLaughlin (1963), was impressively developed by Pascual-Leone (1970) and has been elaborated by Case (1985) and Halford (1988, in press).

There is an important difference, however, between these theories and our theory. These theories assume that a general characterization of the person's competence associated with successive general stages is possible. Thus, they propose very elaborate stage systems that are intended as substitutes for the Piagetian system. According to our theory, this assumption is untenable. Our theory claims that even if it were true that the SSSs do share common processing resources, these are, so to speak, protean in nature. That is, they are amorphous when inactive or in a latent state. Once activated, however, they take the form of the operating and symbolic characteristics of the task responsible for their activation. In effect, they are shaped on the pattern of the SSS that the activating task is affiliated to. Thus, the need to model the development of the different SSSs that control task performance cannot be dispensed with. Finally, none of the neo-Piagetian theories mentioned here involves premises about the development of the hypercognitive system and its effects on the other systems. Modern students of the child's theory of mind (Flavell, 1988; Wellman, 1990) do focus on the hypercognitive system. However, they do not pay attention to the relations between the theory of mind and the systems involved in the other two fronts described earlier.

Gardner's (1983) theory deserves special mention. Some of

Gardner's intelligences do correspond to some of our SSSs. Specifically, spatial and verbal intelligence appear very similar to our spatial–imaginal and verbal–propositional SSSs, respectively. Interestingly enough, some of Gardner's criteria for the definition of an intelligence are similar to some of our organizational principles (especially formal–procedural specificity and symbolic bias). However, there are some important differences between the two theories. On the one hand, our other three SSSs (qualitative–analytic, quantitative–relational, and causal–experimental) are all parts of what Gardner has undifferentiably called logicomathematical intelligence. On the other hand, Gardner refers to domains of activity (i.e., musical and kinesthetic intelligence) that are beyond the present concerns of our theory. Some other differences are more important than those mentioned here. Specifically, Gardner's is a static theory. It neither describes the development of his intelligences nor involves explanatory concepts capturing the dynamics of development. Finally, Gardner's theory seems to underestimate the regulatory role of hypercognition and the constraining role of the processing system.

Having defined our terms, we focus this article on the causal–experimental SSS, the object of the present study. In addition to understanding a particular SSS, this study has two general aims: first, to test if the general structural assumptions of the theory as represented in Figure 1 do hold for this SSS as has been found to be the case for other SSSs (Demetriou et al., 1993; Demetriou, Platsidou, Efklides, Metallidou, & Shayer, 1991); and second, to generate evidence that will shed light on the operation of synergic causality that may lead from within-SSS changes to changes in the general systems, such as the hypercognitive system.

Causal–Experimental SSS

Structure

Interactive reality structures consist of elements that act on each other in a multiplicity of ways. The condition of and the changes in some of the elements involved in the structure are always dependent on the condition of and the changes occurring in other elements of the structure. Interactive structures are often embedded in broader coexistence structures, in which the causal relations between elements are masked by the presence of co-occurring but unrelated elements. In this case, interactive structures cannot be directly perceived. Therefore, there must be a system that enables the person to disembed the causal structures from the more general structures of coexistence in which they are embedded.

This is the function of the *causal–experimental SSS.* That is, this system is directed to the processing of those causal reality structures that have a minimum degree of complexity, so that a minimally disciplined search is required if the relations between elements are to be located and specified in order to be understood and processed (Demetriou & Efklides, 1988). Therefore, the primary characteristic of the causal–experimental SSS is that it is analytic. Being analytic, the system is able to differentiate causal from other coexistence structures and to specify the type of causal relations that connect the factors under consideration. For this reason, the following component abilities are involved.

Combinatorial abilities form the cornerstone of this SSS. This is so because they function as the means used to exhaustively define the broader coexistence structure one would have to operate on in order to generate all possible interactions, of which the observed one is a special case. Therefore, combinatorial abilities are the basic instrument for the implemention of the system's analytic orientation to the processing of causal reality structures.

Hypothesis formation abilities enable the person to provisionally differentiate the combinations conceived into categories of possible causal relations. At a general level, the relations between a possible factor and a given result fall in one of the following five categories: Type *a,* necessary and sufficient; Type *b,* necessary and not sufficient; Type *c,* not necessary and sufficient; Type *d,* not necessary and not sufficient; and Type *e,* negative or incompatible. Relations of Types *a, b, c,* and *e* are interactive, whereas relations of Type *d* are plain coexistence relations. In turn, relations of Types *a, b,* and *c* are causal in the direct sense of the word. That is, they refer to the case in which a given effect *Y* is caused directly, entirely (Types *a* and *c*), or partially (Type *b*) by a given factor *X.* Relations of Type *e* refer to the (neutralizing) effect exerted on the causal factor *X* by some other factor *Z.*

In other words, hypothesis formation abilities enable the person to induce predictions about possible causal connections between the elements involved in a coexistence structure. These predictions are usually based on patterns of covariation between the elements already observed or conceived. The predictions must be taken by the person, at least implicitly, as propositions to be verified or falsified by experimentation. Therefore, hypothesis formation abilities enable the person to formulate a mental map of the experiments he or she would have to carry out and of the results he or she would have to expect.

Experimentation abilities enable the person to "materialize" hypotheses in the form of complementary experiments. The isolation-of-variables ability is probably the cornerstone of this set of abilities. This is equivalent to applying the scheme "all other things should be kept equal except for the factor tested" to see if varying this factor does produce the hypothesized effect.

To be able to fulfill its function, this ability should be supported by a set of complementary abilities. These enable the person to resist the *maximization fallacy* and the *dependent variable fallacy.* The first leads to the belief that if another factor, which is already known to produce the same effect, is varied together with the factor under study, then the target factor's effect will be maximized. Thus, the fallacy goes, the hypothesis related to this factor will be tested more easily. By yielding to it, one is canceling the isolation-of-variables scheme mentioned earlier. The second fallacy leads to the belief that if the dependent instead of the independent variable is manipulated, then the hypothesis would be tested more firmly because reciprocal effects would have to be produced. By yielding to it, one makes one's experiment irrelevant to the hypothesis tested. This is due to the possibility that an unknown factor may be in operation that causes the same type of variation in the dependent variable (Demetriou & Efklides, in press; Demetriou et al., 1993).

Model construction abilities enable the person to properly map the results of experimentation with the original hypothesis in order to arrive at an acceptable interpretative framework or theory. Therefore, a model should involve statements about causes, effects, and extraneous variables, differentiating the sought-after causal structure from the broader coexistence structure in which the causal structure is embedded. Ideally, a model should refer to higher order constructs that are able to integrate phenomenally contradictory evidence. In this way, the model would contain the elements of its own modification. That is, it would indicate what hypotheses have to be tested by a next series of experiments (Demetriou, 1990).

It is interesting that other researchers have recently analyzed causal–experimental or scientific thought, as they call it, in terms very similar to those advanced here. Klahr and Dunbar's (1988) analysis, in particular, is very close to our analysis. They speak about a *search hypothesis space, test hypothesis,* and *evaluate evidence* components. Clearly, these components correspond to our hypothesis formation, experimentation, and model construction components. Moshman's (1979) analysis and the recent studies by Kuhn, Amsel, and O'Loughlin (1988) are concerned with the same components.

Development

The subjects involved in the testing of the foregoing formulation were adolescents. Therefore, the present discussion focuses on the development of abilities involved in the causal–experimental SSS from early adolescence to maturity.

Combinatorial abilities. The first, and one of the most comprehensive studies of the development of combinatorial thinking, was presented by Inhelder and Piaget (1958). According to this study, which has been subsequently replicated by many other investigators (see Neimark, 1975), the development of combinatorial abilities takes place at three general levels. At the first level, the child is able to effect no more than clearly apparent combinations. His or her inability to go further is ascribed to the fact that he or she operates on the level of observable reality rather than on the level of the possible. Thus, the child can conceive 1 × 1 or 2 × 2 combinations because he or she can literally see the elements going together. However, higher order combinations presuppose that the person is able, from the beginning, to conceive the networks of relations between elements at the level of the possible. Once one is able to form this conception, the combinations can be generated as instantiations of the network already mentally constructed. This capability is achieved at the second level. However, at this level, the person is still lacking the skills that would enable him or her to fully control his or her newfound potentialities to operate in the world of the possible. As a result, he or she produces combinations unsystematically and inexhaustively. These limitations are removed at the next level. At this level, the person uses one or more strategies (e.g., specify all 1 × 1 combinations first, then all 2 × 2 combinations) that enable him or her to systematically and exhaustively move throughout the "combinatorial space."

Hypothesis formation. The development of this ability can also be described in terms of three levels. At the first level, the person is able to hypothesize that there is a causal relation between two factors if they systematically covary. However, at this level, the person falls prey to *false inclusion,* in Kuhn et al.'s (1988) terms. That is, the person tends to infer a causal relation between an outcome and a factor only on the basis of his or her observation that this factor was present when the outcome occurred. Thus, at this level, the person tends to form what we call *symmetric hypotheses.* In other words, the person tends to ascribe whatever changes he or she observes in the dependent variable to corresponding changes in one particular dependent variable, without envisaging the possibility that another variable may be operating.

In terms of the analysis of hypothesis formation abilities advanced here, symmetric hypothesizing enables the person to formulate hypotheses about necessary and sufficient or, at most, necessary and not sufficient relations but not about the other types of relations. In fact, at this level, Type *d* and Type *e* relations are taken to be either Type *a* or Type *b* relations. At the second level, the person begins to master *exclusion.* That is, he or she begins to understand that copresence or covariation does not necessarily imply causality. Thus, he or she can formulate hypotheses about not necessary and not sufficient relations. The implication is that, at this level, hypotheses are formulated on the basis of sophisticated scanning of the information patterns available. To hypothesize that a variable is not related to a given outcome with which it co-occurs, one needs to be able to contrast the co-occurrence observed at one point in time with at least one nonoccurrence observed at another point in time. Finally, at the third level, hypotheses about complex interactions can be formulated. This implies that hypotheses are now formulated on the basis of a careful examination of each variable in relation to all other variables present. As a result, the incompatibility relation of Type *d* can be understood (Demetriou et al., 1993, Study 1).

Experimentation. At the first level, the ability to design experiments is limited by the maximization fallacy mentioned earlier. The person does understand that some of the factors mentioned in the hypothesis have to be varied if the hypothesis is to be tested. However, he or she thinks that the effect of the factor under consideration will be better demonstrated if some other factors already known to produce the effect referred to in the hypothesis are allowed to "contribute." At the second level, the person realizes that only one factor has to be varied. However, at this level, he or she may be deceived by the dependent variable fallacy or by the assumption that class matching is sufficient. For instance, a person may grow beans in a sunny place and lentils in a shadowy place to test the effect of light on the growth of plants because "they are both pulses so they are not really different" (Demetriou & Efklides, in press). At the third level, the design of experiments is always directed by the assumption that there is only one way to make an experiment relevant to the hypothesis. This is to vary only the factor that is mentioned as the cause in the hypothesis. Thus, at this level, it is understood that if *n* factors are referred to as "causes" in the hypothesis, then they should be allowed to interact in such a way so that an *n*-way experiment is generated. This is conceived as a composite of "subexperiments" such that each of them tests the effect of one and only one of the factors specified in the hypothesis.

Model construction. According to Kuhn et al.'s (1988) work, the development of model or theory construction is a process that leads the person from the condition of thinking with a model to the condition of thinking about a model. In Moshman's (1979) terms, this is equivalent to moving from the use of theory to the use of metatheory as a frame guiding the generation and evaluation of evidence.

At a first level, which precedes adolescence, a person does not differentiate models from the evidence they are supposed to explain. The models he or she conceives are nothing more than a juxtaposition of statements that describe the results generated by the manipulations attempted. At the second level, the person begins to differentiate evidence from the interpretations that might be advanced in regard to it. As a result, he or she systematically attempts to map the data generated by his or her manipulations with what would have to be expected on the basis of the hypothesis. However, at this level, the person is still lacking the skill that would make this mapping exhaustive. This causes misinterpretations of the data vis-à-vis the hypotheses or misrepresentations of the data in the model. These limitations are overcome at the third level. At this level, it is understood that even the models presently held to be true are, at best, subcases of better models to be constructed in the future. Thus, on the one hand, the person is able to conceive very refined models to account for the results of complex experiments designed to test hypotheses interwoven in very sophisticated ways. On the other hand, he or she is perfectly aware of the temporariness and the (presently unknown) limitations of his or her endeavors (Demetriou, 1990; Demetriou & Efklides, 1985, in press; Demetriou, Gustafsson, Efklides, & Platsidou, 1992; Efklides, Demetriou, & Gustafsson, 1992).

An Empirical Investigation of the Structure and Development of Causal–Experimental Thought

Predictions

Structure. We argued earlier that, on the one hand, each of the four sets of abilities involved in the causal–experimental SSS is characterized by its own representational and processing peculiarities. On the other hand, they all belong to the same SSS. Therefore, according to the general representation of mind depicted in Figure 1, it is to be expected that performance on tasks representing component abilities should be structured, in terms of confirmatory factor analysis, in two types of factors: (a) four ability-specific factors representing each of the component abilities and (b) a general factor. This would have to be either of the first or the second order. The first would represent the common core shared by all abilities. The second would capture the intercoordination of the properties specific to each set into a unified SSS. This general model is analyzed further in the Results and Discussion section. However, we need to state here that the validity of the model would become questionable to the extent to which any of these factors would be found statistically rejectable (Prediction 1).

Development. The assumption about the existence of general factors suggests a further hypothesis regarding the interpatterning of the tasks belonging to the four batteries. They should be reducible to a common developmental scale by means of analyses that test for unidimensionality, such as Rasch analysis. In simple developmental terms, a set of items might be unidimensionally scalable when the pattern of successes and failures on them is such that the items addressed to a given level L are passed by the persons operating on this or all higher levels $L + n$ and failed by the persons operating on the lower levels $L - n$. Inconsistent success–failure patterns cannot be calibrated along a common scale (Fischer, Hunt, & Russell, 1984; Shayer, Demetriou, & Pervez, 1988; Wilson, 1989). Meeting this requirement is the minimum precondition for the design of a developmental model that would aim to account for causal interactions between abilities (Prediction 2). It need not be emphasized here that, according to this theory, this prediction holds for within-SSS but not for across-SSS relations.

Having confirmed this hypothesis, one might then proceed to test two further predictions regarding the placement of the various abilities along this dimension. Specifically, the developmental analysis of the four abilities suggests that they should exhibit overlapping developmental trajectories. The reader is reminded that, on the one hand, the combinatorial ability was considered to be the building ground on which the other abilities are constructed. On the other hand, the model construction abilities seem to be the developmental endpoint of the causal–experimental SSS. Thus, one is justified in assuming that the combinatorial abilities should appear and mature first. The model construction abilities should appear relatively late and mature last. Hypothesis formation and experimentation abilities should make their appearance somewhere in the middle of the development of combinatorial abilities and mature before the model construction abilities arrive at their developmental endpoint (Prediction 3).

Individual differences. One can look for individual differences in regard both to structure and development of any set of abilities. Regarding structure, one may anticipate that the same set of abilities may be structured in the same or in a different way in different social or cultural groups. In as far as the structure of the present abilities is concerned, it is expected that the same factors would be abstracted from the performance of the various age and social groups (Prediction 4). This prediction, which has been confirmed in our previous studies concerning other SSSs, is based on the postulate of the theory that the structural organization of abilities is a strong universal (Demetriou & Efklides, 1985). That is, the organization reflects the structural tuning of the human mind to the structural organization of external reality. Thus, the SSSs are conceived as something like Thurstone's primary abilities or Kant's natural categories of reason. Such a strong argument is supported by recent research in infant cognition. A number of studies have suggested that the abilities underlying understanding of quantitative (Starkey, Spelke, & Gelman, 1990), causal (Spelke, 1984), and spatial relations (Landau, Spelke, & Gleitman, 1984) are present from the very first months of life.

Regarding development, however, the theory is compatible with the assumption that factors traditionally associated with individual differences in cognitive functioning, such as sex or socioeconomic status (SES), may affect the rate of ascension from the lower to the higher levels of a developmental hierarchy. Based on our earlier research (see Demetriou & Efklides, 1985, 1989), we would anticipate that high SES subjects would outperform low SES subjects. Based on the same research, differences between genders must not be expected. We have consistently found that the causal–experimental SSS does not differentiate the genders, although there may be gender differences in regard to other SSS, such as the quantitative–relational (Prediction 5; Demetriou & Efklides, 1985; Demetriou et al., 1991).

Method

Subjects

Table 1 shows the distribution of subjects across grades, SES, and sex. A total of 260 subjects (123 boys and 137 girls) were involved in the study. The age of these subjects ranged from 12 to 17 years. We selected this age range because it is particularly appropriate for providing data that is able to test the predictions of this study. Our own earlier research (Demetriou & Efklides, 1985) and others' (Martorano, 1977; Neimark, 1975; Shayer, 1979) have shown that this age phase is coextensive with the period during which practically all of the abilities to be investigated here appear and are consolidated.

The subjects in each age group were about equally drawn from upper-middle SES (i.e., professionals and businessmen) and from working-class families of urban residence. The main criterion for allocating subjects to SES groups was the education of the parents. Specifically, at least one of the parents of the high SES subjects had to be a university graduate. Both parents of the low SES subjects had no more than primary school education. The subjects in each SES group were about equally divided between boys and girls. All subjects were Greeks living in Thessaloniki, Macedonia's capital.

Tasks

Three task batteries were given to each subject. One was addressed to the combinatorial ability and one to the experimentation ability. The third was a complex battery that involved items addressed to (a) the hypothesis-formation ability, (b) the model-construction ability, and (c) metacognitive awareness regarding the epistemological status of one's own and the others' models. The three batteries are described next in some detail.

Combinatorial ability battery. This battery involved two sets of items. In each of the four items involved in the first set, a diagram was presented to the subject that presumably shows an area in which the subject would have to move. The four diagrams are shown in Figure 2. The task of the subject was to specify all alternative paths that he or she could follow in order to go from the departure point A to the arrival point T. Thus, this part of the battery aimed to test the ability of the subject to execute combinations in a rather familiar context. We manipulated the difficulty of the items by varying the number of paths or their complexity. Specifically, four, four, six, and nine paths had to be specified in Items 1, 2, 3, and 4, respectively. In the case of the two first items, we manipulated the difficulty by varying the perceptual organization of the paths. It can be seen in Figure 2 that two of the paths in Item 2 (i.e., the paths ACDT and ABDCET) are less easily perceived as paths leading from A to T than the corresponding paths in Item 1, as a result of the orientation of the figure.

The second part involved six items. In each item, the subject was asked to specify all possible sequences in which he or she could take out of a box the balls put in it by the experimenter. We manipulated the difficulty of these items by varying the number or the color of the balls or both. For space considerations, let the letters B, G, R, and P stand for the blue, the green, the red, and the pink balls, respectively. Item 1 involved one G and two Rs; Item 2 involved one R, one G, and one B; Item 3 involved two Rs and two Bs; Item 4 involved one G, one B, and two Rs; Item 5 involved one R, one G, one B, and one P; and Item 6 involved three Rs and two Gs. The structure of Items 1 through 6 provides for 3, 6, 6, 12, 24, and 10 possible combinations, respectively. Both Items 2 and 3 involved a maximum of 6 combinations. However, each of the combinations in Item 2 involved three balls, whereas each combination involved four balls in Item 3. Likewise, each of the combinations in Item 5 involved four balls, compared with five balls in Item 6, with less resulting combinations. In this way, we might be able to dissociate the effect of the sheer number of combinations from the effect of the structure of combinations on combinatorial ability.

Experimentation ability battery. This battery involved four tasks. These were concerned with the effect of a number of factors (frequency of irrigation, area, and presence or absence of fertilizer on the productivity of two plants A and B). Each of the tasks presented a hypothesis to be tested by a properly designed experiment, as well as the conditions and materials the subjects had to work with in order to design their experiment. We manipulated the difficulty of the tasks by varying the complexity of the hypothesis to be tested.

In the simplest task, the hypothesis referred to the effect of only one factor on plants in general ("the increase in watering frequency increases the productivity of plants") and asked the subject to use any of Plants A and B and two watering frequencies (two or four times a month) to design an experiment aimed to test the hypothesis. Thus, in this task, only one independent variable would have to be manipulated to test its effect on a single dependent variable. This is a one-way experiment. That is, irrigating either Plant A or B two and four times per month would be sufficient to test the hypothesis.

In the second task, the hypothesis stated that "(a) irrigation increases the productivity of Plant A but (b) it does not affect the productivity of Plant B." Thus, the same independent variable as before would have to be manipulated to independently test its effects on two dependent variables. That is, each of the Plants A and B would have to be irrigated two and four times per month to test the hypothesis. This is a two-way (2 × 2) experiment.

In the third task, the hypothesis stated that "(a) irrigation does not affect the productivity of Plant A in Area I, but it does affect it in Area II; however, (b) irrigation does not affect the productivity of Plant B in Area II, but it does affect it in Area I." To test this hypothesis, both Plants A and B would have to be irrigated both two and four times per month in both Areas I and II. This is a three-way (2 × 2 × 2) experiment.

Table 1
Composition of the Sample

Grade	Age (months)	High SES Boys	High SES Girls	Low SES Boys	Low SES Girls	*N*
7	152	13	15	12	18	58
8	165	15	17	16	16	64
9	176	14	18	16	17	65
10	190	7	9	6	8	30
11	199	17	14	7	5	43
Total		66	73	57	64	260

Note. SES = socioeconomic status.

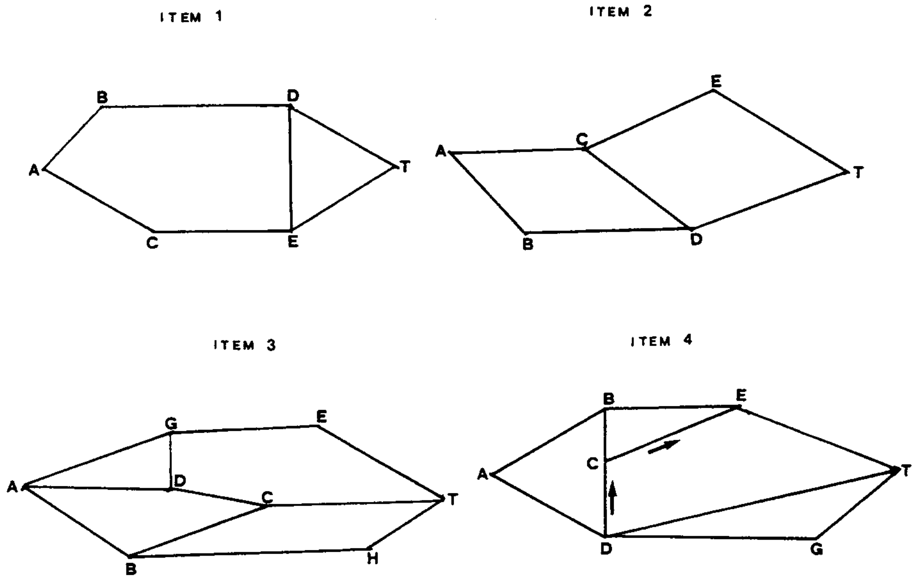

Figure 2. Some of the combinatorial thought items. (The subjects were instructed to write down all possible paths leading from the departure point A to the arrival point T. The arrows indicate one-way paths.)

In the fourth task, the hypothesis involved the use of fertilizers in addition to the factors involved in the third hypothesis. To test this hypothesis, one would have to design an experiment in which each of Plants A and B would have to be fertilized both once or twice per month in each of Areas I and II under each of the two irrigation frequencies. This is a four-way (2 × 2 × 2 × 2) experiment.

Hypothesis-evidence handling battery. This was a complex battery that involved eight items addressed to the ability to (a) induce hypotheses on the basis of evidence, (b) deduce the crucial evidence that would be consistent with a hypothesis, (c) develop models that would integrate a series of hypotheses into a comprehensive system, and (d) conceive of the epistemological status of scientific theories. All items revolved around the following summarized story.

According to this story, there were some farmers whose cultivation of wheat in a mountainous region failed. The farmers hypothesized that the cause of failure was because their region was always cloudy, and they invited an agriculturist to help them solve the problem. The agriculturist supposedly had his own hypothesis, but he did not mention whether he agreed with the farmers or not. To test his hypothesis, he carried out several experiments that were shown in a table. This table involved three columns. In the first column, there were four entries, each corresponding to an experimental manipulation. Specifically, he cultivated the wheat in (a) an area at low altitude that was always cloudy, (b) an area at low altitude that was never cloudy, (c) an area at high altitude that was always cloudy, and (d) an area at high altitude that was never cloudy. The second column presented the results expected from each experiment: that is, that the crop would succeed, succeed, fail, and fail, respectively. The third column presented the results found from each experiment: that is, that the crop succeeded, succeeded, failed, and failed, respectively.

In the first part of the test, the subjects were asked to study this table and answer the following questions.

1. Guess the hypothesis of the agriculturist by taking into account the results that he expected to obtain. Specifically, the subject was asked to decide if the hypothesis of the agriculturist was that (a) cloudy conditions alone, (b) cloudy conditions in combination with low altitude, (c) high altitude alone, or (d) cloudy conditions in combination with high altitute was the cause of the failure.
2. Guess the conclusion of the agriculturist by taking into account the results that he obtained. The four aforementioned hypotheses were now stated as conclusions, and the subjects were asked to choose the correct one and explain their selection.
3. Guess whether the agriculturist would conclude that the results were (a) in agreement or (b) contrary to his hypothesis, and explain why.
4. The subject was then given the second part of the battery. In this part, both the hypothesis of the agriculturist was revealed (i.e., "that cloudy conditions do not have any effect, high altitute alone is the cause") and the conclusion was suggested that the results were in agreement with the hypothesis. The story then went as follows: "However, a second agriculturist claimed that his colleague who had run the experiments was wrong. Specifically, he argued that the cause of the failure was not the high altitude alone but the combined effect of high altitude and a kind of fungi which usually grow at high altitude. Thus, the second agriculturist decided to run a second series of experiments aimed to test this hypothesis." These experiments, which were given in a second table, were the eight possible combinations between the two levels of each of the three variables involved (i.e., low and high altitude, cloudy and noncloudy conditions, and presence or absence of fungi). The subject's task was to specify the result that would be expected from each experiment (success or failure), taking into account the hypothesis tested.

Having answered this question, the subject was given a table presenting what would be expected from each experiment according to the hypothesis and what was actually found. Specifically, it would be expected that the cultivation would fail only in (the three) experiments in which both high altitude and fungi were present. Supposedly, the results obtained were consistent with this hypothesis with only one exception. Specifically, the cultivation also failed in the "low altitude, with clouds, and with fungi" experiment. On the basis of this information, the following questions were asked.

5. Specify if the hypothesis was (a) entirely wrong, (b) entirely correct, or (c) involved right and wrong elements, or (d) specify if there was an additional factor that affected the cultivation, and explain why.
6. The subjects were then reminded of all three hypotheses (i.e., that of the farmers, that of the first agriculturalist, and that of the second agriculturalist) and of the results obtained (contrary to the former, in

total agreement with the second, in partial agreement with the latter). The subjects were then asked to state their own hypothesis that would be able to explain all three hypotheses and accommodate the results of all experiments that have been conducted. The last two questions were concerned with epistemological issues. They were as follows.

7. Given that your hypothesis is better than the others, would you say that all experiments that you might run in order to test it would produce results in agreement with it?

8. Who would you regard as more confident about his views regarding the failure of the cultivation? The farmers or the agriculturist? In other words, when do you think that one believes that one's knowledge is right: When one has not subjected it to any control, or when one has started testing it? Explanations were required in all cases.

Scoring Criteria

Combinatorial ability battery. Performance on this battery was scored on the basis of a 3-point scale. Specifically, a score of 0, 1, or 2 was ascribed to the performance attained on an item when (a) less than half, (b) more than half but not all, and (c) all of the combinations of the given item, respectively, were found.

Experimentation ability battery. Performance on this battery was also scored on the basis of a 3-point scale. A score of 0 was given to all responses indicating that the subject was not able to apply the scheme, "all other things should be equal except for the factor tested." A score of 1 was ascribed to experiments indicating that the subject was able to apply this scheme but the experiments proposed were not able to test the hypothesis of the experiment (see the aforementioned fallacies). A common response that falls in this category is to have an experiment in which all factors but one are held constant but the factor manipulated is not the one that has to be tested, according to the hypothesis. Finally, a score of 2 was given to experiments in which all factors were held constant except for the one that had to be tested, according to the hypothesis.

Hypothesis-evidence handling battery. Performance on this battery was scored on the basis of a 4-point scale. Specifically, a score of 0 was always ascribed to no, nonsensical, or invalid responses.

A score of 1 was ascribed to correct but totally unexplained or partially correct but explained responses in the case of Items 1, 2, 3, 4, and 5. In the case of Item 6, a score of 1 was ascribed to responses simply reiterating or juxtaposing the three hypotheses given. In the case of Item 7, this score was ascribed to responses indicating a belief that the experiments would confirm the hypothesis. In the case of Item 8, this score was ascribed to responses indicating that confidence in the correctness of one's assumptions is a function of this person's authority (e.g., "The agriculturist would be more confident because he is more knowledgeable than the farmers").

A score of 2 was ascribed to correct but partially explained responses in the case of Items 1, 2, 3, 4, and 5. In the case of Item 6, this score was ascribed to models indicating that the subject attempted to integrate at least two of the three factors referred to in the three hypotheses into a common framework. In the case of Item 7, this score was ascribed to responses indicating a recognition of the fact that theories may always be refuted, although they may be superior in relation to other theories. In the case of Item 8, this score was ascribed to responses indicating that confidence to the correctness of one's assumptions is a function of the measures that one takes to test these assumptions (e.g., "The agriculturist would be more confident because his assumptions came out of careful control and they are not a matter of belief or impressional observations").

A score of 3 was ascribed to correct and fully explained responses in the case of Items 1, 2, 3, 4, and 5. In the case of Item 6, this score was ascribed to models introducing an extra factor able to explain the effects of the three factors mentioned in the hypotheses at one and the same time (e.g., "Humidity might be the real cause because it is usually humid at high altitutes and when it is cloudy, and fungi normally thrive in humid environments"). In the case of Item 7, this score was ascribed to responses clearly specifying that unforeseen or presently inconceivable factors may be in operation that might partially or fully invalidate the theory. In the case of Item 8, this score was ascribed to responses coordinated with those given to Item 7. That is, that one who tests his or her assumptions would, on the one hand, be more confident in their validity for the reasons referred to here. On the other hand, he or she may feel uneasy because he or she is aware of the possibility that every assumption may eventually be refuted.

Origin and Reliability of the Three Batteries

The primary aim of this study was to test the assumptions of our theory regarding the structure and development of the causal–experimental SSS. However, the study also aimed to provide a more refined and integrated picture of this SSS than could be formulated on the basis of earlier research. These two aims made it necessary that the characteristics of the tasks to be used had to be known in some respects and new in others.

Specifically, the *combinatorial ability* battery was devised for the purposes of the present study. However, it was based on similar tasks first used by Piaget (Inhelder & Piaget, 1958; Piaget & Inhelder, 1975) and others (Demetriou & Efklides, 1985, in press) and that are known to be good indicators of combinatorial ability. Our earlier studies have already shown that this ability loads on the same factor with tasks tapping other components of the causal–experimental SSS. The alpha reliability of this battery was .88.

The *experimentation ability* battery was first used in our studies that tested the intra- and inter-SSS transfer of training and the relations between our SSSs and psychometrically defined general intelligence (Demetriou et al., 1993, Study 2; Efklides et al., 1992). These studies have clearly demonstrated that the abilities represented by this battery, although related, are clearly different from psychometrically defined fluid intelligence or the quantitative–relational SSS. The alpha reliability of this battery was .69.

The *hypothesis-evidence handling* battery was first used in the present study. The design of the battery was guided by the need to have items representing as clearly as possible the ability to induce hypotheses from a database, as well as the need to grasp the real and epistemological relations among hypotheses, experimental manipulations, and models. Previous research had shown that these abilities are clearly distinct from those tapped by the other two batteries (Demetriou & Efklides, 1985, 1989; Kuhn et al., 1988). However, their structural and developmental relations with these other abilities had not been systematically investigated within the same experiment. The alpha reliability of this battery was .82. Because the scoring of performance on this battery was based on both judgments and explanations, the interrater agreement for each of the eight items was calculated. It was high and ranged from 90% to 98% (M = 95%, SD = 2.82%).

Procedure

The subjects were tested in groups during school hours. The presentation order of batteries was randomized across subjects. However, the items within a battery were presented in the order described earlier. That is, difficult items were always presented after easy items. This decision was based on two reasons. First, presenting the difficult items first might negatively predispose the subjects even to those items that were within the range of their ability. Second, the transfer of practice from the easy to the difficult items may be regarded as a kind of support that might enable the subjects to approach their optimum level as fully as possible (see Fischer et al., 1984). No time limitations were imposed on the subjects. However, in the order they were described earlier, an average of about 15, 30, and 45 min were required for the completion of the three batteries, respectively. Specifically, the first two batteries were administered during the same 45-min school period. The longer battery was administered during a second school hour. A 15-min break was allowed between the two testing sessions. One of the authors was always present throughout testing to provide clarifications as needed.

Results and Discussion

Structure

To test the predictions concerned with the structure of abilities, we analyzed the data through a sequence of confirmatory factor analysis models, fitted with the latest version of the EQS program (Bentler, 1989). These models were fitted on mean rather than on raw scores. Specifically, three mean scores were created for each subject to represent performance on the combinatorial ability battery. Mean Score 1 involved Item 1 of the paths set and Items 1 and 5 of the balls set. Mean Score 2 involved Item 4 of the paths set and Items 2 and 4 of the balls set. Finally, Mean Score 3 involved Items 2 and 3 of the paths set and Items 3 and 6 of the balls set. Performance on the experimentation battery was reduced to two mean scores, one for Tasks 1 and 3 and one for Tasks 2 and 4. Performance on the eight items of the hypothesis-evidence handling battery was reduced to five scores, specifically, a mean score for Items 2 and 4 (going from evidence to a conclusion or from a hypothesis to evidence), Items 3 and 5 (evaluating a hypothesis on the basis of the evidence obtained), and Items 7 and 8 (epistemological awareness). Items 1 (hypothesis formation) and 6 (model construction) were involved in the analysis as raw scores because no other similar items were involved in the battery. Thus, whenever possible, we attempted to have different difficulty levels and content variations represented in each of the mean scores of a given battery or a category of items (Figure 4 shows the means in which each score was collapsed).

This approach was adopted for technical as well as substantive reasons (Gustafsson, 1988). According to theorists of measurement, averaging over raw scores reduces bias variance in the measurements representing observed variables (Humphreys, 1976). This facilitates the identification of latent variables or factors. One needs to be particularly sensitive to factor identification problems when confirmatory factor analysis and structural equation modeling is used. These methods pose very strong demands on the reliability of the measurements fed into the analysis.

It may be noted here that the attempt has been made to have more than two score indicators for a factor, whenever this was possible. Satisfying this requirement removes possible factor identification problems that may be encountered in the case of factors associated with only two score indicators. We need to emphasize, however, that factor identification was never a problem in this study. This is so because the fixation of factor variance to 1 (Model 2A), the estimation of the paths from the first-order factors to the second-order factor (Model 2B), and the estimation of covariance between the factors (Model 3C; see below) act as extra indicators of each factor (Bentler, 1989).

The zero-order correlations between the scores representing the same ability were high: mean r = .648, SD = .112. The correlations between scores representing different batteries were considerably lower than the within battery correlations: mean r = .314, SD = .078.

According to our prediction (Prediction 1) regarding structure, the 10 scores specified earlier should be structured in two types of factors: (a) four ability-specific factors representing each of the component abilities and (b) a general first- or second-order factor. Specifically, validating a model that presumes a first-order general factor would indicate, on the one hand, that the observed variables themselves do share one or more common components. On the other hand, this model enables one to test whether each of the ability-specific factors does indeed stand over and above the general factor as an organizer of some aspects of mental activity. The nested-factor method, pioneered by Gustafsson (1988), is the method par excellence for resolving these questions. Mathematically, this method is based on the assumption that the variance of a variable is a linear function of the set of factors that, according to theory, determine the dimensions of performance represented by this variable. In the present case, each of the 10 scores was prescribed to load on the general factor. In addition, the three mean combinatorial scores, the hypothesis formation raw score and the two hypothesis-evidence mapping mean scores, the two design of experiments mean scores, and the model construction raw score together with the epistemological awareness mean score were prescribed to load on the combinatorial, the hypothesis handling, the experimentation, and the model construction factor, respectively (see Figure 3). The maximum likelihood method was used for the estimation of all models.

To test whether the contribution of each single factor is significant, one proceeds in a stepwise fashion: Progressing from the most general to the most narrow factor, only one factor is added to the model each time. Thus, it is possible to test whether the factor added at each step accounts for a significant portion of the variance (i.e., the residual variance) not accounted for by the factors already introduced in the model. If so, the additional factor introduced into the model should result in a significant reduction of the chi-square statistic, given the difference of the corresponding degrees of freedom. Table 2 presents the statistics of the models tested in the five successive runs, which started with only the general factor and ended with all five factors specified in the previous paragraph. It can be seen that the introduction of each of the factors resulted in a highly significant improvement of model fit. It can also be seen in Table 2 that the fit of the complete model, which involved all five factors, was excellent, $\chi^2(24) = 30.848$, $p = .158$, comparative fit index (CFI) = .994; conventional fit indexes exceeding .9 are regarded as signifying that the model is a good representation of the data (Tanaka, 1987). This is the model shown in Figure 3A. We need to stress that all of the loadings—which can be interpreted as correlations—going from an ability specific factor to an observed variable were significant. Likewise, all loadings of the variables on the general factor were significant. Therefore, all four subsystems of abilities involved in the causal–experimental SSS appear to stand as autonomous systems that share common components.

One might object that the first-order factor identified here is confounded with the factor that, according to the general representation of mind suggested in Figure 1, would represent the processing system. The objection is justified, because to separate the two factors, one would need to involve in the analysis

Table 2
Results From Tests of Fit of Nested Factor Models for the 10 Causal–Experimental Scores

Model	Factor included	Fit of models			Change		
		χ^2	*df*	CFI	χ^2	*df*	*p*
1	CE SSS	419.522	34	.680			
2	+CO	206.340	31	.854	213.182	3	.001
3	+HH	141.779	28	.905	64.561	3	.001
4	+DEx	77.960	26	.957	63.819	2	.001
5	+MOD	30.848	24	.994	47.112	2	.001

Note. The factors tested are the factors shown in Figure 3A. CFI = comparative fit index; CE SSS = causal–experimental Specialized Structural System; CO = combinational ability; HH = hypothesis handling ability; DEx = experimentation ability; MOD = model construction ability.

measures directly representing the processing system. Under this condition, the PS factor would be geared on the PS-specific measures. No such measures were used here. When they were used in a previous study, however, they resulted in a PS factor accounting for between 28% and 54% of the variance of the SSS-specific measures, compared with 32%–64% of the variance of these measures accounted for by their SSS-specific factors (Demetriou et al., 1993, Study 5). Applying this relation to the present results leads to the conclusion that at least half of the variance accounted for by the first-order general factor represents the common core of the causal–experimental SSS.

The model shown in Figure 3B tested the assumption that the general factor is of the second order. The fit of this model to the data was very good, $\chi^2(30) = 39.868$, $p = .108$, CFI = .992. In fact, this model is statistically indiscriminable from the model discussed in the previous paragraph. Attention is drawn to the fact that the loadings of the ability-specific factors on the general factor were generally about the same and very high. This indicates that the causal–experimental SSS does function as a powerful attractor that forces its component systems to remain interrelated.

The models shown in Figures 3A and 3B suggest that the four subsystems of abilities investigated here do share a common core (Figure 3A) and are interrelated through their connection to the same SSS (Figure 3B). However, none of these models shows how each of the subsystems is directly related to the others. The answer to this question is suggested by the model shown in Figure 3C, $\chi^2(29) = 32.775$, $p = .287$, CFI = .997. In this model, all four subsystem-specific factors were taken to be independent factors that were allowed to correlate. All correlations were significant. Nevertheless, not all of them were equally high. Specifically, the correlation between the combinatorial factor and the hypothesis handling or the experimentation factor was considerably higher than the correlation between the combinatorial factor and the model construction factor. The correlation between the model construction and the hypothesis handling factor was considerably higher than the correlation between the model construction and any of the other factors. It will be seen below that the factors correlating higher are those representing abilities that follow overlapping developmental trajectories. Therefore, this pattern of correlations may be taken as a sign of the strength of reciprocal interactions that may be occurring between abilities during development.

Finally, one may justifiably ask here why the loadings of the observed variables on the ability-specific factors in the model shown in Figure 3A are lower than their corresponding loadings in the other two models. This is because the variance of the observed variables was partitioned between two factors in Model 3A but was associated with only one factor in the other two models. Thus, a part of the higher loadings in Models 3B and 3C represents common rather than ability-specific components.

This possibility may be a problem when interpreting the meaning of the second-order factor in Model 3B or the between-factors correlations in Model 3C. That is, one might claim that they too represent the common core in disguise rather than the intercoordination of the component abilities as such. To examine this possibility, we subjected Models 3B and 3C to a second test. In this test, each of the 10 paths going from the ability-specific factors to observed measures was fixed to be equal to its corresponding value in Model 3A. In principle, this manipulation purifies the paths from the possible contribution of the first-order general factor. The values of the paths from the second- to the first-order factors (Model 3B) or the between-factors correlations (Model 3C) should be appreciably reduced if the first-order general factor were responsible for their magnitude. The values produced by this manipulation are shown in parentheses in Figure 3. All of them were still significant and only slightly lower than before. Therefore, as assumed in the introduction and specified in Prediction 1, each of the three models does represent a different aspect of the dynamics underlying cognitive organization.

Development

Rating scale analysis was applied on the set of 22 items involved in the four batteries to build a unified hierarchy of causal experimental abilities and specify the exact position of each item in this hierarchy (Rasch, 1980; Wright & Masters, 1982). For the purposes of the present analyses, the scores of 2, 2, and 2 were defined as "success scores" on the items involved in the combinatorial, the experimentation, and the hypothesis handling battery, respectively.

It may be noted that these 22 items seem to meet two of the main assumptions of the Rasch (1980) model. According to the first of these assumptions, the "pass" response must not be attained by guessing. According to the second, all items would

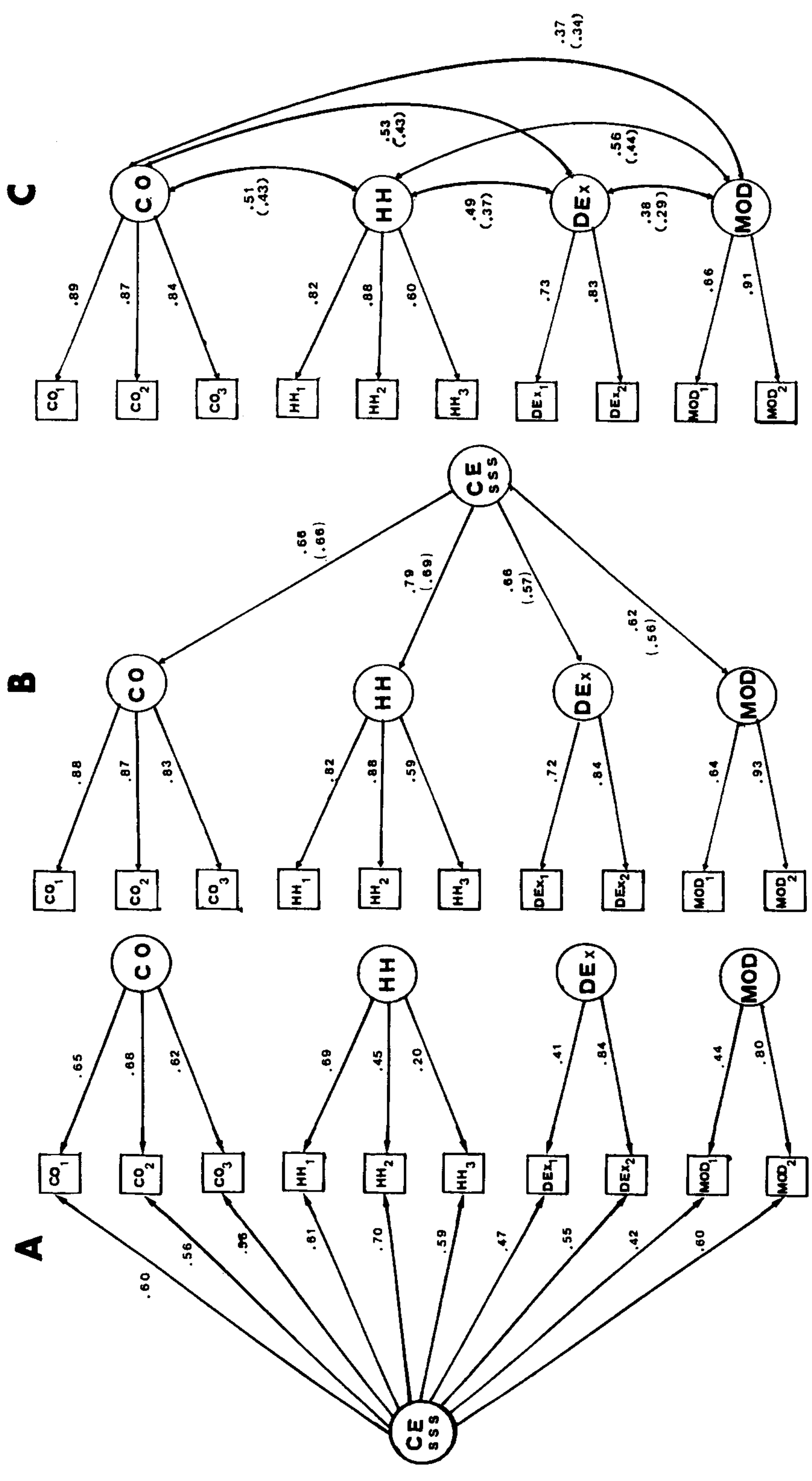

Figure 3. The three complementary models of the structure of causal–experimental thought. (The symbols CO, HH, DEx, and MOD stand for the combinatorial, the hypothesis handling, the experimentation, and the model construction ability, respectively. The symbol CE SSS stands for the causal–experimental Specialized Structural Systems.)

have to be equal in terms of measurement of the underlying construct. In as far as the first assumption is concerned, the reader is reminded that our system of scoring criteria rules out the possibility that a task would have been solved by guessing. In as far as the second assumption is concerned, the models shown in Figures 3A and 3B may be invoked. That is, the basically similar relation of all 10 scores to the first-order factor and of the 4 ability-specific factors to the second-order general factor strongly suggests that the items represented by these scores or factors did relate to the underlying SSS-specific construct in a similar way.

Figure 4 summarizes the results of this analysis. The horizontal dimension that structures this figure is the unified dimension on which the whole set of 22 items involved in all four batteries were calibrated. Thus, the position of each item in relation to the rest can easily be inspected. This position is defined by a logit score that reflects the difficulty of the item in relation to the other items. It must be underscored that this dimension was reliable. The person- and item-separation reliability index was .79 and .99, respectively. In the Guttman scaling sense, these indexes suggest that a subject of a given estimated ability did succeed on the items discriminating up to this ability or below. Alternatively, they can suggest that an item occupying a given position on the ability scale is performed by subjects found to have this or "more" ability. Evidently, this finding is in agreement with Prediction 2.

Inspection of Figure 4 suggests that the various abilities are intertwined in interesting ways during development. Specifically, the combinatorial abilities lie at the bottom of the hierarchy. The simpler experimentation, hypothesis-formation, and hypothesis-evidence mapping items, together with complex combinatorial skills, were placed in the middle of the hierarchy. The model construction, complex experimental, and epistemological items were placed at the top of the hierarchy. This placement of the various abilities along the scale is consistent with the between-factors correlations shown in Figure 3C. That is, the correlation between the factors associated with items placed at the two extremes of the developmental scale (model construction vs. combinatorial) was lower than the correlation between factors associated with items adjacent to each other on the developmental scale (see Model 3C). Thus, these findings are in line with Prediction 3.

Attention must be drawn here to factors that define developmental ascension. Complexity as defined by the number of the elements involved in a task is certainly a factor. It can be seen that tasks involving more elements proved to be generally more difficult than tasks involving less elements (see the combinatorial items or the experimentation tasks). This is evidently in agreement with our (Demetriou et al., 1993) and other theories (Case, 1985; Halford, in press; Pascual-Leone, 1988) claiming that the quantitatively defined processing capacity constrains the kinds of problems that can be solved at a given age. However, the intuitiveness of the relations involved was an equally strong factor. It can be seen that Item 2 in the combination of paths battery was more difficult than Item 1, although they both involved the same number of paths. Likewise, Item 6 in the combinations of balls battery was more difficult than the corresponding Item 5, although it involved fewer combinations. In both cases, the more difficult tasks were organized in a way that made the relations involved not readily representable. Thus, the strategies that direct the thinker how to represent different problems so as to circumvent their deceptiveness is no less important than the sheer expansion of processing capacity. The generation of these strategies is discussed in the General Discussion.

The results shown in Figure 4 are particularly relevant to an issue that is still hot in developmental psychology. This is the continuity–discontinuity issue. Some theorists, especially those working in a neo-Piagetian tradition, believe that development ascends along a sequence of distinct, qualitatively different stages (e.g., Case, 1985; Halford, in press). Others, particularly those associated with the novice–expert paradigm (e.g., Carey, 1985; Chi, 1978), claim that development is continuous and that no sharp change boundaries really exist. Finally, a third group of scholars adopt a compromise position. For example, Fischer (Fischer et al., 1984; Fischer & Silvern, 1985) argues that development is discontinuous between major stages and continuous within stages.

To obtain evidence directly related to this issue, the Saltus model recently proposed by Wilson (1989) was applied on the present data. The Saltus model is an extension of the Rasch model. It enables developmental researchers to test whether the change in difficulty from item to item is the same all along the scale generated by Rasch scaling or whether there is one or more regions on the scale at which changes are significantly greater than before or after these regions. In other words, the model aims to test if there are points on the scale that differentiate between levels. In such a case, differences between adjacent items belonging to different levels would be greater than differences between adjacent items within a level. According to Wilson, in this case, one may argue that these regions define second-order discontinuities or boundaries between major developmental levels. If no such regions exist, discontinuity would be of the first order, and the simple Rasch model would be sufficient to represent it.

To test a Saltus model, one needs to specify the level membership of the items involved. For the present purposes, we assumed that the 22 items are developmentally structured in three levels. The broken lines in Figure 4 separate the levels to which each of the items was allocated. It may be noted here that the allocation of items to levels was based on the sequencing of the various abilities as prescribed in Prediction 3. Globally, the first, the second, and the third level may be described as the combinatorial, the hypothesis formation and testing, and the model construction-epistemological awareness level, respectively. The loglikelihood value for this model was 5,726.69. The loglikelihood value for the simple Rasch model was 5,737.45. Thus, the Saltus model increased the chi-square fit by only 10.26 at the cost of 46 more parameters. This improvement is not significant. This finding strongly suggests that the development of the abilities investigated here is continuous rather than discontinuous. We elaborate more on this issue later on in the discussion.

Individual Differences

Structure. We anticipated that there would be no individual differences in the structure of abilities. To confirm this prediction, we tested the three models shown in Figure 3 separately across two age groups (junior secondary school subjects in

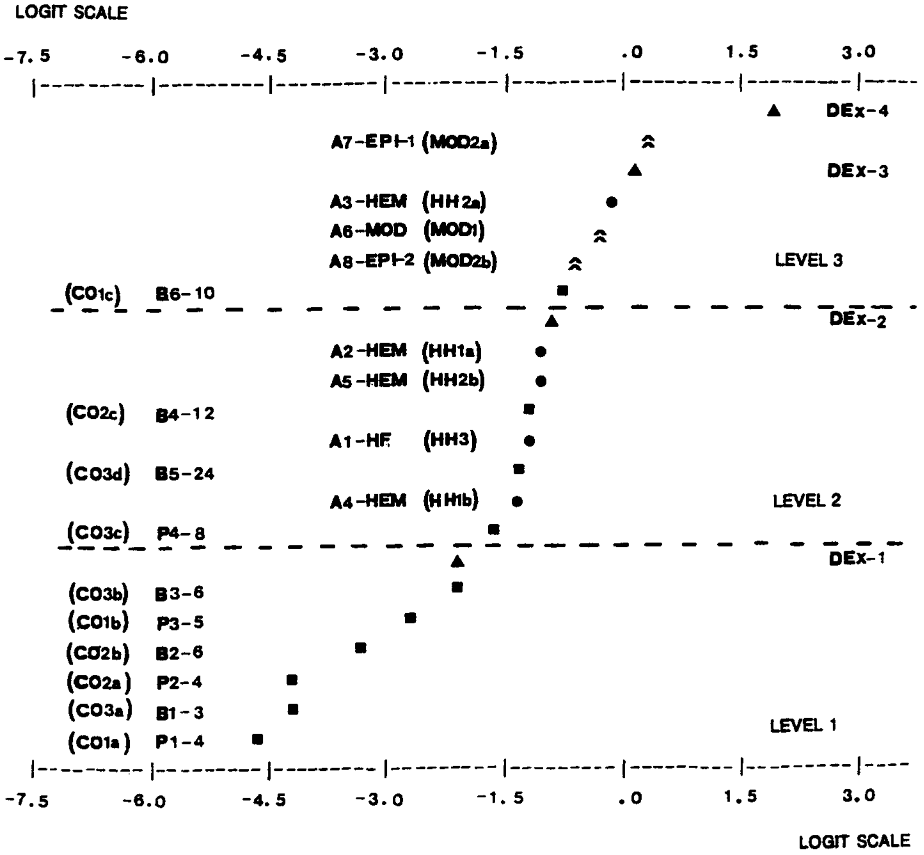

Figure 4. The scaling of the 22 causal–experimental items by Rasch rating scale analysis. (The symbols P and B refer to the combinations of paths and balls, respectively. The first number refers to the item number and the second to the number of combinations involved in the item. The symbol A stands for the agriculturist hypothesis-handling battery. The symbols HF, HEM, MOD, and EPI refer to hypothesis formation, hypothesis-evidence mapping, model construction, and epistemological awareness, respectively. Symbols in parentheses are the same as in Figure 3. The number attached to each of these symbols indicates the mean into which it was collapsed. The symbols ■, ▲, ●, and ^ stand for the combinatorial, the experimentation, the hypothesis handling, and the model construction items, respectively.)

Grades 7, 8, and 9 and senior secondary school subjects in Grades 10 and 11, averaged across sex and SES), two sex groups (averaged across age and SES), and two SES groups (the group of high SES vs. the other two SES groups, averaged across age and sex). The fit was excellent for all three models across all groups, as shown by the fact that all chi-squares were very low ($p > .05$ and CFI > .95 in all cases). From a technical point of view, these results suggest that the excellent fit of the models shown in Figure 3 was not caused by variability due to age, sex, or SES. From a substantive point of view, the results strongly suggest, in agreement with Prediction 4, that the structure of causal–experimental abilities is universal over the age, sex, and SES groups tested in this study.

Development. Rasch analysis provides a logit score for each subject that indicates his attainment of the abilities represented by a given dimension. Given that the unified causal–experimental thought dimension that emerged out of the present study was very reliable, the individual logit score was subjected to a 5 (age) × 2 (gender) × 2 (SES) analysis of variance. This aimed to uncover the effect exerted by the various individual difference factors on the attainment of causal–experimental abilities. Figure 5 summarizes mean logit attainment across age and SES. It may be noted that the logit score associated with a given group of subjects indicates that the subjects in this group solve the tasks that, according to scaling of the items (i.e., according to their position on the dimension shown in Figure 4), are associated with this or a lower logit.

It can be seen that the main effect of age, $F(4, 259) = 7.268$, $p = .000$, and SES, $F(1, 259) = 38.060$, $p = .000$, was highly significant. The Age × SES interaction was also significant, $F(4, 259) = 2.627$, $p = .035$. No sex effect was found. Inspection of Figure 5 makes it clear that the modal developmental level at the age of 12 and 13 years for the high SES subjects is middle-level combinatorial abilities and the first-level experimentation ability. There is a spurt between 13 and 14 years of age, which raises these subjects to the level of complex combinatorial abilities, second level of experimentation abilities, and the lower levels of hypothesis formation and hypothesis-evidence mapping abilities. The complex experimentation and model construction abilities, let alone epistemic awareness regarding the status of theories, is well beyond the reach of even the 16-year-old adolescents. Interestingly enough, the low SES subjects stay at the level of the lowest combinatorial abilities until the age of 15 years. However, they spurt somewhere between 15 and 16 so that, by 16, they almost catch up with their high SES peers. Therefore, the results came in agreement with Prediction 5, which anticipated SES but no sex differences in the acquisition rate of the causal–experimental abilities.

General Discussion

At a general level, the results of this study are concerned with the nature of (a) cognitive structures, (b) developmental processes, and (c) individual differences in regard to structure and development. At a more specifialized level, this study is concerned with causal–experimental thought. In fact, the ultimate aim of the study was to test a number of general assumptions about the human mind through the study of a specialized system of thought without losing sight of either the general or the more specific characteristics of mind. This aim directs the discussion of the three following issues.

Structure: The Composition of SSSs

What is an SSS in general and the causal–experimental SSS in particular? The alternative models shown in Figure 3, which have been verified for all five SSSs described by the theory (Demetriou et al., 1993; Demetriou et al., 1991; Efklides, Demetriou, & Metallidou, 1992), suggest that an SSS has to be viewed as a complex functional entity that involves three discrete kinds of constructs.

First, it involves a core or nuclear component that resides, in one form or another, in every particular skill or ability that an SSS may be composed of. This core may be more an attitude or orientation to the processing of problems rather than a specifiable pattern of cognitive activity. For instance, in the case of the causal–experimental SSS investigated here, this core may be identified with the analytical orientation to problems referred to in the introduction. In structural models, this construct is represented by the general factor that is of the same order with the ability specific factors (see the $_{cr}$SSS factors in Figure 1).

The second kind of constructs involved in an SSS can easily be specified. They are the component abilities themselves. As assumed by the theory, the causal–experimental SSS involves combinatorial, hypothesis handling, experimentation, and model construction abilities. Each of these abilities has to be viewed as a rather extended universe of tightly interconnected skills or patterns of cognitive activity that preserve, nevertheless, a considerable degree of functional autonomy for each ability. For example, one may be quite proficient in combinatorial skills but not in experimentation. Likewise, one may be a good experimenter but not a very good model inventor and vice versa.

It is proposed that there are good reasons behind the relative functional autonomy of the integral component abilities involved in an SSS. Specifically, different component abilities usually represent different phases in the problem-solving process. For instance, combinatorial abilities need to be activated first to enable one to define alternative causal paths between a given effect and a series of factors before a hypothesis can be formulated that one specific path is more probable than another. Moreover, there needs to be such a hypothesis, self-ascribed or given, if an experiment is to be designed that would be able to test it. Finally, at least one hypothesis must be tested by at least one experiment if a model is to be constructed. Therefore, each of the component abilities involved in an SSS serves a specialized function in the representation and understanding of the reality domain to which the given SSS is connected, which is distinct from that of the others.

The third kind of construct that one must invoke to under-

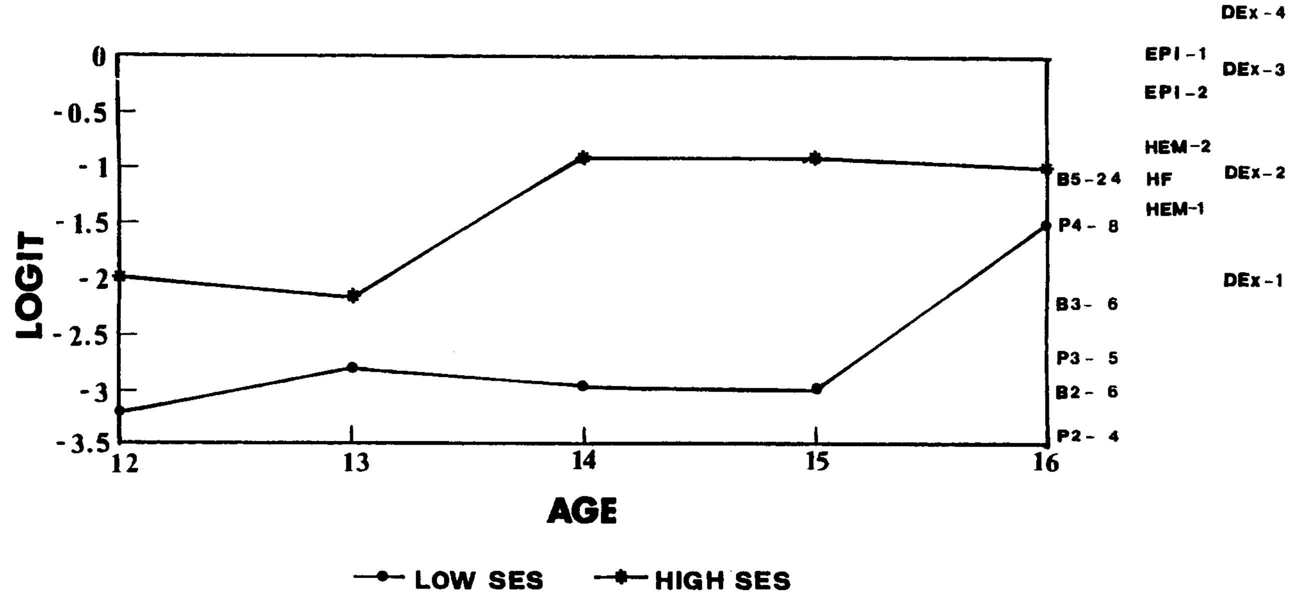

Figure 5. The mean logit attainment of causal–experimental thought as a function of age and socioeconomic status (SES). (The symbols DEx, EPI, HEM, and HF refer to the experimentation, epistemological awareness, hypothesis-evidence mapping, and hypothesis formation, respectively, and indicate the relation between the item scale shown in Figure 4 and the subject variables involved in the study.)

stand the operation of an SSS refers to the relations that interconnect the different component abilities into a smoothly running unit. Technically, the second-order general factor in a structural model was considered as indicative of these relations (see Figure 3B and the $_{pr}$SSS factors in Figure 1). In fact, we were able to decompose this factor into the web of paths that shows how the four abilities are interrelated (see Figure 3C). How are these relations regulated? According to the theory, the five organizational principles referred to in the introduction represent the regulatory forces that direct particular abilities to be fused or interrelated into broader functional systems (Demetriou, in press; Demetriou & Efklides, 1988).

In conclusion, an SSS is regarded as requiring a general problem-solving orientation that generates and permeates a set of specialized abilities that are functionally tuned to each other under the guidance of the organizational principles. Why and how does an SSS develop? This is the question to be discussed next.

Development: Layers of Mind and Synergic Developmental Causality

According to the theory, developmental causality is a synergic force. That is, it is assumed that developmental change is possible precisely because the cognitive system consists of multiple structures so coordinated that they are in a sense functionally tuned to one another. Thus, a change in any of them at both the intersystemic and the intrasystemic level is able, given certain other conditions being met (see Demetriou, in press; Demetriou et al., 1993), to initiate a series of changes in the other structures to which it is connected. This conception of developmental causality is supported by the basically continuous nature of the development of abilities, as suggested by the findings presented earlier in this study: that is, the fact that different abilities seem to follow neither entirely successive nor entirely parallel, but rather overlapping developmental sequences. This makes the mind appear as though it is programmed in such a way that, in order to develop, it should always be in a state of some sort of imbalance between at least some of its integral parts. This condition ensures that even reasons that, by themselves, may appear insufficient to cause large reorganizations (such as, for example, a slight change in speed of processing or the acquisition of a memorizing strategy) are able to function in this way once they are integrated in a change loop that interlocks them in a reaction chain.

How might such a reaction chain function in the case of the causal–experimental SSS? To be able to envisage how this reaction chain might function, we first need to depict in some detail the general metaphor of developing mind from early adolescence to youth, as suggested by the present and our previous studies.

In early adolescence (about 11 to 13 years of age), the person becomes able to see through the veil of the apparent. The emergence of the combinatorial ability signifies the attempt of the person to fix his or her suppositions about the nonapparent in forms that would make possible their further manipulation, real or mental. However, the person is not yet in good command of either the suppositions or the process generating the suppositions. Thus, as an experimenter, the young adolescent is clumsy and susceptible to his or her preconceptions about the issue under investigation. At the same time, the young adolescent's understanding of his or her own cognitive organization is very inaccurate (Demetriou & Efklides, 1989; Demetriou et al., 1993). This is the *suppositional* mind.

The suppositional mind is transformed into an *inquiring* mind during middle adolescence (about 14 to 16 years of age). That is, the adolescent can now formulate his or her suppositions into alternative hypotheses and map them onto experiments capable of testing them. An accurate, although global, theory of cognitive organization may now be present. It is during this phase that the subjects were able to solve the most complex of the combinatorial items.

The inquiring mind is transformed into a *modeling* mind during late adolescence and youth. During this phase, the thinker becomes able to conceive of models capable of unifying other models and still recognize their limitations. By the end of this phase, the modeling process itself is viewed as an object of modeling. Thus, on the one hand, the inherent limitations of models can now be recognized. On the other hand, the personal theory of cognitive organization becomes detailed enough to involve premises associating different cognitive structures, such as our SSSs, with different processes and strategies (Demetriou et al., 1993, Study 3). As before, it was during this phase that the subjects were able to design the most complex experiments.

Therefore, the inquiring mind appeared to interlock, so to speak, the suppositional mind and the modeling mind that develop in succession. This pattern has been observed in the development of all other SSSs (Demetriou, in press; Demetriou et al., 1991; Efklides et al., 1992). We have assumed that the abilities that interlock other abilities during development may operate as the catalyst that enables the system to shift from the state of functioning characterizing the lower block of levels to that characterizing the higher one. Applying this model to the development of the causal–experimental SSS, we would argue that combinatorial abilities of the lower levels serve to set up the mental space in which the first hypotheses can be formulated. At a subsequent phase, hypothesis formation abilities enable the young adolescent to fix the combinations he or she conceives in structures that can function as frames that can direct his or her manipulations of reality. Thus, systematic experimentation begins to be possible. Once it is acquired, experimentation acts in two directions. On the one hand, it provides strategies that can be put into the service of the combinatorial ability itself. As a result, the construction of the higher level combinatorial abilities is facilitated. On the other hand, it generates a data space that has to be understood. Evidently, a good way to gain an understanding is to map this data space onto the space of the hypotheses that served as the starting point of experimentation. This is the beginning of model construction. In turn, when model construction advances to a certain level, it functions as a frame for the conception of more complex hypotheses and consequently for the design of more sophisticated experiments. These then may loop back by providing the material, external or mental, for the construction of more complex models, and so on and so forth.

This is equivalent to saying that, according to our theory, the multistructural nature of mind is one of the main causes of the development of mind. That is, the interplay of structures during development causes the differentiation between the cognitive structures or skills themselves as well as between the structures and the cognitive activity that generates the structures. This differentiation enables the person to deliberately make both his or her cognitive structures and cognitive activity the object of further cognitive activity. Thus, as assumed in the introduction, the interaction between the different components of an SSS elevates this SSS to ever higher levels of cognitive functioning. At the same time, it generates experience of experience, which is the basis for the development of the hypercognitive system. Evidently, the changes in the hypercognitive system are transferable over different domains. In conclusion, development is viewed here as the result of a kind of dialogue between mental spaces. The more they interact, the more they are differentiated from each other; in turn, the more they are differentiated, the better they become knowable by the thinker. In this process, the thinker acquires knowledge about reality and herself or himself.

A note of caution is in order here. One might argue that reference to layers brings in the stage concept again through the back door. This is certainly not the case. The notion of layer used here differs from the Piagetian (Piaget, 1970) or the neo-Piagetian notion of stage (Case, 1985; Fischer & Silvern, 1985; Halford, in press; Pascual-Leone, 1988) in at least two important respects. First, it refers to an SSS and not to the whole cognitive system. Two different SSSs may be operating on noncorresponding layers during the same age phase (Demetriou & Efklides, 1988; Demetriou et al., 1993). Second, and probably more important, the layers as used here are differentiated from each other in relative terms. That is, two successive levels may differ in their modal characteristics, but they also share common characteristics as they interpenetrate each other. In fact, this structural property of mind has been regarded here as a part of the dynamics underlying the continuous change of mind.

Individual Differences

In agreement with all our earlier studies (e.g., Demetriou & Efklides, 1985, 1989; Demetriou et al., 1991), this study suggested that there are no individual differences in the structural organization of cognitive abilities. However, there are individual differences in the rate of development of these abilities. Two main conclusions are suggested by these findings. First, the structures described by our theory seem to be universal. Therefore, they should be taken as the cognitive instruments that result from and direct our interactions with the different aspects of the real and the symbolic environment we live in. Second, we need to understand the factors responsible for the fact that development is slower in some groups of the population as compared with others. This is the first step for the development of methods that would minimize or remove the effects of these factors. Our studies (Demetriou et al., 1993, Study 2; Efklides et al., 1992), Shayer's (1988), and others' (see the studies reviewed by Goossens, 1992) represent an attempt to specify how the factors blocking the development of some social groups may be compensated for by specialized training. In conclusion, the evidence generated by the present study proved to be congruent with both the general representation of developing mind and the analysis of causal–experimental thought advanced by our theory.

References

Bentler, P. M. (1989). *EQS: Structural equations program manual.* Los Angeles: BMDP Statistical Software.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Case, R. (1985). *Intellectual development: Birth to adulthood.* San Diego, CA: Academic Press.

Case, R. (1992). *The mind's staircase.* Hillsdale, NJ: Erlbaum.

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73–96). Hillsdale, NJ: Erlbaum.

Connell, J. P., & Tanaka, J. S. (Eds.). (1987). Structural equation modeling. *Child Development, 58,* 1–175.

Demetriou, A. (1990). Structural and developmental relations between formal and postformal capacities: Towards a comprehensive theory of adolescent and adult cognitive development. In M. L. Commons, F. A. Richards, C. Armon, & J. Sinnot (Eds.), *Beyond formal operations: Vol. 2. The development of adolescent thought and perception* (pp. 147–173). New York: Praeger.

Demetriou, A. (in press). On the quest of the functional architecture of developing mind. *Educational Psychology Review.*

Demetriou, A., & Efklides, A. (1985). Structure and sequence of formal and postformal thought: General patterns and individual differences. *Child Development, 56,* 1062–1091.

Demetriou, A., & Efklides, A. (1988). Experiential structuralism and neo-Piagetian theories: Toward an integrated model. In A. Demetriou (Ed.), *The neo-Piagetian theories of cognitive development: Toward an integration* (pp. 173–222). Amsterdam: North-Holland.

Demetriou, A., & Efklides, A. (1989). The person's conception of the structures of developing intellect. *Genetic, Social, and General Psychology Monographs, 115,* 371–423.

Demetriou, A., & Efklides, A. (in press). Dynamic patterns of intra- and inter-individual change in the acquisition of complex thinking abilities from 10 to 17 years of age: A longitudinal study. *Scientific Annals of the Department of Psychology of the Aristotelian University of Thessaloniki* (Greece).

Demetriou, A., Efklides, A., & Platsidou, M. (1993). Experiential structuralism: A frame for unifying cognitive developmental theories. *Monographs of the Society for Research in Child Development, 58*(5, Serial No. 234).

Demetriou, A., Gustafsson, J.-E., Efklides, A., & Platsidou, M. (1992). Structural systems in developing cognition, science, and education. In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *The neo-Piagetian theories of cognitive development: Implications and applications for education* (pp. 79–103). London: Routledge.

Demetriou, A., Platsidou, M., Efklides, A., Metallidou, Y., & Shayer, M. (1991). The development of quantitative-relational abilities from childhood to adolescence: Structure, scaling, and individual differences. *Learning and Instruction: The Journal of the European Association for Research in Learning and Instruction, 1,* 19–43.

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

Efklides, A., Demetriou, A., & Metallidou, D. (1992). *The structure and development of propositional thought.* Unpublished manuscript, Aristotelian University of Thessaloniki, Department of Psychology.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87,* 477–531.

Fischer, K. W., & Farrar, M. J. (1988). Generalizations about generalization: How a theory of skill development explains both generality and specificity. In A. Demetriou (Ed.), *The neo-Piagetian theories of cog-*

nitive development: Toward an integration (pp. 137–171). Amsterdam: North-Holland.

Fischer, K. W., Hunt, H. H., & Russell, S. L. (1984). The development of abstractions in adolescence and adulthood. In M. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations* (pp. 43–73). New York: Praeger.

Fischer, K. W., & Silvern, L. (1985). Stages and individual differences in cognitive development. *Annual Review of Psychology, 36,* 613–648.

Flavell, J. H. (1988). The development of children's knowledge about the mind: From cognitive connections to mental representations. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 244–267). Cambridge, England: Cambridge University Press.

Flavell, J., & Wohlwill, J. F. (1969). Formal and functional aspects of cognitive development. In D. Elkind & J. H. Flavell (Eds.), *Studies in cognitive development: Essays in honor of Jean Piaget* (pp. 67–120). New York: Oxford University Press.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Goossens, L. (1992). Training scientific thinking in children and adolescents: In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *The neo-Piagetian theories of cognitive development: Implications and applications for education* (pp. 161–179). London: Routledge.

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

Gustafsson, J.-E. (in press). Hierarchical models of intelligence and educational achievement. In A. Demetriou & A. Efklides (Eds.), *Intelligence, mind and reasoning: Structure and development.* Amsterdam: Elsevier.

Halford, G. S. (1988). A structure-mapping approach to cognitive development. In A. Demetriou (Ed.), *The neo-Piagetian theories of cognitive development: Toward an integration* (pp. 103–136). Amsterdam: North-Holland.

Halford, G. (in press). *Children's understanding: The development of mental models.* Hillsdale, NJ: Erlbaum.

Humphreys, L. G. (1976). A factor model for research on intelligence and problem-solving. In L. R. Resnick (Ed.), *The nature of intelligence* (pp. 329–339). Hillsdale, NJ: Erlbaum.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* London: Routledge & Kegan Paul.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12,* 1–48.

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

Landau, B., Spelke, E. S., & Gleitman, H. (1984). Spatial knowledge in a young blind child. *Cognition, 16,* 225–260.

Martorano, S. H. (1977). A developmental analysis of performance on Piaget's formal operations tasks. *Developmental Psychology, 13,* 666–672.

McLaughlin, G. H. (1963). Psycho-logic: A possible alternative to Piaget's formulation. *British Journal of Educational Psychology, 33,* 61–67.

Moshman, D. (1979). To really get ahead get a metatheory. In D. Kuhn (Ed.), *Intellectual development beyond childhood* (pp. 59–68). San Francisco: Jossey-Bass.

Neimark, E. D. (1975). Intellectual development during adolescence. In J. Horowitz (Ed.), *Review of child development research* (Vol. 4, pp. 541–594). Chicago: University of Chicago Press.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32,* 301–345.

Pascual-Leone, J. (1988). Organismic processes for neo-Piagetian theories: A dialectical causal account of cognitive development. In A. Demetriou (Ed.), *The neo-Piagetian theories of cognitive development: Toward an integration* (pp. 25–64). Amsterdam: North-Holland.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Charmichael's manual of child psychology* (Vol. 1, pp. 703–732). New York: Wiley.

Piaget, J. (1973). *Structuralism.* London: Routledge & Kegan Paul.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children.* London: Routledge & Kegan Paul.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests.* Chicago: University of Chicago Press.

Shayer, M. (1979). Has Piaget's construct of formal operational thinking any utility? *British Journal of Educational Psychology, 49,* 265–276.

Shayer, M. (1988). Neo-Piagetian theories and educational practice. In A. Demetriou (Ed.), *The neo-Piagetian theories of cognitive development: Toward an integration* (pp. 245–266). Amsterdam: North-Holland.

Shayer, M., Demetriou, A., & Pervez, M. (1988). The structure and scaling of concrete operational thought: Three studies in four countries. *Genetic, Social, and General Psychology Monographs, 114,* 307–376.

Spelke, E. S. (1984). Perception of unity, persistence, and identity: Thoughts on infants' conceptions of objects. In J. Mehler & R. Fox (Eds.), *Neonate cognition: Beyond the blooming, buzzing confusion* (pp. 409–430). Hillsdale, NJ: Erlbaum.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction in human infants. *Cognition, 36,* 97–127.

Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58,* 134–146.

Vernon, P. E. (1961). *The structure of human abilities.* London: Methuen.

Wilson, M. (1989). Saltus: A psychometric model of discontinuity in cognitive development. *Psychological Bulletin, 105,* 276–289.

Wellman, H. (1990). *The child's theory of mind.* Cambridge, MA: MIT Press.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis.* Chicago: Mesa Press.

Received July 16, 1991
Revision received July 28, 1992
Accepted July 29, 1992 ■