

The Development of Inductive Reasoning: Cross-sectional Assessments in an Educational Context

Benő Csapó

Attila József University, Szeged, Hungary

This paper links two research paradigms, one that studies attributes and mechanisms of inductive reasoning and one that tries to make school learning more meaningful and knowledge better understood and more easily applied, by examining how inductive reasoning develops during a significant age range of schooling and how it relates to certain other cognitive functions. Six tests of inductive reasoning (number analogies, verbal analogies, number series, verbal series, coding, exclusion) were devised and administered to 3rd, 5th, 7th, 9th, and 11th grade students ($N > 2000$). Data were also collected on students' school achievement, and a test of applied science knowledge was administered to the two oldest samples. The comparison of age groups indicated that the fastest development of inductive reasoning took place between the 5th and 9th grades; a major development was detected before the 5th grade, and only modest changes were found after the 9th grade. Regression analysis models indicated that inductive reasoning accounted for around twice as large a proportion of the results of the test that measured the applied science knowledge in everyday situations as did school knowledge (represented by school grades).

Inductive reasoning is considered a basic component of thinking, and it is one of the most broadly studied procedures of cognition. The inductive method, or teaching by examples, is one of the oldest methods of instruction. In addition, induction, or rather its role in generating scientific knowledge, is

Requests for reprints should be sent to Benő Csapó, Department of Education, Attila József University, H-6722 Szeged, Petőfi sgt. 30-34, Hungary.

The data collection was carried out in the framework of the project "Integrating learning procedures of declarative knowledge and development of higher order thinking skills", supported by the Hungarian National Science Foundation (OTKA-1622, T018577). Preliminary results of this project have already been presented at conferences (Csapó, 1993) and were published in Hungarian (Csapó, 1994). This paper was completed while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences. I am grateful for financial support provided by the Andrew W. Mellon Foundation and the Johann Jacobs Foundation.

one of the most enduring problems of philosophy. Although recent cognitive research has resulted in a vast body of new information about inductive reasoning and has refined our understanding of its functioning, school instruction still seems far from using it effectively and several problems can be attributed to this scanty attention paid to the role of induction in learning. In this study we examine how inductive reasoning develops during a significant age range of schooling and how it relates to certain other cognitive functions. We try to estimate its contributions to school achievement and to the application of knowledge in schools today.

The centrality of inductive reasoning can best be demonstrated by the fact that many research schools have dealt with it, in the frameworks of several research paradigms, and it has been discussed in relationship to most of the higher-order cognitive skills. The mechanism of rule induction was central in some early information-processing models of *problem-solving* (Egan & Greeno, 1974; Simon, 1974). Inductive reasoning, or its components, have often been related to or identified with *general intelligence*. For example, Sternberg (1977) discussed analogical reasoning in this context, and several intelligence tests contained inductive reasoning tasks or consisted exclusively of them. Various conceptualisations have interpreted inductive reasoning by reformulating the classical inductive-deductive categorisation and have described it by contrasting or comparing it with *deductive reasoning* (Ennis, 1987; Johnson-Laird, 1988; Shy, 1988; Sternberg, 1986). In these systems and taxonomies, inductive reasoning appears as one of the components of *critical thinking* (Ennis, 1987) or is discussed in a framework that also accommodates *creativity* (together with associations, computing, and deduction; see Johnson-Laird, 1988). Inductive reasoning may be one of the mechanisms of hypothesis generating and *hypothesis testing* (Gilhooly, 1982), as well as a means of *concept development* (Egan & Greeno, 1974; Gelman & Markman, 1987; Markman, 1989).

The significance of induction in the acquisition of personal knowledge is accentuated by the diversity of research on inductive reasoning carried out in close relationship with learning. Holland, Holyoak, Nisbett, and Thagard (1986) presented a comprehensive theoretical analysis and a number of studies have examined its specific functions. Depending on the paradigm and approach, inductive reasoning is interpreted as one of the basic *learning abilities* (Pellegrino & Glaser, 1982), or *learning skills* (Ropo, 1987); whereas some recent works use inductive tasks to measure *learning potential* (Resing, 1993; Tissink, Hamers, & van Luit, 1993).

Inductive thinking, especially analogical reasoning, is a means of *transfer*—applying knowledge acquired in one context in new situations. Contemporary cognitive research has revealed that the majority of our knowledge is content-bound or domain-specific; it does not transfer easily from one field to another. Wason's famous card task (1968) inspired

numerous studies that proved that knowledge demonstrated in one problem does not transfer automatically even into a structurally isomorphic problem. On the other hand, transfer obviously does exist, and analogies are the most frequently studied forms of transfer (Gick & Holyoak, 1983; Holyoak, 1985). Experiments aiming directly at transfer mechanisms (Klauer, 1989; Phye, 1989, 1990) or inductive reasoning in general (Klauer, 1990) have shown that inductive reasoning can be improved. Thus, through improving inductive reasoning we can also improve learning.

A large number of studies have examined the structure of inductive reasoning by using various types of tasks (most frequently analogies, series completions, and classifications) with different contents (e.g. verbal, figural). Most of the data sets were highly intercorrelated; for example, the study by Sternberg and Gardner (1983, p. 106) resulted in the conclusion that: "... highly similar processes are used on the various tasks, but these processes operate on different knowledge stores and possibly different forms of representation for the different contents." These examinations were carried out under different circumstances (mainly in laboratory conditions) and different techniques of data collection (e.g. measuring response time) were used than our planned measurements in educational contexts. Therefore, it seemed plausible: (1) to devise a test battery of different types of inductive tasks with various contents; (2) to examine whether they highly intercorrelate; and if so (3), to represent inductive reasoning with one single variable for further computations.

The concepts just highlighted cover nearly every relevant cognitive aspect of school instruction, and as they relate to educational applications, the results of previous research have great value. Many of them have shaped our understanding. However, without further field work they can hardly be applied in real-life instructional situations. Others, especially some results of the most recent research, can easily find their way into instructional applications.

The paradigm shift formulated by Kuhn (1989, p. 261) can also be observed in research on inductive reasoning:

It was not long ago that relevant research was strictly second class. Research that related directly to real world issues and problems was labeled "applied", and strongly implied by that label was lack of rigor. Today, in contrast, we are concerned that our research be relevant. Competently designed and executed research programs are vulnerable to criticism and dismissal on the grounds of irrelevance, and the priority of our concerns regarding internal versus external validity has largely been reversed.

Several recent works clearly mark this new orientation (e.g. Halford & Boulton-Lewis, 1992; Klauer, 1990; Resing, 1993; Tissink et al., 1993).

Use of knowledge has become an increasingly important field of research that approaches the problem of meaningful learning from another direction: The improvement of school instruction. Several studies have indicated that the type of achievement schools acknowledge and praise may be different from the type that can be used in a practical context. For example, Nunes, Schliemann, and Carraher (1993) distinguish “street mathematics” and “school mathematics”; Sternberg and Wagner (1994) organise the contexts of cognitive functioning under the titles “academic tasks” and “everyday tasks”; and Gardner (1991) distinguishes three types of achievements performed by three types of students (the “intuitive learner”, the “traditional student”, and the “disciplinary expert”) and describes the gaps between them.

In our interpretation, inductive reasoning is one of the mental tools that is used not only to acquire new knowledge but also to make the acquired knowledge more readily applicable in new contexts. Despite efforts on both sides, there is still a gap between studies that describe inductive reasoning on the one hand and those that try to make school learning more meaningful and knowledge better understood and applied on the other hand. In this study we relate these two research paradigms by assessing the developmental level of inductive reasoning in five age groups in a school context and estimating its relationships to some important variables of school instruction.

To examine the role inductive reasoning plays in present school learning, we distinguished two types of performance: (1) “school knowledge”; and (2) “applicable knowledge”. School knowledge is the type of knowledge that is mastered at school and is performed in school-related situations, usually in the same context in which it has been acquired. It is the basis of the usual assessments (quizzes, tests, exams) carried out at schools, it forms the basis of grading, and it determines a student’s career within the school system. Applicable knowledge is the form of knowledge that can be used in situations other than the one in which it was acquired. We presume that: (1) school knowledge does not transfer automatically into applicable knowledge, so those students whose school performance indicators are high do not necessarily perform well in application tasks; and (2) inductive reasoning is important in the transfer procedures, so students whose inductive reasoning is more developed will be better able to apply their knowledge.

As for the more specific aims of this study, we hypothesise that higher-order thinking skills can best be trained when their development is still in progress, especially when they are in a fast-growing phase. This assumption is a generalisation from the finding of our previous training programme that operational abilities that were still in their growing phase could be better modified by training (Csapó, 1992). Through cross-sectional

measurements, we outline the developmental tendencies and estimate the age range when the major phase of development takes place, and we try to locate those periods where we can expect the best effects from a modified school instruction. By analysing the relationship between inductive reasoning and certain other variables of school instruction, we try to decide what modifications of school instruction should be proposed and what benefits could be expected from improved instructional practices.

METHODS

Samples

The type of problems studied here did not require composing nationally representative samples; however, we needed samples that were large enough to bear the major typical characteristics of schooling. Thus, samples were drawn from the schools of Szeged (one of the major cities of Hungary) and its metropolitan area. Whole school classes were chosen for group testing, and we tried to achieve the best representation of schools in the area in terms of quality and type of schooling. Other surveys on school achievement have indicated that the target area of our study did not differ much from the national norms in terms of school achievements.

The eight-year primary schools (schooling begins at the age of 6 years) use the same centrally developed curriculum and the same set of textbooks, and they are fairly homogeneous. There are three types of secondary schools in Hungary (grammar school or *gymnasium*, technical school, and vocational school), and in order to achieve a representative sampling we chose larger samples in secondary schools (9th and 11th grades). The earliest age when our tests can be used is the 3rd grade of primary school; below this age we faced reading difficulties. The oldest age group that can be tested within the educational system is the 11th grade. Between these two points, measurements took place in the odd-numbered grades. Further data on the samples are summarised in Table 1.

TABLE 1
The Samples of the Survey

<i>Grade</i>	<i>n</i>	<i>Boy/Girl Ratio (%)</i>	<i>Mean Age (yrs)</i>	<i>Mean Grades</i>	<i>No. of Classes</i>
3	373	46.3/53.7	9.4	4.2	16
5	410	44.3/55.7	11.4	3.8	14
7	290	50.8/49.2	13.5	3.5	12
9	650	47.2/52.8	15.4	3.5	24
11	701	51.8/48.2	17.4	3.4	19
Total	2424	—	—	—	85

Instruments

Although several tests for inductive reasoning were described in the literature, we decided to develop new ones. The reason for this decision was that: (1) verbal tests may translate poorly; and (2) our tests were prepared for further use in other surveys and training experiments where we need very sensitive and reliable measurement instruments.

Using the descriptions found in the literature and our experience with similar tests, we devised a battery of six tests. The *number analogies*, *number series*, *letter series*, and *coding* tests consist of open-ended tasks, whereas the *verbal analogies* and *exclusion* tests are composed of multiple-choice items. Sample items of the inductive reasoning tests are presented in Table 2.

Although we did not devise figural tests, we adapted one of the most widely used intelligence tests, Raven's progressive matrices. From previous measurements we knew that the standard set (SPM) was too easy for the older age groups, so we composed a new scale by choosing the more difficult items from the standard set and adding items from the more complex Set II. The reliability coefficients (Cronbach α) of the version of tests used in the present study can be found in Table 3.

To examine various cognitive processes related to inductive reasoning, we used two tasks, both found in the literature. Wason's four-card task (Wason, 1968) has been used in several studies; the proportion tasks were described, for example, by Lawson, Karplus, and Adi (1978). The four-card task can be considered as a bridge between inductive and deductive reasoning. On the one hand, it has an explicit solution, which can be found through logical reasoning. On the other hand, the solution is to find the cards that can falsify

TABLE 2
Sample Items of the Inductive Reasoning Tests

Number analogies

14→17 :: 18→21 :: 23→_____

Verbal analogies

CHAIR : FURNITURE = DOG : ?

a CAT b ANIMAL c SPANIEL d TABLE e DOGHOUSE

Number series

1 7 13 19 25 31 _____ _____

Letter series

a c e g i k m _____ _____

Coding

Models: Monday + Wednesday = Thursday; Tuesday + Thursday = Saturday

Task: Wednesday + Thursday = _____

Exclusion

a SCARF b SHOES c HAT d CHAIR e SHIRT f GLOVES

TABLE 3
Reliability of Tests used in the Study

<i>Test</i>	<i>No. of Items</i>	α
Number analogies	14	.92
Verbal analogies	36	.94
Number series	16	.85
Letter series	14	.91
Coding	6	.86
Exclusion	18	.71
Total inductive items	104	.97
Raven	42	.95

a given rule, and falsification or finding a counter example is one of the main processes of induction. (Or *the* process, at least in the terms of philosophy, as Popper, 1972, argued.)

In the proportion task, a picture of two containers with different diameters was presented, indicating that the water that occupied 4 units in the wide container occupied 6 units when poured into the narrow one. Students were asked to estimate how high another quantity of water would be in the thin container, if it originally occupied 6 units in the wide one.

In this study “school knowledge” was represented simply by the data that are most commonly used in the Hungarian schools as an indicator of overall academic school performance: the means of the grades given by the teachers (comparable to the Grade Point Average in the United States). It involves all of the values on which teachers judge their students’ knowledge, plus the teachers’ subjective perceptions and biases. Teachers may use five numeric values to express their judgements with “1” being the worst and “5” the best grade. These data are used throughout in the school system as a basic indicator of students’ overall academic performance and they determine students’ choices in the elementary-to-high school and high school-to-higher education transitions. The results of the semester preceding the administration of inductive reasoning tests were used in the data analysis.

To measure “applicable knowledge”, an applied science test was devised. The test consisted of 45 questions that required the application of scientific knowledge in everyday situations. Most of the questions were not directly discussed in the science courses. Nevertheless, each question could be answered correctly by applying knowledge learnt in science, but it could not be answered simply on the basis of everyday experiences. For example: “Why do we see the flash of lightning first and hear the sound of thunder only some time later?” (The reliability of the test: $\alpha = .90$.) This measure was not as far from school knowledge as real-life applications were (e.g. as far as

school mathematics is from street mathematics). It measured applicable knowledge in a school context in a test situation, but it excluded teachers' subjective judgements, and the test could not be solved by simply recalling the facts and figures of disciplinary knowledge.

Data Collection

The tests were administered to 7th, 9th, and 11th grade students during the spring of 1993. When the results showed that the tests were not too difficult even for the 7th grade students, tests were administered to the 3rd and 5th grade samples during the spring of 1994. Although the data were collected at different times, the measurements took place within two years, and we do not know of any relevant differences between the two testing periods. Therefore, we analysed the data as if they were the results of simultaneous cross-sectional measurements. The same procedures of group testing were applied in each case.

The applied science test was administered during spring 1994 to the students who had been in our 7th and 11th grade sample in 1993 when the inductive reasoning tests were given, thus, these students were in the 8th and 12th grade when we tested their science knowledge. Because the vocational schools end with the 11th grade, in the 12th grade we tested only the students who attend gymnasium or technical school. This data cannot be considered as representative for the entire cohort, but rather for the students who attend schools that make them eligible to enter higher education.

RESULTS AND DISCUSSION

Differences between the Age Groups

In the first step, we computed the test results as percentages of the maximum score. The means and standard deviations of the percentage scales are summarised in Table 4. (In the case of the two tasks, the means indicate the percentage of students giving the correct solution.) As a general tendency, larger standard deviations were found in the younger age groups. An analysis of the frequency distributions showed that this larger variability can be attributed mainly to a few students in the younger age groups who did much better than their peers.

The mean achievements of the students of grades 3 through 11 in the two tasks are depicted in Fig. 1. Only the curve of proportion task is similar to a regular, logistic-like developmental curve. The considerable increase in the percentage of correct answers can be attributed to school instruction, especially to the mathematics skills acquired during the elementary school years. The proportion task can be considered one of the simplest forms of numeric analogies, and similar improvement can be expected in the solution of numeric analogy tests.

TABLE 4
The Results of Tasks and Tests in Percentages of the Maximum Score

<i>Tasks/Tests</i>	<i>Grade 3 m (SD)</i>	<i>Grade 5 m (SD)</i>	<i>Grade 7 m (SD)</i>	<i>Grade 9 m (SD)</i>	<i>Grade 11 m (SD)</i>
Wason	6.3 (24.2)	6.0 (23.7)	7.4 (26.2)	10.1 (30.2)	10.2 (30.3)
Proportion	1.5 (12.1)	7.3 (26.0)	19.9 (40.0)	50.7 (50.0)	62.0 (48.6)
Number analogies	20.9 (23.5)	31.8 (29.2)	63.4 (20.8)	74.4 (17.9)	78.3 (17.2)
Verbal analogies	30.3 (16.4)	41.1 (21.2)	56.0 (23.2)	71.9 (21.3)	75.7 (19.7)
Number series	39.4 (18.2)	46.8 (19.9)	61.1 (20.2)	66.9 (19.9)	71.8 (18.2)
Letter series	45.1 (31.2)	63.1 (30.3)	77.9 (23.3)	85.0 (20.1)	88.5 (18.8)
Coding	31.0 (30.0)	45.6 (34.0)	56.4 (38.3)	69.0 (35.4)	74.4 (32.4)
Exclusion	31.1 (12.5)	39.5 (10.5)	46.1 (11.9)	53.8 (15.2)	58.4 (14.8)
Raven	36.6 (17.9)	52.5 (18.9)	60.3 (18.4)	76.6 (16.6)	81.9 (14.0)

The solution of Wason's task did not show a major improvement over the eight years covered by the measurements, and the results were hardly better than random guessing. (The probability of finding the correct solution by guessing is 6.25%.) This is not surprising, considering the results of numerous similar studies. However, because the new mathematics curricula emphasise logical operations, we expected somewhat higher achievements. This result indicates again the poor transfer and the content-bound nature of some reasoning skills.

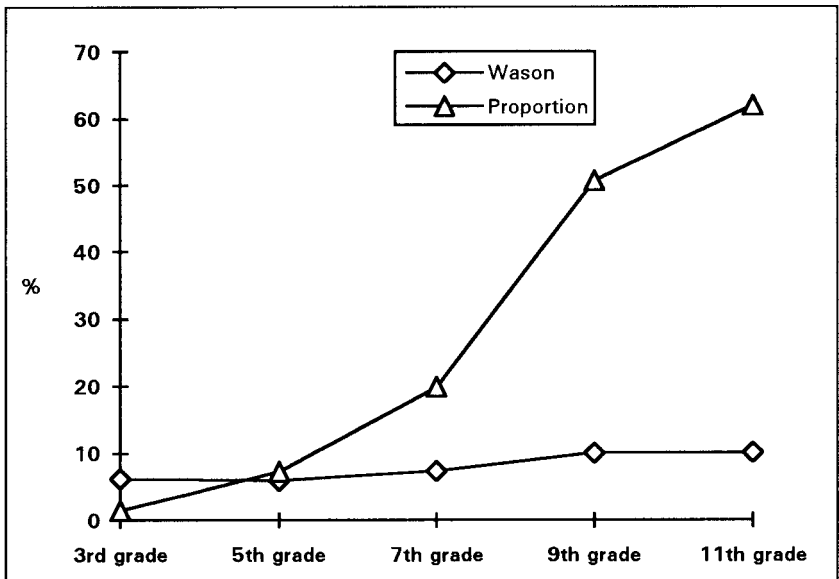


FIG. 1. Differences in the achievements of the age groups in the Wason and proportion tasks.

For a theoretically better established representation of age-related differences, we analysed test data by using the procedures of Item Response Theory (IRT). For each of the six inductive reasoning tests an IRT model was developed separately. We did the same for the Raven test, and then for all 104 items, as if they formed one single scale. This method offers a theoretically less ambiguous solution for summarising the measure of the developmental level of inductive reasoning in one single variable. In order to obtain a simpler and more purposeful scale, we made a linear transformation on the person parameters: 5 was added to the person parameters (expressed in logit units), and this sum was multiplied by 100. In this way we devised a scale on which the test results range approximately from 0 through 1000. (A similar scale was constructed to represent and compare science achievements in several countries; see Keeves, 1992.) The differences between the age groups in the result of inductive reasoning tests (represented on the scale described) are depicted in Fig. 2.

The curves of Fig. 2 indicate regular developments. Each of them can be identified with a segment of a logistic curve. The greatest differences between the younger and older samples can be observed in numeric analogies. In this case, especially in the younger children, not only did the inductive reasoning skills, which are required to solve numeric analogies develop rapidly but also their mathematical prerequisites. (See the similar curve for the proportion task, Fig. 1.) Although the curves run parallel at

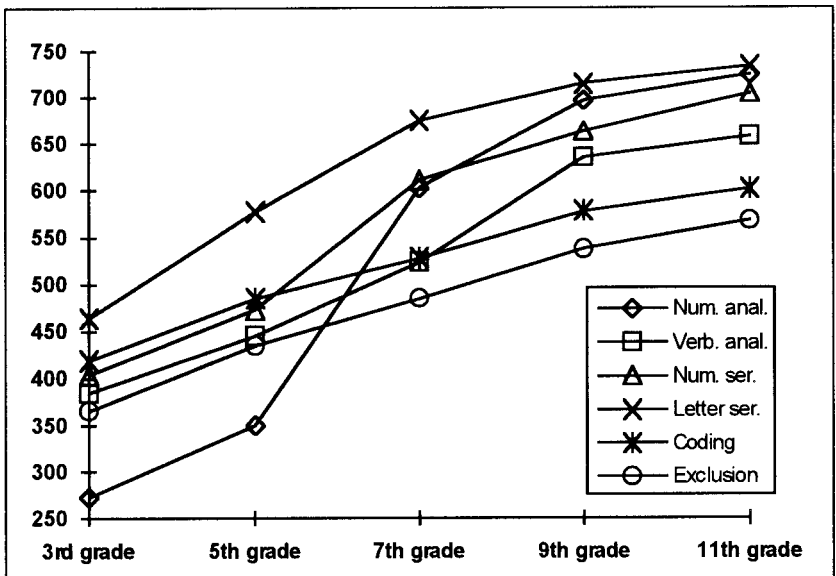


FIG. 2. Differences in the achievements of the age groups in the inductive reasoning tests.

both ends of the covered age range (between the 3rd and 5th grades and the 9th and 11th grades), some of them cross between the 5th and 9th grades. This middle age range seems to be a period of major structural reorganisation. The relative positions of achievement on tests with numeric content and with analogy operations improved similarly, so it is consistent with these observations that the greatest improvement was found in numeric analogies. The large age group differences in the numeric and analogical tasks can mostly be attributed to the influence of school learning.

The curves of Raven intelligence and inductive reasoning (based on the person parameters computed from the results of all 104 inductive items) are separately graphed in Fig. 3. Inductive reasoning showed a regular developmental trend. The largest change took place between the 5th and 7th grades. The turning point, where the accelerating improvement begins to slow down, can be found near the 6th grade. Furthermore, the curve also indicates that substantial development took place before the 3rd grade and some improvement of inductive reasoning can also be expected after the 11th grade.

Achievements on the Raven test showed minor irregularities in the 5th and 7th grades. The scores of 5th graders were somewhat higher and those of 7th graders were somewhat lower than could have been expected on the basis of a single logistic curve model. No specific explanation was found for

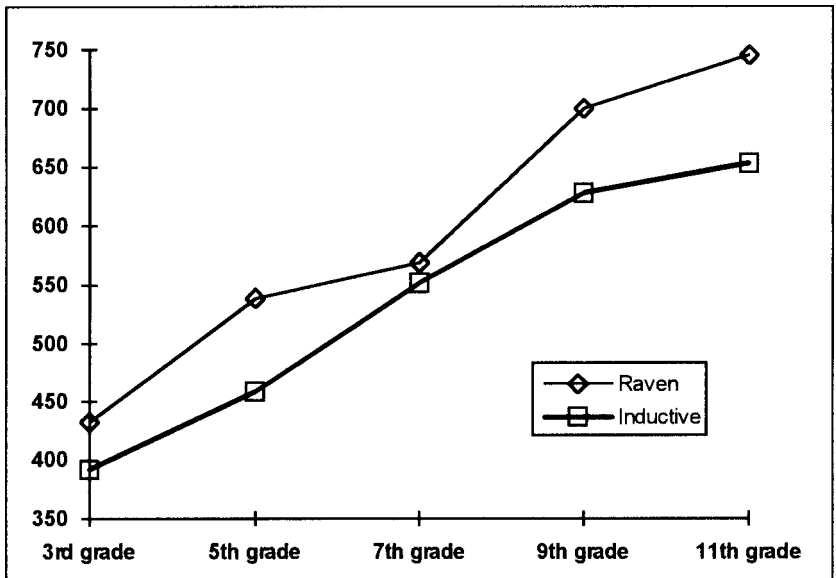


FIG. 3. Differences in the achievements of the age groups in the Raven intelligence and inductive reasoning tests.

this result. Data were carefully checked and no error was found. Because no such irregularities were found on other tests, sampling error can also be excluded, but statistical fluctuation may be considered part of the explanation (see the data in Tables 1 and 4 for computing the standard errors of measurement). On the other hand, this was the only test with figural content and this feature could contribute to its unique behaviour. The other tests indicated a structural reorganisation around this age (as do most developmental theories). If we hypothesise a developmental curve consisting of two consecutive logistic curves, results in the other measurement points seem to support such a hypothesis. Nevertheless, further investigations are necessary to find a proper explanation.

Correlations between Tests and Tasks

The correlations between the tasks and tests are presented in Table 5. Because of the poor results in the solution of Wason's task, all of its correlations were very low. The proportion task correlated at the highest level (.55 and .60) with the two analogy tests, as we also predicted.

The correlations between the tests were usually high, a result which is comparable with the similar data in the literature. For example, Pellegrino and Glaser (1982) examined verbal and numerical analogies and the correlations they found were between .55 and .74. Sternberg and Gardner (1983) reported somewhat higher correlations between solution times for analogy, series completion and classification tasks. In our study, in the case of those pairs of variables where either the inductive procedure (numerical analogies–verbal analogies, $r = .67$) or the content of tasks (numeric analogies–number series, $r = .62$, verbal analogies–exclusions, $r = .64$) was identical, the correlations were somewhat higher than the others.

TABLE 5
Correlation of Tasks and Tests

	1	2	3	4	5	6	7	8	9
1. Wason task									
2. Proportion	07								
3. Number analogies	07	55							
4. Verbal analogies	09	60	67						
5. Number series	08	48	62	60					
6. Letter series	06	38	58	55	60				
7. Coding	11	45	49	52	54	50			
8. Exclusion	07	50	57	64	57	51	48		
9. Raven	11	56	66	69	61	60	55	60	
10. All inductive tasks as test	10	64	82	90	79	75	67	77	78

Note: Decimal points were omitted; $N > 2000$; correlation coefficients over 0.07 are significant at the $P < .001$ level.

Correlations were slightly higher with Raven intelligence (ranging from .55 through .69) than the inductive tests correlated with each other. These high correlations (as well as the even higher Raven—all inductive items correlation, $r = .78$) empirically support the assumption that the Raven test can be used to measure the developmental level of inductive reasoning (besides the theoretical arguments based on the inductive nature of the progressive matrices). On the other hand, the differences between the correlations also indicate that specific contents and operations of the tests also play an important role. Especially in an educational context, the concept of inductive reasoning may be more useful than the abstract concept of general intelligence.

Knowledge and Inductive Reasoning

To examine the role inductive reasoning may play in learning, especially in meaningful learning that results in well understood and applicable knowledge, we analysed the relationships of inductive reasoning with two indicators of performance. The correlations of “school knowledge” (represented by the mean of grades given by teachers), applicable science knowledge (measured by the applied science test), inductive reasoning and Raven intelligence are summarised in Table 6. To eliminate age from the relationships, we separately computed the correlation coefficients for the 7th and the 11th grades. In this way, we can also compare the relationships of inductive reasoning at these two ages.

It was a consistent result of the previous research that inductive reasoning highly correlated with academic achievements. Pellegrino and Glaser (1982) reported high correlation between inductive reasoning tasks and vocabulary, reading and language tests. Now we may ask, what type of achievements described in this paper (school knowledge vs. applied knowledge) is more closely related to inductive reasoning? As we expected, there were high correlations between inductive reasoning and applied science knowledge at both ages and there were also significant (although weaker) correlations between inductive reasoning and school grades. Especially in the older age group, inductive reasoning proved to be in a

TABLE 6
Correlations between the Test Results and School Performance

	<i>7/8th grade</i>			<i>11/12th grade</i>		
	<i>Inductive</i>	<i>Raven</i>	<i>Grades</i>	<i>Inductive</i>	<i>Raven</i>	<i>Grades</i>
Raven	.60			.51		
School grades	.45	.31		.36	.21	
Applied science	.49	.36	.37	.46	.36	.22

closer relationship with the results of the applied science test than did school grades. Taking into account that the science application test was administered one year after the inductive reasoning tests and the school grades data were collected, these correlations mean that those students who performed better in the inductive tests tended to perform better in the applied science test a year later. This was also the case with school grades, but the relationship was stronger with inductive reasoning. In this sense, we may say that inductive reasoning skills measured at the 7th (11th) grade level were better predictors of applicable knowledge at the 8th (12th) grade level than were school grades.

To shape these observations, we developed a regression model with the applied science test as dependent variable. Regression analyses were carried out separately in each of the two examined ages in the same way. The independent variables were simultaneously entered in one single step. The magnitude of the multiple regression coefficients were similar for the two age groups ($R^2 = 0.27$ and $R^2 = 0.24$) and approximately one-quarter of the variance of the applied science test was explained by these three independent variables. *F*-tests indicate that the relationship between the entire set of the independent variables and the dependent variable was highly significant ($P < .00001$) at both ages. As for the *beta coefficients*, only that of the Raven for the younger group and school grades for the older group were not significant at least at $P < .05$ level.

The main results of regression analyses are depicted in Fig. 4; the proportions of variance of dependent variable explained by the independent variables are indicated in percentages. (Beta coefficients multiplied by the correlation coefficients and by 100.)

At both ages, approximately 17% of the variance of applicable knowledge can be attributed to inductive reasoning. This is not a very high proportion in itself, but it is much higher than that of the other variables. The proportion of variance explained by “school knowledge” was still significant at the age of 13 years, but by the end of secondary education it became marginal. Several other regression models were also computed, including other variables. The percentages were somewhat different, but the influence of inductive reasoning on the use of science knowledge was usually around twice as high as that of school knowledge. (For example, when we did not separate the two age groups, the ratio of variance of applied science test explained by school grades was 13%, and 21% was explained by inductive reasoning.)

It is also worth observing that the Raven test explained far less variance than did inductive reasoning. This finding is consistent with the expectation that specific and concrete mechanisms of thinking are more important in transferring knowledge than is the more abstract general intelligence. It is not obvious in this case, however, why the Raven test explained less variance if it also consisted of inductive reasoning tasks. We explain this difference by

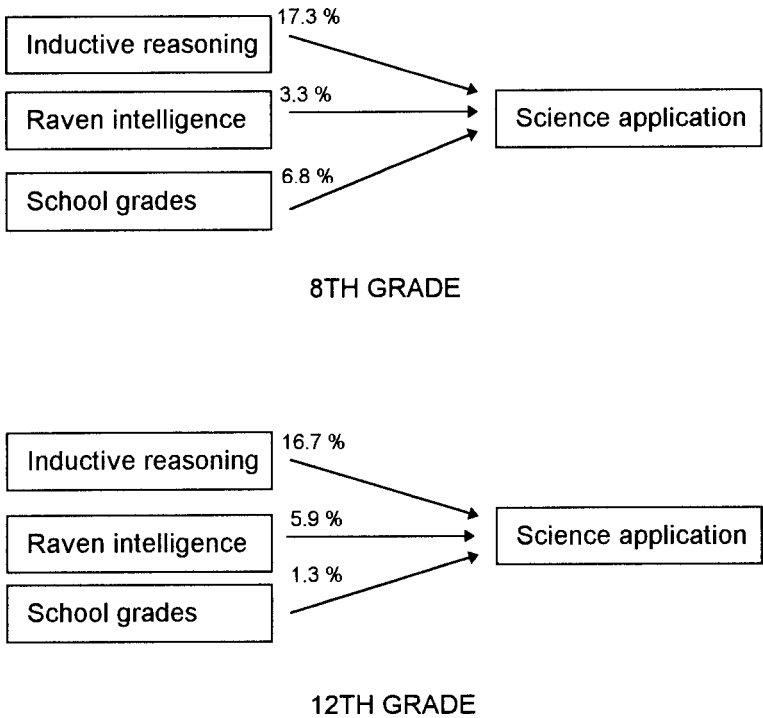


FIG. 4. Regression model of the effect of inductive reasoning on the application of science knowledge.

the different operations and content of the Raven test. The more “content-rich” inductive reasoning tests consisted of tasks with content (words, numbers, letters) and operations (e.g. analogies) familiar in knowledge transfer, whereas Raven’s progressive matrices are deliberately as far from familiar contexts as possible. A comparison of the two charts in Fig. 4 also indicates that school knowledge and more concrete thinking skills are more significant for younger children, whereas abstract thinking skills are somewhat more important for the older students.

GENERAL CONCLUSIONS

The results support the proposition that inductive reasoning develops during a broad age range and that the developmental process covers the whole period of elementary and secondary education. These results are also consistent with studies that report successful training experiments at younger and older ages (Klauer, 1990). Nevertheless, the most rapid development took place during the elementary school years, so that may be the most sensitive period. If we take into account the nature of inductive

reasoning as a means of further learning, the greatest effects could be expected from programmes that aim at younger age groups.

As far as it can be handled by means of statistics, we have found that inductive reasoning is essential to the use of knowledge. In education, we think it more fruitful to use the more easily described concept of inductive reasoning (or its specific mechanisms, such as analogies) than abstract concepts, such as intelligence. Studying the more readily identifiable specific mechanisms may result in more substantial findings for instructional practice, although the overall concept of inductive reasoning may be useful for broad statistical analyses.

As for the generalisability of the results, we have no reason to question that in other cultures we would also find that inductive reasoning develops during a long period of life, and that it is significant in applying knowledge in new contexts. However, if we take into account that inductive reasoning is not only a means of learning but also a product of learning, we may conclude that different educational environments may result in different developmental trends. Certain school curricula and educational practices may result in accelerated development of inductive reasoning, and others may not provide students with sufficient materials to train their higher-order thinking skills.

As for Hungarian students, we may find indicators that seem to support the proposition that the educational system may train their thinking skills effectively, as well as indicators that contradict this proposition. On the one hand, the mathematics and science achievement of Hungarian students is high when compared to other countries, especially around the age of 14 years (see Keeves, 1992), leading to a conclusion that the learning that resulted in these achievements trains their reasoning skills as well. On the other hand, an analysis of school curricula and observation of teaching practices suggests that Hungarian students are closer to the kind of learners Gardner (1991) refers to as “disciplinary”.

Our results have also uncovered some problems for further research. For example, further elaboration of the concepts of school knowledge and applicable knowledge, especially finding more variables to represent them, may allow us to develop more sophisticated models for the role of inductive reasoning. Some of the results of this study cannot be generalised into other educational contexts, but cross-cultural surveys may reveal general trends and allow us to identify the effects of specific cultural contexts, especially the effects of various educational systems, teaching methods, and school curricula, all of which may moderate development. Cross-national studies of educational achievement (e.g. the International Association for the Evaluation of Educational Achievement studies) have indicated several differences between countries, but they have mostly focused on school knowledge and mastering school disciplines (e.g. mathematics, science).

Schooling has other important outcomes as well, such as fostering higher-order thinking skills, and studying them in other cultures or contexts may not only highlight the conditions that best help their development, but also the nature of development.

Manuscript received May 1995

Revised manuscript received April 1996

REFERENCES

- Csapó, B. (1992). Improving Operational Abilities in Children. In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *Neo-Piagetian theories of cognitive development. Implications and applications for education* (pp. 144–159). London: Routledge.
- Csapó, B. (1993, June). *Development of inductive reasoning in adolescence*. Paper presented at the Eleventh International Congress of World Association for Educational Research, Jerusalem, Israel.
- Csapó, B. (1994). Az induktív gondolkodás fejlődése [Development of inductive reasoning]. *Magyar Pedagógia*, 94, 53–80.
- Egan, G.E., & Greeno, J.G. (1974). Theory of rule induction: Knowledge acquired in concept learning, serial pattern learning and problem solving. In L.W. Gregg (Ed.), *Knowledge and cognition* (pp. 43–104). Potomac, MD: Lawrence Erlbaum Associates Inc.
- Ennis, R.H. (1987). A taxonomy of critical thinking dispositions and abilities. In J.B. Baron, & R.J. Sternberg (Eds.), *Teaching thinking skills* (pp. 9–26). New York: W.H. Freeman.
- Gardner, H. (1991). *The unschooled mind. How children think and how schools should teach*. New York: Basic Books.
- Gelman, S.A., & Markman, E.M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, 58, 1532–1541.
- Gick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1–38.
- Gilhooly, K.J. (1982). *Thinking: Directed, undirected and creative*. London: Academic Press.
- Halford, G.S., & Boulton-Lewis, G.M. (1992). Value and limitations of analogies in teaching mathematics. In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *Neo-Piagetian theories of cognitive development. Implications and applications* (pp. 183–209). London: Routledge.
- Holland, J.H., Holyoak, K.J., Nisbett, R.E., & Thagard, P.R. (1986). *Induction. Processes of inference, learning, and discovery*. Cambridge, MA: MIT Press.
- Holyoak, K.J. (1985). The pragmatics of analogical transfer. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 19, pp. 59–87). Orlando, FL: Academic Press.
- Johnson-Laird, P.N. (1988). A taxonomy of thinking. In R.J. Sternberg & E.E. Smith (Eds.), *The psychology of human thought* (pp. 429–457). Cambridge: Cambridge University Press.
- Keeves, J.P. (1992). *The IEA study of science III. Changes in science education and achievement: 1970 to 1984*. Oxford: Pergamon Press.
- Klauer, K.J. (1989). Teaching for analogical transfer as a means of improving problem-solving, thinking and learning. *Instructional Science*, 18, 179–192.
- Klauer, K.J. (1990). A process theory of inductive reasoning tested by the teaching of domain-specific thinking strategies. *European Journal of Psychology of Education*, 5, 191–206.
- Kuhn, D. (1989). Making cognitive development research relevant to education. In W. Damon (Ed.), *Child development today and tomorrow* (pp. 261–287). San Francisco: Jossey-Bass.

- Lawson, A.E., Karplus, R., & Adi, H. (1978). The acquisition of propositional logic and formal operational schemata during the secondary school years. *Journal of Research in Science Teaching*, *15*, 465–478.
- Markman, E.M. (1989). *Categorization and naming in children. Problems of Induction*. Cambridge, MA: MIT Press.
- Nunes, T., Schliemann, A.D., & Carraher, D.W. (1993). *Street mathematics and school mathematics*. Cambridge: Cambridge University Press.
- Pellegrino, J.W., & Glaser, R. (1982). Analyzing aptitudes for learning: inductive reasoning. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 2, pp. 269–345). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Phye, G.D. (1989). Schemata training and transfer of an intellectual skill. *Journal of Educational Psychology*, *81*, 347–352.
- Phye, G.D. (1990). Inductive problem solving. Schema inducement and memory-based transfer. *Journal of Educational Psychology*, *82*, 826–831.
- Popper, K.R. (1972). *Objective knowledge. An evolutionary approach*. Oxford: Clarendon.
- Resing, W.C.M. (1993). Measuring inductive reasoning skills: The construction of a learning potential test. In J.H.M. Hammers, K. Sijstma, & A.J.J.M. Ruijsenaars (Eds.), *Learning potential assessment. theoretical, methodological and practical issues* (pp. 219–242). Amsterdam: Swets & Zeitlinger.
- Ropo, E. (1987). Skills for learning: A review of studies on inductive reasoning. *Scandinavian Journal of Educational Research*, *31*, 31–39.
- Shy, S. (1988). Inductive and deductive reasoning: A structural reanalysis of ability tests. *Journal of Applied Psychology*, *73*, 308–311.
- Simon, H. (1974). Problem solving and rule induction: a unified view. In L.W. Gregg (Ed.), *Knowledge and cognition* (pp. 105–128). Potomac, MA: Lawrence Erlbaum Associates Inc.
- Sternberg, R.J. (1977). *Intelligence, information processing and analogical reasoning: The component analysis of human ability*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Sternberg, R.J. (1986). Toward a unified theory of human reasoning. *Intelligence*, *10*, 281–314.
- Sternberg, R.J., & Gardner, M. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General*, *112*, 80–116.
- Sternberg, R.J., & Wagner, R.K. (Eds.) (1994). *Mind in context. Interactionist perspectives of human intelligence*. Cambridge: Cambridge University Press.
- Tissink, J., Hamers, J.H.M., & van Luit, J.E.H. (1993). Learning potential tests with domain-general and domain-specific tasks. In J.H.M. Hammers, K. Sijstma, & A.J.J.M. Ruijsenaars (Eds.), *Learning potential assessment. Theoretical, methodological and practical issues* (pp. 243–266). Amsterdam: Swets & Zeitlinger.
- Wason, P.C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*, 273–281.