CAPACITY CONSTRAINTS ON REASONING: DEVELOPMENTAL ASPECTS

by

GUEORGUI STEFANOV PACHEV

B.A., University of Sofia, 1986

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Psychology)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

March 1996

Abstract

The development of children's reasoning has often been associated with the attainment of strategies, higher-order concepts, or a broader knowledge-base. Several researchers have argued that these achievements are parallelled and influenced by a child's capacity to process a fixed amount of information in the context of a reasoning task. It is assumed that children's processing resources increase with age and that this increase allows children to acquire the more complex forms of reasoning, advanced strategies, etc. Insufficient resources, on the other hand, prevent children from performing at a higher level.

The present study had two goals: (1) to provide evidence for the increase of resources with age, and (2) to explore the effects asociated with the limited-capacity of the processes involved in reasoning. Two dual-task paradigms were combined in the pursuit of these goals. One paradigm allowed for identifying the level of a task, at which performance was capacity-limited. Age-groups with different amount of resources were expected to exhibit capacity-limited performance at different task-levels. The second approach, based on introducing additional processing load, allowed for comparing the effects of charging the capacity-limits of different processes involved in reasoning. Eighty-six children from three age-groups were given a matrices-completion task at four levels of difficulty. The task was performed either alone or concurrently with a second task. The secondary task was administered in the beginning or in the middle of some trials, thus disrupting processes at the initial stage and processes at the executive part of the solution.

Capacity-limited performance was detected at the third and fourth level of the task for the first two age-groups, respectively. There was an indication that the oldest subjects would exhibit capacity-limited performance at levels beyond the fourth one. Reasoning performance deteriorated when the secondary task disrupted the operation of the executive processes. The additional load introduced in the beginning of a trial, however, did not affect the level of reasoning performance or resulted in improvement in certain conditions. The results were interpreted as supporting the hypothesis that the capacity factor operates as a necessary but not a sufficient condition for the development of reasoning.

Table of Contents

Abstract ii

List of Tables vi

List of Figures viii

Chapter 1: Introduction 1

The Problem 1

The Concept of Capacity: Information-Processing Perspective 5

Chapter 2: Review 23

Neo-Piagetian Models of the Relation between Reasoning and Capacity 23

The Relation between Reasoning and Capacity in Neo-Piagetian theory: Evidence and Problems 33

Chapter 3: Hypotheses and Method 51

Objectives and Hypotheses 51

Method 55

Participants 55

Equipment 56

Tasks 57

Measures and Design 60

Procedure 71

Chapter 4: Results 74

Association between Concurrent Performance on the Secondary Task and Primary Task Performance 74

Group 1 76

Group 2 91

Group 3 101

Effects of Distraction on Reasoning Performance 108

Group 1 109

Group 2 116

Group 3 120

Chapter 5: Discussion 123

Age-Differences in Capacity 123

Effects of Capacity Limits on Reasoning Performance 129

References 134

List of Tables

1. Summary of the empirical evidence 50

2. Trials for the one-attribute level of the reasoning task: conditions, attributes, shapes, and correct responses 66

3. Trials for the two-attribute level of the reasoning task: conditions, attributes, shapes, and correct responses 67

4. Trials for the three-attribute level of the reasoning task: conditions, attributes, shapes, and correct responses 68

5. Trials for the four-attribute level of the reasoning task: conditions, attributes, shapes, and correct responses 69

6. Capacity-limited performance tests for Group 1: perception condition 77

7. Capacity-limited performance tests for Group 1: memory condition 78

8. Group 1: Means and standard deviations for the two measures of primary task performance (no-distraction trials) across task conditions and levels of task difficulty 84

9. Group 1: Means and standard deviations for the secondary task performance measures 86

10. Capacity-limited performance tests for Group 2: perception condition 92

11. Capacity-limited performance tests for Group 2: memory condition 93

12. Group 2: Means and standard deviations for the two measures of primary task performance (no-distraction trials) across task conditions and levels of task difficulty 96

13. Group 2: Means and standard deviations for the secondary task performance measures 99

14. Group 3: Means and standard deviations for the two measures of primary task performance (no-distraction trials) across task conditions and levels of task difficulty 103

15. Group 3: Means and standard deviations for the secondary task performance measures 105

List of Figures

1. Hypothetical performance-resource function 13

2. Sharing of resources between two concurrent tasks 15

3. Unobservable parameters, measures, and their interrelations according to Hunt and Lansman's (1982) model 19

4. An example of a trial for the "perception" condition 61

5. An example of a trial for the "memory" condition 62

6. Group 1: Secondary task performance by task condition, level of difficulty, and distraction condition. 87

7. Group 2: Secondary task performance by task condition, level of difficulty, and distraction condition. 100

8. Group 3: Secondary task performance by task condition, level of difficulty, and distraction condition. 106

9. Hypothetical data pattern 110

10. Group 1: Primary task performance (percentage correct) by performance measure, task condition, level of difficulty, and distraction condition. 112

11. Group 2: Primary task performance (percentage correct) by performance measure, task condition, level of difficulty, and distraction condition. 117

12. Group 3: Primary task performance (percentage correct) by performance measure, task condition, level of difficulty, and distraction condition. 122

Chapter 1

Introduction

The Problem

The purpose of this work is to explore the question of how children's performance on a reasoning task at different ages is influenced by limits on their processing capacity. In brief, it is hypothesized that: (1) there is an age-related increase in the quantitative characteristics of different processes involved in the solution of a reasoning problem, and (2) reasoning performance is constrained by the capacity characteristics of these processes in that a lack of sufficient resources would prevent individuals from performing at a higher level. If these hypotheses are confirmed, the results will add to the description of a developmental factor that parallels and influences the attainment of more effective strategies, higher order concepts, or more complex forms of reasoning with development. This, in turn, can help the explanation of a number of developmental phenomena.

To illustrate, an analogy can be made with adults' performance on a six-term transitive reasoning problem. Adults will usually fail unless there is some way of visualizing the elements and their relations. This failure will not be due to an inability in making transitive inferences; it is the number of elements or the number of intermittent steps to be carried out that make this task difficult. By analogy, one may suppose that a five-year-old experiences similar difficulties when confronted with a three-term transitive task. The child may be able to carry out the necessary comparisons between the elements of each pair but the number of required comparisons may exceed the child's processing resources.

This explanation seems obvious and consistent with the common-sense. Its conceptualization and empirical test, however, have proved to be difficult. First, capacity is not the only factor that determines performance. Several common findings from the area of reasoning development suggest that there are important qualitative differences in the way younger and older subjects approach and solve a reasoning problem. For example, young subjects often exhibit systematic patterns of mistakes; older subjects (i.e., subjects with presumably sufficient processing resources) sometimes fail to solve a reasoning problem correctly; the influence on the solution process of domain-specific knowledge and experience with the type of task is also well documented. Thus, one source of difficulty is associated with the need to distinguish between the effects of capacity limitations and the effects of other factors on performance. Second, a number of unresolved and controversial questions arise with respect to the notion of capacity: what is capacity; how is it quantified; how is it measured? In the context of the example above, one may ask whether the inferior performance of younger subjects on the three-term task and older subjects on the six-term task is due to inability to apprehend all necessary components of the task, to a deficient memory for the premises of the task, or to a failure to coordinate all premises into a logical inference.

Several attempts to answer these and similar questions have been made. In the cognitive developmental literature these attempts are most often associated with the so called "neo-Piagetian" tradition. The authors from this tradition have either proposed alternatives to Piagetian models of development by including additional developmental factors, or have tried to extend Piaget's theory to areas not covered by empirical research from a Piagetian perspective. These efforts of the neo-Piagetian theorists are aimed at providing a more accurate and detailed account of children's cognitive development.

The assumption that an increase in individuals' capacity underlies age-related regularities in cognitive development has been accepted as central by several theorists (Case, 1985; Halford, 1982, 1993; Pascual-Leone, 1970, 1984). On the basis of this assumption they have addressed important issues and offered explanations for a number of developmental phenomena. For example, Pascual-Leone and Sparkman (1980) have argued for the advantages of Pascual-Leone's theory in explaining the transitions between Piagetian stages, the phenomenon of horizontal decalage, and the effects of a task's information-processing load on performance. Horizontal decalage is one of the phenomena considered in Chapman's (1987) structural-functional model. By assuming maturational capacity constraints on cognitive development, Case (1985) offered an explanation for a number of facts: the failure of training studies to produce stage advancement in some children; the considerable cross-task parallels in intellectual development; the similarities in the rate in which cognitive and physical development decelerate; the relatively universal character of cognitive development up to the ages of 16 - 18. To account for these problems and phenomena, neo-Piagetian theorists have proposed several lines of evidence for the role of capacity as a necessary but not a sufficient condition for cognitive development.

The general goal of the present work is consistent with this line of research. The proposed study, however, differs from the "mainstream" research in two important aspects.

First, although the general hypothesis of neo-Piagetian theories that capacity acts as a necessary but not sufficient condition for the development of reasoning is preserved, the approach to the test of the hypothesis is new. So far, the main body of empirical evidence comes from correlational studies, or studies where performance on tasks with established capacity demands has been predicted by subjects' performance on tasks designed to measure capacity directly. Although the latter approach is more reliable than the former, both types of evidence rely heavily on task analytical procedures and "direct" measures of capacity with questionable validity. An exception is provided by several studies by G. Halford and his colleagues (Halford, 1993; Halford, Maybery & Bain, 1986; Maybery, Bain & Halford, 1986). A dual-task procedure for detecting capacity-limited performance instead of direct measures of capacity have been used in these studies. This alternative approach is followed in the present work but it is applied to subjects from different age groups. It is argued that if there are age-related changes in the amount of processing resources, then capacity-limited performance will be detected at the easier versions of the task for younger subjects and at the more demanding task versions for older subjects.

Second, it is assumed that reasoning, as a complex activity, is constrained by the capacity limits of different processes that comprise a particular task. So far, the studies have concentrated on a single process, according to the model of reasoning activity adopted by the particular author. The present study is aimed at exploring the capacity characteristics of several component processes in the context of a reasoning task and their relative influence on performance through the use of a dual-task methodology.

In what follows, the concept of capacity and capacity limitations as used in the adult cognitive literature is considered first. Two approaches to capacity, which the present work builds upon, are briefly reviewed and issues related to measurement of capacity are discussed. The next chapter reviews four contemporary developmental models of the relation between reasoning and capacity and summarizes the empirical evidence generated by these models in support of the neo-Piagetian hypothesis that capacity is a necessary but not a sufficient condition for the development of reasoning. The assumptions that underlie the present study, the particular hypotheses and the description of the method are outlined in Chapter 4. The final two chapters contain the results of the study and their discussion in terms of the stated objectives and hypotheses.

The Concept of Capacity: Information-Processing Perspective

The idea that there are quantitative limits to the human ability of processing information receives special attention within the information processing perspective, where the view of the cognitive system as a channel through which information flows and is transformed makes the description of the channel's limitations an important study task. Despite this importance, there is no commonly accepted approach to the conceptualization and study of capacity. Two approaches that are relevant to the goals of the present study will be described briefly below.

The first one depicts capacity as a characteristic of a mechanism. A particular cognitive function is presented as carried out by a finite collection of elements, each performing a well defined operation on the information. The result is additive, failures of a particular element influence the overall result in a specific way, with the magnitude of the influence depending upon the role of the element in the mechanism. Explanation of performance is in terms of the qualities of the participating structures.

Capacity, in this view, characterizes the elements (and the overall mechanism) in terms of the quantity of specific, *concrete work* that they perform. Thus, the capacity of a store is the number of items it can hold, the capacity of a filter is quantified as the number of items it lets through, the capacity of a process is characterized by the number of items that can be manipulated for a certain amount of time. This approach to capacity will be referred to here as "specific" capacity view.

The important problem in approaching the relation between reasoning and capacity from this point of view is the conceptualization of the "workspace" of reasoning and the particular processes involved in the reasoning activity. Two models of the "workspace" have prevailed in the field: Atkinson and Shiffrin's (1968, 1971) short-term memory store model and Baddeley and Hitch's (1974) working memory model.

R. Shiffrin (1976) reviewed the capacity limitations as revealed by studies in several areas of memory research and concluded that all of them (except for the results on masking) "can be traced to a relatively small set of limitations in a single system: the active memory system, called short-term store (STS)" (Shiffrin, 1976, p.213). The review is based on a version of a model proposed by Atkinson and Shiffrin (1968, 1971). According to the model the memory system consists of sensory input channels and two memory structures, an active but temporary structure called short-term store (STS) and a permanent repository called long-term store (LTS). The STS is considered to be the activated subset of LTS. Sensory information enters the system and is encoded in a series of stages. The process of encoding is the activation of the inactive features contained in the LTS through their contact with the sensory input and, in fact, this is the process of constructing the temporary STS. The loss of information from STS is assumed to be equivalent to the reversion of a currently active feature to a stable inactive mode in LTS. The cause of such transformation is interference by an activity in STS that prevents the maintenance of information in an active state through rehearsal or other control processes.

Thus, the limitation with most far-reaching consequences is the rate of loss of information from STS. The same constraint on the quantity of information in active state applies to the stages after selection when certain information is designated for retention or for use in the processes of controlled search or decision making. For example, the evidence for automatic processing implies that with practice the process of search can be speeded up, processing accuracy can be increased, as can be the number of items selected in the search (see Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). In addition, there are several ways to expand the limits of the controlled processing (e. g., through categorizing and unitizing, as discussed by Shiffrin, 1976). However, no matter how fast and accurate the search process and how many items identified in the search, the loss-rate limitation always stays and restricts the output. Therefore, the capacity characteristics of the participating processes will influence the overall capacity of the system only within the range set by the capacity for storage. In this case the short-term memory span would be a relatively accurate measure for the capacity of the system. This notion of capacity and the way of its measurement are relevant to several problems in memory research. The question is whether it would be relevant to the idea that the short-term store is at the same time the workspace of reasoning.

A. Baddeley and G. Hitch (1974) presented evidence against the view of the STS as the workspace of reasoning and proposed the concept of working memory as their alternative for a system that serves the complex cognitive tasks of language comprehension, learning and reasoning. The argument against the STS model can be summarized in two points. First, if the short-term store acted as a working memory necessary for the performance on a reasoning task, one would expect patients with a grossly defective short-term store to show many other cognitive problems. In fact, such patients often seem to encounter very few practical problems in coping with the information-processing demands of everyday life. Second, if the capacity of the system determined the number of items that can be held in an active state, then the maintenance of a memory load of a number of items while performing a reasoning task would impair the solution. The experimental test of this proposition yielded similar results for reasoning, comprehension and learning tasks: with the increase of concurrent load, performance declined but the degree of disruption was far less than predicted. The disruption affected mainly the time for performance, while the error rate remained more or less constant. These results indicate that memory load does interfere, implying some overlap of processing with the reasoning task, but even loading subjects' memory to capacity still leaves them able to reason accurately. In addition, the pattern of disruption by the concurrent task suggests that storage capacity is not the main capacity constraint on reasoning. What is necessary is a more detailed view of the processes involved in performing a reasoning task and the capacity limits of these processes.

In the model of working memory proposed by Baddeley and Hitch (1974) and elaborated in later works (for reviews, see Baddeley, 1983, 1986, 1992), storage and control processes are (at least theoretically) separated and viewed as different subsystems. In brief, it is hypothesized that the core of the system is a central executive responsible for coordinating the information from the subsidiary systems. The central executive is assumed to function like a limited capacity attention system capable of selecting and operating control processes and strategies. Unfortunately, it has proven the most difficult both to analyze and conceptualize. Thus, the research has concentrated on the subsidiary slave systems with the hope being the gradual whittling down of the functions that need to be assigned to the central processor (Baddeley, 1983, p.315).

The articulatory loop is one of the stores explored in more detail. It stores speech coded information and makes use of a subvocal rehearsal system. The loop is considered to comprise two components -- a phonological store that can hold acoustic or speech based information for 1 to 2 seconds, and an articulatory process, analogous to inner speech. It was established that memory span for words is inversely related to spoken duration of the words. Subjects can generally remember about as many words as they can say in about 2 seconds (the time for which the traces fade away in the passive phonological store). Thus, the model provides an explanation of the tendency for the digit span of children to increase with age: as children get older, they are able to rehearse faster (Hitch & Halliday, 1983).

The second system explored in some detail is the visuo-spatial scratch-pad. Its functions are considered to be maintenance and manipulation of visuo-spatial images. The scratch-pad seems to comprise, in analogy to the articulatory loop, a store linked with a rehearsal process -- in this case the one used voluntarily to control eye movements. Hitch and Halliday (1983) report interesting findings of age related changes in the use of the two systems in memory tasks. Older children tend to use the articulatory loop to remember picture names as indicated by the phonemic similarity effect and the disruption by articulatory interference. Younger children's performance is not affected by either of these factors but is sensitive to visual similarity.

In general, an advantage of the working memory model, as compared to the short-term store model discussed previously, seems to be the potential to separate analytically the different functional parts of the system and thus, to provide a more detailed account for the system's operation. However, the important task of providing a way for estimating the processing capacity involved in reasoning is not as yet solved. The potential of the working memory model in this direction seems to be associated with the analysis of the central executive, which still seems to be the "area of residual ignorance" (Baddeley, 1983, p.315).

In summary, both models under this view represent performance as strictly dependent upon the specific structures that participate in carrying out the solution of a reasoning problem. In order to estimate the impact of capacity constraints on reasoning performance one should first specify an explicit model of reasoning activity. The model should specify the participating processes and a consistent mode of operation should be assumed under all conditions. Two general empirical approaches can be applied on the basis of this model. One can correlate reasoning performance with measures of the capacity characteristics of the participating processes in an attempt to construct a model that explains in full performance on the reasoning task. The other approach consists in disrupting the operation of the separate processes by offering a specific concurrent task (e.g., additional memory load in order to disrupt the operation of the short-term store) in a dual-task situation. Both approaches have been used but the results indicate that the assumption of a consistent mode of operation does not always hold and that subjects are quite flexible in overcoming the disruption of the secondary task.

The second approach to capacity, discussed below, is an attempt to account for this flexibility by including an additional factor that determines performance, namely, the amount of effort (or capacity) invested in the solution of a task. The approach is best exemplified by the resource theories of attention (e.g., Kahneman, 1973; Norman & Bobrow, 1975; Navon & Gopher, 1979). These theories treat capacity limits as limits in the amount of processing, that is, mental work, that can be devoted to a task. A basic assumption of this perspective is that there is a general limit on people's capacity to perform mental work. Thus, the inability to perform a task with excessive capacity demands or two tasks at once may not derive from a structural bottleneck at any particular stage of processing, but rather from a non-specific depletion of a limited pool of resources. Put figuratively, according to the former view the limits of capacity are imposed by the "walls" of the channel; according to the latter, it is the processing efficiency between these walls that further determines the permeability of the channel.

Another assumption shared by capacity theories is that the level of performance depends upon the amount of resources allocated to the task. In Kahneman's (1973) theory, for example, the efficiency of processing within the limits imposed by structural constraints is controlled by four factors: (1) enduring dispositions, assumed to reflect the rules of involuntary attention; (2) momentary intentions; (3) evaluation of demands; (4) arousal level.

Examples of situations, in which results are better when one is more concentrated, are numerous. Performance failures, from this point of view, can be due not only to structural factors like unavailability of appropriate strategies or lack of the necessary operations, but also to a depletion of individuals' resource pools. Figure 1 presents graphically the relation between resources and performance on a hypothetical task. The sector from the origin to point A shows that the quality of performance is directly dependent upon the amount of resources applied. Norman and Bobrow (1975) proposed the term "resource-limited" to designate this type of performance. The sector between points A and B depicts what is known as "data-limited" performance. That is, no matter how many resources are allocated to the solution of a problem, there is a point beyond which performance will not become any better. An example is the situation of listening to a radio when the signal is masked by static. At a certain point the signal-to-noise rate of the radio transmission is such that despite increased efforts invested in understanding the message, it does not become intelligible.

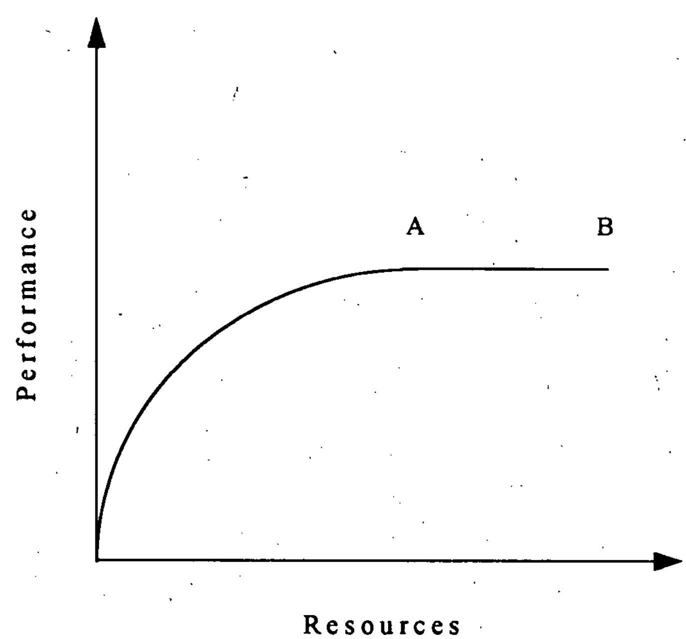


Figure 1. Hypothetical performance-resource function

This second assumption has been further specified with respect to the dual-task situations. Capacity theories assume that when two activities are to be carried out concurrently, the available resources are divided between them. In cases where the joint demands of the two tasks exceed the available resource pool, performance on one or both is at lower levels, or, is capacity-limited. Such a situation is presented graphically by plotting performance on one task against performance on a second, concurrent task on Figure 2. The sectors AB and CD reflect data-limits on Task 1 and Task 2 respectively. The region from B to C depicts the capacity trade-off between the two tasks which in this hypothetical case results in deterioration of performance on both. This kind of plot is known as "performance operating characteristic" (POC) (Norman & Bobrow, 1975).

Comparing the performance plots on Figure 1 and Figure 2, it should be noted that the theoretically assumed, unobservable resource variable on the abscissa in Figure 1 is substituted with an observable performance variable in Figure 2. That is, in a dual-task situation, the quality of performance on one task may serve as an index for the resource expenditure on the other task.

Due to this possibility of indexing resources with performance on another task, the dual task paradigm has been accepted as the main method for estimating demands and resources. The rationale behind the application of this procedure requires additional assumptions and specific trade-off arrangements. In particular, the demands of the two tasks together should exceed the available resources. If the tasks were too easy then performance on both would be data-limited and no interference would be observed.

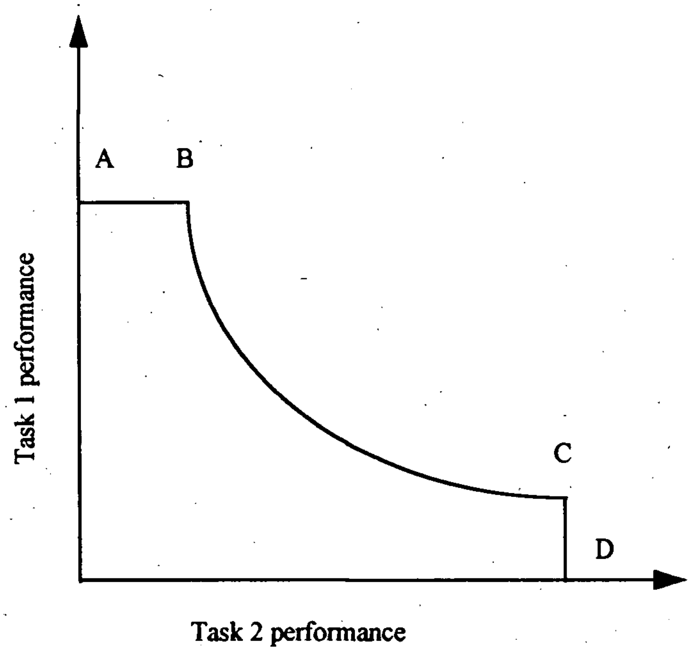


Figure 2. Sharing of resources between two concurrent tasks.

Second, the same amount of resources should be allocated to one of the tasks when it is performed alone and when it is performed concurrently with the second task. This is achieved by manipulating the instructions: one of the tasks is designated primary and subjects are instructed to give it priority in the dual-task situation. When these conditions are met, the quality of performance on the other (secondary) task will depend upon the spare resources, that is resources that remain unused after the necessary amount is allocated to the solution of the primary task. In cases where the secondary task is sensitive to resource variation, the quality of performance on it will allow inferences about the capacity demands of different primary tasks or about the relative amount of resources applied by different individuals to one and the same primary task.

For example, performance on two tasks that have different capacity demands may be indistinguishable if both are within the capacity range of the subjects. The dual task situation allows for distinguishing them. When paired with one and the same secondary task, the task with fewer demands on resources will be accompanied by a better performance on the secondary task.

For the purposes of the present work it is important to evaluate the extent to which the dual-task approach allows us to distinguish between individuals or groups of individuals with respect to their available capacity. Logic similar to that involved in distinguishing between capacity demands of tasks can be applied. Individuals with different resources may perform equally well on an easy task. In a dual-task situation, their performance on the secondary task will differ: the individuals with more available capacity will perform better. Thus, secondary task performance can be used as an index of individuals' capacity.

The problem is how to use this index in dealing with the question of the role of capacity in development. In testing the hypothesis that capacity is a necessary but not sufficient condition for the development of reasoning, the important issue is to show that the amount of available resources sets limits on the level of difficulty at which an individual can perform successfully. In terms of the dual-task approach, individuals with a worse performance on the secondary task in the concurrent-tasks condition should not succeed on the more difficult versions of the primary task. In addition, one should expect a strong association between secondary and hard primary tasks performance, if the difficulty levels of the primary task differed in terms of demands on resources only.

Predicting performance on the basis of secondary task indices of capacity is central for Hunt and Lansman's (1982) formal model of the role of resources in determining individual performance. In brief, it is assumed that individuals differ in three general characteristics that determine performance. These are: (1) structural parameters pertaining to primary task performance; (2) structural parameters pertaining to secondary task performance; and (3) total resource capacity (Hunt and Lansman, 1982, pp. 218-219). Performance on the primary and secondary tasks in both the single and dual task condition is expressed as a function of various combinations of these unobservable variables. Information theory is then used to generate specific predictions concerning the relationships between performance measures.

The idea behind this approach is that if there is a causal relationship between a set of unobservable characteristics and each of a set of observable measures, then knowledge of one observable measure may provide information concerning another. The implementation of the idea in a particular model is depicted in Figure 3. The unobservable variables E 1 (structural parameter pertaining to primary task performance), E 2 (structural parameter pertaining to secondary task performance), and R (total resources), are connected to the observable variables by arrows, whose direction illustrates causation. Performance on the hard primary task is designated as a target variable (i.e., the measure that should be expressed in terms of the other observable variables).

It is clear from Figure 3, that performance on the hard primary task depends on E 1 and R. Therefore, any measure that depends on one or both parameters will provide information concerning these parameters and, thus, improve the prediction of the target variable performance. In particular, the authors demonstrate that performance on the secondary task in the dual task condition provides information concerning performance on the difficult version of the primary task in that performance on both can be expressed as being determined by the same unobservable variables. Statistically, one should expect that a reliable linear relation exists between the two measures if the postulated causal relations are true.

Two other measures that depend on the unobservable variables participating in the prediction and may influence the observed relation, are the performance on the easy primary task and performance on the secondary task, when both are performed alone.

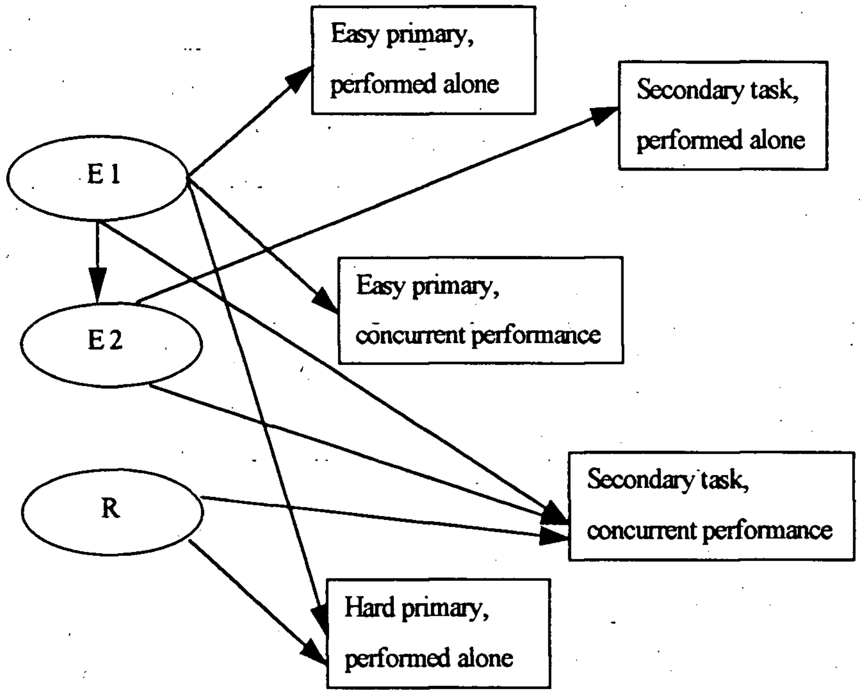


Figure 3. Unobservable parameters, measures and their interrelations according to Hunt and Lansman's (1982) model.

The authors propose that eventual shared variance between these two measures and the performance on the hard version of the primary task is partialled out in the statistical test of the relation.

In summary, the dual-task approach to measuring capacity provides an index of individual capacity (secondary task performance), which can be used for predicting performance on a reasoning task (Hunt and Lansman's "easy-to-hard" paradigm). A success in the statistical prediction of performance on the difficult primary task with performance on the secondary task in the dual-task condition would indicate that the observed differences in primary task performance are due to limited resources. Such a result bears direct relevance to the claim that capacity acts as a necessary but not sufficient condition for the development of reasoning.

The second view of capacity, as exemplified by resource theories of attention, makes no specific architectural predictions, or more precisely, accepts certain architectural assumptions of the alternative approach. Explanation of performance, however, is based not only on the quantitative characteristics of the specified processes, but also in terms of the allocated capacity relative to task demands. That is, the approach allows for taking into account an additional, "intensive" aspect of performance.

The purpose of measuring capacity under this view is to express the result in terms of the quantity of *abstract work* invested in it. As such, capacity can be used as a characteristic of a task (task demands), of performance (allocated capacity), of a structure or a process (processing capacity), or of an individual (available resources). Approaching the problem about the relation between reasoning and capacity from this point of view avoids the need for an *ad hoc* identification of the processes or structures participating in a solution and the need for specifying the units into which capacity of the particular process or structure is quantified. Performance on one and the same secondary task, accompanying the different stages of the solution, can serve as an index of the capacity demands of the compartment processes.

There is, definitely, an overlap of the phenomena that the two uses of the term capacity, described in this section, are intended to capture. It can be argued that these are, in fact, two ways of expressing and describing the capacity factor in performance. The first view describes capacity in terms of specific processing units. The second view depicts capacity in terms of abstract, "pure" quantity and allows for including in the explanation the "intensive" properties of performance. Thus, the two approaches could usefully complement each other.

The concept of capacity, as used in the present work, bears upon both views described above and is an attempt at capitalizing on the advantages of each approach. First, similar to the "specific" approach to capacity, it is assumed that the term refers to a *characteristic* of a process, task or individual. Descriptions of capacity as "mental energy", "mental space", "(resource) pool", etc., can be accepted only as useful but limited metaphors. Second, similar to the "non-specific" capacity view, it is assumed that capacity characterizes the *quantity* of abstract mental work that an individual, applying certain processes, can handle at a time, or the quantity of mental work that is required for carrying out successfully a particular task. Finally, it is assumed that the overall success on a task depends on how much effort is invested in solving the problem. Thus, in a dual-task situation, although the two tasks involve different operations and processing units, the success on each will depend on how much effort the individual can allocate to them.

Chapter 2

Review

Neo-Piagetian Models of the Relation between Reasoning and Capacity

In the literature on cognitive development the study of capacity limitations is associated mainly with the work in what is known as the neo-Piagetian tradition. The goal of this research is to provide a more detailed and accurate account of cognitive development by adopting a more functionalist stance, often borrowing concepts and models from the information-processing approach and the problem solving research. At the same time, neo-Piagetians have more or less tried to preserve basic tenets from Piaget's structural theory of cognitive development.

As stated in the previous chapter, several neo-Piagetian theorists have raised the hypothesis that capacity acts as a necessary but not sufficient condition for the development of reasoning. In brief, insufficient capacity would prevent performance at higher levels of reasoning. In contrast, the more capacity available, the more sophisticated form of reasoning would be exhibited, given that all other conditions are met. Different views, however, characterize the understanding of capacity in neo-Piagetian theories. A brief description of these views will precede the discussion of the evidence in support of the hypothesis.

In Pascual-Leone's theory (1970, 1984, 1987), performance is considered to be determined by the dynamic interplay of two groups of factors. The first group comprises the basic units of the mental apparatus: operative schemes, figurative schemes, and executive schemes. The second includes as factors the most general architectural constraints of the system that is studied. The M-operator (the capacity construct in Pascual-Leone's theory) denotes the nonspecific mental attentional energy or mental space. The function of the M-operator is to boost the mental schemes necessary for performance. Its action is monitored by the executive schemes dominant at the given moment.

Operationally, the power of the M-operator (M-power) is defined as the maximum number of schemes that can be activated in a single mental centration. The power of the M-operator, according to Pascual-Leone, increases endogenously with age. M-power is partitioned into two additive components: *e* -- capacity used by the task executive to represent the problem goal and the initial strategy; and, *k* -- the capacity for activation of additional schemata. The *e*-component is assumed to develop up to the second or third year of life and to remain constant afterwards, while the *k*-component increases to late adolescence. The rate of growth is constant and the range is from one additional scheme at age 3 -- 4, to six or seven additional schemes at approximately the age of 16. These estimates were originally inferred from a logical analysis of Piagetian tasks, and later were empirically verified using a variety of tasks (see Pascual-Leone & Goodman, 1979, for references).

Despite the parallels that can be drawn between this concept of M-capacity and the "nonspecific" capacity notion as described in the previous section, Pascual-Leone's approach to the empirical study of capacity has been quite different from the approach to capacity measurement under the "nonspecific" capacity view. This difference is due to the assumption that M-capacity can be quantified in terms of the number of schemes that are applied to a solution of a problem. Thus, Pascual-Leone proposes an extensive and detailed procedure of task analysis for establishing the capacity demands of a task. Available capacity is estimated by means of tasks considered to be direct measures of individuals' capacity. These estimates of individuals' M-power and the assessment of task demands through task analysis have been used to provide evidence for the relation between reasoning and capacity.

Task analysis has been most often identified as the weakness of the theory (Gelman & Baillargeon, 1983; a similar argument is leveled by Case, 1985). As mentioned above, Pascual-Leone has proposed a detailed and elaborate procedure for task analysis. Nevertheless, there is some arbitrariness in determining the units of analysis because the task analytical procedure is not derived from the theory in a rigorous manner (see also Chapman, 1987).

Another problem is the prediction of performance with measures of M-power. The M-operator is not the only activating factor according to the theory. Other operators can serve as scheme-boosters. For example, the interactions of the operators for learning (L-, and C-operators) with the M-operator are considered to determine both the scheme that is formed and the speed of its formation (see deRibaupierre & Pascual-Leone, 1979). The direct influence on activation is the I-operator, which serves to inhibit the irrelevant to the current task structures (see Pascual-Leone 1987, 1991). The joint action of all these factors produces a dynamic synthesis of schemes in the field of activation. However, since performance is a function of the joint action of several factors, why is it predicted with M-power measures? Averaging the performance level across several tasks of the same demand has been proposed as a safeguard against such bias. This is supposed to reduce the variance caused by factors other than M-power in predicting performance on cognitive tasks. Nevertheless, there is no guarantee that the reduction is sufficient.

The other three theories adopt an approach to the definition of capacity that is closer to the "specific" view of capacity. That is, specific assumptions are made regarding the processes and structures involved in the particular cognitive activity and capacity characterizes quantitatively the operation of these processes or structures.

Executive Processing Space is the capacity construct in Case's (1985) theory, and is defined as the "maximum number of independent schemes that a child can activate at any one time" (p. 289). It is further subdivided into "operating space" (allocated to the activation of new schemes) and a "short-term storage space" (the proportion of the total processing space devoted to maintenance and retrieval). This subdivision does not imply two different capacities, each with their own limit, but a single capacity that can be allocated to the two functions.

Case (1985) provided data from several studies concerning the nature of the growth of short-term storage space with age. It was hypothesized that the total processing space within each period remains constant, while the increase of storage reflects the decrease of operational space. In short, the greater the operational efficiency, the more space available for storage. Supporting evidence for this hypothesis is the high correlation between speed and span as revealed in a series of studies with each age group (see Case, 1985, pp. 354-365). In addition, a study of adults whose counting rate was artificially reduced to the level of six-year old children by counting in a nonsense language, showed the same relation. The span performance of these adults was comparable to that of six-year olds.

For the explanation of these changes in capacity, Case raises a tentative hypothesis for the possible physiological correlate of the changes in efficiency. In short, increased processing efficiency is related to the established fact of progressive myelinization of nerve fibres with development. The myelinization is related to the increased speed of linear transmission and reduced amount of lateral transmission.

One advantage of Case's approach to capacity, as compared to that of Pascual-Leone, is the attempt at specifying in more detail the different processes involved in solving a problem and the quantitative limitations associated with them. Reasoning activity is described as involving two types of processes: the processes of storage and maintenance of information, and the processes of active organization and manipulation of this information for producing the solution. The assessment of capacity in this case, however, is more or less the assessment of the capacity of the short-term storage space, which is characteristic for a particular stage of development. In face of the evidence for the relative independence and different processing areas of the executive and storage processes (Baddeley & Hitch, 1974), and for the general increase of processing speed with age (Kail, 1988), such a treatment of capacity, at least when the question is about the capacity constraints on reasoning, seems insufficient. A more detailed account of the processes involved in reasoning is necessary as is an approach to measuring the capacity limits of these processes.

Another important aspect, which is missing in Case's treatment of reasoning, has been pointed out by D. Kuhn (1983) in a critique of an earlier version of the theory. According to Kuhn, Case analyzes subjects' ability to execute the sequence of steps that lead to a success on a task. What is missing is the more developmentally challenging aspect of knowing that these are the appropriate strategies to apply, i.e., the process of active construction of the problem situation by the subject (see Kuhn, 1983, pp. 94 -99).

The capacity constructs proposed in Chapman's (1987) model and in Halford's (1982, 1993) theory are both examples of the "specific" approach to capacity in that they characterize the limitations of particular processes. Both differ from Case's treatment of capacity in that they attempt to concentrate on the quantitative aspects of processes that are specific to reasoning, rather than on the capacity limitations of the overall system involved in a performance of a task. The conceptualization of these processes, however, is quite different in the two models.

Chapman (1987) based his model on a constructivist approach to reasoning and, following Piaget, described the formal properties of the age-specific forms of reasoning as rooted in the interiorization of action and the coordination of mental operations in operatory structures. Three types of schemes are the functional units of the model: representational, procedural and operational schemes. "Representational schemes" refer to sensory, perceptual and cognitive representations of the permanent and simultaneous properties of comparable objects or classes. They provide the content that is coordinated in the process of solving a task by the operational and procedural schemes. "Procedural schemes" are defined as "transformations effected by the child successively in time and in pursuit of a goal" (Chapman, 1987). Finally, in defining "operational schemes", Chapman emphasized the importance of understanding them as internalized reversible actions. Reversibility here means that the operation is integrated in a structure with other operations that compensate the transformation it designates. In this case, the application of the operation is a simultaneous coordination of implications and does not involve temporal sequence. This simultaneous coordination marks deeper understanding and reaching the conclusion by necessity.

The model is aimed at investigating the "form of reasoning", which is defined as the type of inferential relation uniting children's judgments (conclusions) with the explanations (premises) of those inferences. The operation of the model involves simultaneous coordination of the values provided by representational schemes in an "inferential scheme" (Chapman, 1987). In a later version of the model, this process is referred to as assigning a value to an operatory variable (operatory variables are defined as the "aspects or dimensions of the task situation that the subject recognizes as potentially varying within the experiential context of the task" (Chapman & Lindenberger, 1989). Regardless of the difference in terminology, a basic assumption of the model is that the structural act of assigning a value to an operatory variable corresponds to the functional consumption of a fixed amount of attentional capacity. This fixed amount of capacity is considered a "unit". In terms of the operation of the model then, the capacity requirements of a given form of reasoning will be equal to the number of operatory variables that are assigned values simultaneously in employing that form of reasoning in a particular task. The concept of capacity demands can be defined as the number of representational schemes that must be coordinated for the solution of a given task. Defined this way, the model provides guidelines for the analysis of the capacity demands of the tasks, for the estimation of the capacity necessary for a particular form of reasoning, and for the clarification of the requirements towards the tasks used as independent measures of individuals' capacity.

In task analysis, the estimation of task demands has usefully complemented the analysis of the structural aspects based on Piaget's operatory logic. The advantage of this approach is that it allows for analyzing the tasks in terms of their formal properties and for deriving the capacity demands of those tasks from the quantitative dimensions of those properties. Task analysis of typical Piagetian tasks indicated that demands increase regularly by stage (see Chapman, 1987, pp. 310-311).

With respect to measuring capacity, Chapman did not propose new measurement tasks but used task analysis to demonstrate the relative validity of measures proposed by Pascual-Leone and Case. For example, as a result of the comparison between forward- and backward digit span as measures of capacity, the latter was evaluated as more relevant to the task of measuring capacity involved in reasoning because it entails the assimilation of items to a reversible scheme of temporal order, i.e., a scheme in which the forward order simultaneously implies its inverse.

Halford's treatment of reasoning is based on a complexity metric derived from category theory (MacLane, 1972), which allows for a unified approach to the assessment of the complexity of a task and the structural complexity of the representations and structures used in a solution. The purpose of this is to anchor the complexity classification of the tasks and the levels of reasoning on objective criteria. As Halford put it (Halford, 1982, p. 360), the fact that transitive reasoning has the structure of the binary operations characterizing Level 2 reasoning is a mathematical truth and not an intuitive judgment.

The core of the metric is the process of structural mapping, which is defined as the rule for assigning elements of one structure to elements of another in such a way that any functions, relations or transformations between elements of the first structure correspond to functions, relations and transformations in the second structure (see Halford, 1993, p. 71). Four levels of task complexity are described, which differ in the number of elements and relations determining the problem space.

According to Halford, the level at which a task will be approached depends upon the complexity level of the representations that the subject is capable of processing. The different levels of structure mapping require relevant means of representation, that is, concepts at the respective level of structural complexity. The complexity of concepts, according to Halford, is determined by their dimensionality, or by the number of independent units of information required to define a concept. The units themselves may have arbitrary informational size. Their number is related to the number of arguments in a predicate. Thus, one-dimensional concepts are predicates with one argument (e. g., category membership); two-dimensional concepts are defined as predicates with two arguments (binary relations and bivariate functions); etc.

In brief, the notion of limited capacity in Halford's theory (1993) is associated with the limits in capacity for representing structure and is quantified in terms of the independent dimensions that can enter into a representational structure. Reasoning is capacity limited in the sense that the structure mapping processes impose processing demands that depend on the dimensionality of the mapped structures. Thus, tasks that require concepts of high dimensionality to be mapped will impose high processing load that exceeds the capacity of young children.

Halford's concept of capacity can be clearly classified as an example of the "specific" approach to capacity. What distinguishes it from the other such concepts, is the attempt to consider the capacity limitations of processes that are intrinsically related to the activity of reasoning. Even in his earlier (1982) book, where short-term memory is assumed to be the "workspace" for reasoning, Halford emphasizes the "... information-processing load imposed by the requirements of matching the symbol system to the environment system in a consistent way" (p. 361) and takes no account of loads imposed in any other way. In the later (1993) book, an attempt is made at the explication of the concept of representational dimensionality on the basis of a computational model with connectionist architecture. In terms of the model, the independent dimensions are represented by separate vectors in the processing space. The tensor product of these vectors characterizes the concept that is applied. The argument is based on the established properties of neural networks that discriminability of items is proportional to the number of units used for the representation and the conduction speed of the neural computation. Thus, keeping the two factors constant, the increase in the number of items (vectors) entered into the computation will decrease their discriminability. Although still speculative, this hypothesis offers a way for accounting for the effects of capacity limitations and for eventual age changes in these limitations. In brief, it is suggested that the number of dimensions that can be processed in parallel increases with age, possibly due to differentiation of representation. This would not increase overall capacity but will permit more complex structures to be processed and more complex concepts to be understood. It is clear that the overall amount of information is not limited, as far as the separate vectors that enter into a computation are of arbitrary informational size. The changes are in the number of such independent vectors that enter into any one computation or decision. Also limited in this way are the orders of interaction and levels of structure that can be represented.

The Relation between Reasoning and Capacity in Neo-Piagetian Theories: Evidence and Problems

The theoretical models described above, although similar in several aspects, offer quite different interpretations of capacity and the influence of capacity limitations on reasoning. The different views should be kept in mind in evaluating the evidence because they determine to a great extent the processes upon which the research efforts of the authors were concentrated.

The evidence provided by neo-Piagetian theorists for the relation between reasoning and capacity can be divided into three groups according to the approach taken for the empirical validation of this relation.

The first group is comprised of evidence from correlational studies. Assuming that capacity has the role of a developmental constraint, it is natural to predict that there will be a close correspondence between levels of reasoning and levels of capacity. Two objections have been leveled to this type of evidence. First, some critics have argued that the observed correspondences can be explained as an artifact of correlated age changes (Brainerd & Reyna, 1989; Howe & Rabinowitz, 1990). Second, several authors (e.g., Halford, 1993; Howe & Rabinowitz, 1990) have noted that the magnitude of the correlations does not exceed that typically found between many pairs of cognitive tasks and there is no method of determining whether the correlations are based on capacity.

Both objections are valid, but only if the correlational evidence were the only evidence presented by neo-Piagetian theorists. Several studies in the frame of Pascual-Leone's theory include data about the association between performance on M-capacity measurement tasks and performance on cognitive tasks, the capacity demands of which have been assessed by means of task analysis. This relation has been established for a number of cognitive tasks and for different ages (for a review, see Chapman, 1981; see also Johnson & Pascual-Leone, 1989; Morra, Moiso & Scopesi, 1988; deRibaupierre & Pascual-Leone, 1979). However, this is not the only evidence and the reported correlational data are used as an initial stage in the studies. A similar argument can be made about several studies reported by Case (1985). He found, for example, that the average scores for both cognitive tasks and capacity scores coincided closely with those predicted for each age (Chapters 6-11). It should be noted, however, that these chapters cover the initial "descriptive stages" of his project.

Chapman (1987, 1990) argued that the mere detection of correspondences between capacity and cognitive development does not directly test the necessity-but-not-sufficiency relation between capacity and reasoning predicted by neo-Piagetian theories.

The former pattern of results might exist even if the latter relation did not occur. Chapman and Lindenberger (1989) proposed a direct test of the specific prediction on the basis of the statistical technique of prediction analysis (Hildebrand, Laing & Rosental, 1977). In brief, this is a technique which allows for testing predictions based on a logical relation between nominal or ordinal variables by partitioning the contingency table into "permitted" and "non-permitted" (error) cells. The comparison between the observations actually found in error cells and the expected errors (determined from the marginal totals) yields the test statistic (DEL) which is a measure of the extent to which the number of observed errors is less than expected by chance. For example, in testing the hypothesis that a certain level of cognitive development is necessary for a corresponding level of moral development one should expect the cell determined by performance on cognitive tasks under the specified level and performance on moral tasks at and over the respective level, to be empty.

In Chapman and Lindenberger's study, a task analysis based on Piaget's operatory logic was used to determine the task demands of typical Piagetian tasks for class inclusion, transitivity, multiplication of classes, and multiplication of relations reasoning. A minimum of three units of capacity was found to be necessary for successful performance on such tasks. Two tasks, Backward Digit Span and Pascual-Leone's Figural Intersection Test, were used as measures of individuals' capacity. The results of 120 first, second, and third grade children on the measurement tasks were used to predict their performance on the reasoning tasks. More specifically, it was expected that the tasks would be solved by children with at least three units of capacity as determined from their performance on the measurement tasks. The prediction was confirmed for all tasks but the class inclusion one, which was solved by nearly all children suggesting that the particular version of the class inclusion task could be solved without operational reasoning.

Another line of research, pursued in Chapman's laboratory, addressed the controversial question about the "true" ages at which particular cognitive competencies develop. This topic is characteristic of neo-Piagetian research. Several attempts had been made to demonstrate through task analysis that children could solve versions of Piagetian tasks at an earlier age, because those versions had lower capacity demands (e.g., Case, 1985, Chapter 11; Halford, 1987; Pascual-Leone & Smith, 1969).

Pachev, McBride, Carpendale and Chapman (1993) applied the technique of prediction analysis and tested particular performance predictions in addition to ordering the versions of the tasks according to their demands through task analysis. The first experiment compared the performance of forty eight children from three age groups (4 - 6, 7 - 8, 9 - 11) on two versions of a transitivity task. One version of the task was assessed as requiring the standard (according to Chapman's task analysis) for transitivity tasks: 3 units of capacity. The second version was designed to permit a solution by means of a functional scheme that requires 2 units only. Subjects' capacity level was assessed by means of the Backward Digit Span and the Opposites Test. One of the predictions tested in the study is of interest here. It was expected that children who have only two or less attentional capacity "units" should be able to pass the task when a functional scheme is applicable, but not in cases when a functional solution is impossible. This predictions was confirmed. For both capacity measures and in both conditions of the transitivity task, the number of children who passed the task with two or less capacity units was much less than the number expected to fall into this category solely by chance under the null hypothesis of independence. In addition, clear evidence for a tendency to infer weight as a function of size was obtained, but only for children who would otherwise have used a nonoperational form of inference.

The second experiment compared performance on a standard and easier version of a class inclusion task. It was estimated that in order to solve the standard version of the class inclusion task the child must simultaneously attend to the supraordinate and subordinate classes. Therefore, three variables must be evaluated and the operation of class addition would require a minimum of three units. The decrease in the capacity demands of the easy version of the task was achieved through omitting the comparison of the subclasses by asking the subjects to compare them before the test question was posed. Sixty-eight children, divided evenly into two age groups (5 - 6, 7 - 8) were given two class inclusion tasks (differing in the materials used and the dimension by which the supraordinate class was labeled) and two capacity measures (Backward Digit Span and Opposites Test). Each class inclusion task was given under one of two conditions: the "prior question" condition (subjects were asked to compare the subclasses before the class inclusion question was posed) and, the "no prior question" condition (standard version). It was hypothesized that children with two or less "units" of capacity should be able to pass the class inclusion task in the "prior question" condition, but not in the "no prior question" condition. This prediction was confirmed when Backward Digit Span scores, and the mean scores from the two capacity measures were used in the prediction. The Opposites Test (Case, 1985) failed to yield a significant prediction.

DeRibaupierre and Pascual-Leone (1979) proposed a binomial test method for testing the necessity-but-not-sufficiency of capacity for the development of reasoning. The model resembles the prediction analysis method of Hildebrand, Laing and Rosental (1977) in that it is based on determining error (here called "critical") cells in the contingency table. The procedure afterward is different. Expected frequencies for each critical cell are computed from the marginal frequencies. These expected frequencies are added and divided by the total number of responses in the contingency table. The result is the expected probability p that a response falls by chance within the critical cells and q of falling in the noncritical cells, with N equal to the total frequency in the table and X equal to the total frequency in these critical cells. Using the binomial tables, one can find the probability, with which the obtained pattern is due to chance alone.

In the study, deRibaupierre and Pascual-Leone applied the binomial-test model to test the prediction for the formal-operational stage that subjects with M-power lower than $e + 6$ would not be able to perform above a certain level on formal-reasoning tasks. The M-power of the subjects (12 and 15 year olds) was assessed by means of the Figural Intersection Test and the Compound Stimuli Visual Information task. Both tests were proposed by Pascual-Leone (1970, 1978; Pascual-Leone & Smith, 1969) as measures of M-capacity as defined in his theory. The performance level on the cognitive tasks was determined by task-analysis of subjects' performance on five formal tasks: versions of Balance Task, Projection of Shadows, Pendulum, Flexibility of Rods (Inhelder & Piaget,

1958) and Control of Variables (Scardamalia, 1977). Several other tests were included in the test battery which controlled for the influence on performance of the other "operators" postulated in Pascual-Leone's theory (e. g., F- and L-operators were controlled by including Witkin's Embedded Figures Test and the Hidden Figures Test). In addition, the order of administering the tasks in four separate sessions, allowed the results on the first two sessions to be treated as "pretest" scores and the results from the second two sessions as "posttest" scores.

The predictions of interest here were confirmed for all but one of the four scores determined by the combination of age group and type (pretest vs. posttest) of score: Group 12 pretest ($p = 0.16$), group 12 posttests ($p = 0.02$), group 15 pretests ($p = 0.03$), group 15 posttests ($p = 0.05$). The predictions based on the total scores were also significant: group 12 total ($p = 0.01$), group 15 total ($p = 0.004$), total pretest ($p = 0.005$), total posttest ($p = 0.001$). It should be noted that these results were achieved by using in the prediction the average score of subjects across tasks and the higher score from the results on the M-power measurement tasks. This type of approach was considered by the authors as consistent with the interpretation of the relation between reasoning and capacity in Pascual-Leone's theory. Having in mind, however, that the problem of direct measurement of capacity is still a controversial issue, and the weaknesses of Pascual-Leone's task-analytical procedure mentioned in the previous section, the results should be regarded with caution.

More recent studies in the framework of Pascual-Leone's theory (e. g., Johnson & Pascual-Leone, 1989; Morra, Moizo & Scopesi, 1988; Stewart & Pascual-Leone, 1992) that deal with the relation between capacity and particular forms of reasoning usually approach the specific hypothesis by means of the binomial test. Stewart and Pascual-Leone (1992), for example, in their study of the relation between M-power and moral reasoning report a probability of $p = 0.04$ that the pattern they found is due to chance. This result was supported by a prediction analysis procedure (DEL = 0.527; Z(DEL) = 4.444; $p < .001$). The problems with measuring capacity, however, are evident in this study as well. This time the average of the scores of each subject on the two measures used (Figural Intersection Test and Compound Stimuli Visual Information test) was included in the statistical tests as a more reliable measure of mental capacity.

The prediction analysis and the binomial-test approaches to testing the relation between reasoning and capacity provide more compelling evidence for the role of capacity as a developmental constraint when compared to the mere registration of associations between measures of capacity and measures of cognitive development. It should be noted, however, that each of the studies described above required certain "adjustments" (averaging across reasoning tasks, averaging across measurement tasks, use of highest results, etc.) for a successful prediction in some cases. This indicates one of the weaknesses of this approach: it relies on direct measures of capacity which have questionable validity. In addition, the approach relies on a undifferentiated view of capacity, and does not allow for a detailed study of the compartment processes of reasoning and their quantitative characteristics. This last task is usually addressed in studies with experimental or quasi-experimental designs.

One set of such studies, reported by Case (1985, pp. 331-348), deals more directly with the question of whether the size of the short-term storage space sets limits on the complexity of the executive control structures assembled. All studies use similar logic: if two groups with different short-term storage space are exposed to the opportunity to learn a new structure, only the group with sufficient short-term storage capacity will benefit from the training. This hypothesis has been confirmed for dimensional tasks. Of interest here is an experimental study with adults. The different span of the storage space has been experimentally induced by providing a different amount of training in counting in an artificial language. When exposed to an opportunity for learning a new structure, which involved the trained operation as a component, the group with induced higher short-term storage space benefited considerably more.

The strength of the evidence from these studies should be evaluated keeping in mind the context of the theory that generated these results. As mentioned in the previous section, the important aspect of approaching the task and constructing it as a problem situation by the subject was not well addressed in this model. A particular algorithm for solving a task can be learned, a strategy for approaching the problem can be learned, as well, but the problem is whether the application of the algorithm or the strategy in this case reflects a developmental achievement. From this point of view, the regularities found in these studies are important but they pertain to the role of capacity as a constraint on learning rather than to capacity as a constraint on development.

Halford's application of a mathematical scheme derived from category theory (MacLane, 1972) to the problem of task analysis has been positively evaluated as adding rigor to the problem of task analysis. The application of the complexity metric to the problems of measurement of capacity and of deriving the task demands, however, has been criticized. More specifically, in his earlier works Halford (1982) had relied on the use of short-term memory span measures as measures of capacity.

In his later works, however, Halford adopts an approach that avoids the measurement of capacity by independent means. Instead of independent measures of capacity, Halford applies methods for detecting capacity-limited performance (Norman and Bobrow, 1977), namely Hunt and Lansman's (1982) easy-to-hard paradigm for assessing individual differences in resources. Another feature of Halford's later works is the use of the dual-task paradigm for identifying the key processes involved in reasoning and describing their specific functions. It is this new approach to the problem of the relation between reasoning and capacity in the aspect of development that is of greatest interest for the present work. Two sets of studies on transitive and class inclusion reasoning using the dual-task methodology will be described briefly below.

Consistent with the proposed understanding of capacity constraints as constraints on the dimensionality of representations, Halford predicted that the main difficulties experienced with transitive tasks would be associated with the integration of premises when the relation between nonadjacent premise components is to be judged. Maybery, Bain and Halford (1986) provide evidence for this in a study using a dual task approach with adults. In their study, subjects were presented with successive displays of the premises and the target relation and had to indicate whether the target was consistent with the premises by pressing different buttons for consistent or inconsistent target relations.

The task had two conditions. In one condition, involving premise integration, the target demanded the establishment of the relation between non-adjacent elements. For example, " J is above T", "N is below T" (premises); "N is above J" (target). In the second, control condition, the target display involved establishing the relation between elements from one of the premises (e. g., premises: "R is above G" and "L is below S"; target: " "L is above S"). Vocal reaction to a tone has been used as a secondary task, with probe reaction time as an index for interference. The tone was administered with each premise and target as well as before and after the presentation of the primary task stimuli. Premise integration was expected to occur with the presentation of the second premise, resulting in longest reaction time to the probe at this phase. The results supported the prediction: reaction time was increased for the probe, which accompanied the presentation of the second premise and in the experimental condition only. Other factors, like processing of negatives or increased problem solution time had no effect on this pattern of results. In addition, the pattern was not changed by the decrease of the solution time over trials.

Halford, Maybery and Bain (1986) report two experiments using the dual-task approach to transitive reasoning with children. The first experiment is an attempt at replication of Baddeley and Hitch's (1974) findings that memory load interfered with reasoning only for a near the limit load and for more difficult problems. Eighteen children, (5 - 6-year-olds) were given two- and three-term transitive problems as primary tasks and had to perform them alone or concurrently with either articulation of a word repeated several times or with a short-term retention and rehearsal of two color pairs. Pilot experiments showed that passive retention did not interfere with reasoning. Both the active rehearsal and the articulation condition (the latter to a lesser degree) increased the time and decreased the accuracy of the solution. These effects were larger for the more difficult three-term task. These results, and results like those reported by Maybery et al. (1986) demonstrate the possibility for applying the dual task-approach to studies with children. The problem with this version of the dual-task paradigm is that the observed interference might be due to either structural or output effects. The second experiment of Halford et al. (1986) provides evidence for the capacity nature of the interference. In this experiment the two- and three-term transitive tasks were used as the easy and hard primary task respectively, and the secondary task was remembering and rehearsing two color pairs. Subjects were 36 children within the age range 3;4 -- 5;9. The easy primary task and the secondary task were performed both alone and together, while the hard primary task was performed alone only. The hypothesis for capacity limited performance was assessed by using Hunt and Lansman's (1982) paradigm. The results confirmed the capacity nature of the interference as evidenced by the successful prediction of performance on the three-term transitive task (criterion) by both accuracy and latency measures of performance on the secondary task, performed concurrently with the easy primary task.

This approach has been extended to class inclusion reasoning. Halford proposed that class inclusion inferences were made by mapping the problem into a pragmatic-reasoning schema (i.e., an induced from experience familiar analogue of the inclusion relation). The class inclusion concept, according to Halford's task analytic scheme, is a ternary relation because it involves relations between a superordinate and two subsets. At least part of the difficulties experienced by children, it was hypothesized, could be explained with the system mapping required by the task.

Halford (1993) described two studies (Leitch, 1989; Halford & Leitch, 1989) in which the structure-mapping hypothesis has been tested on the basis of tasks having the logical structure of class inclusion. These tasks preserve the inclusion hierarchy but avoid difficulties caused by the unusual linguistic form of the inclusion question. The minimum case of an inclusion hierarchy would consist of two elements that have at least one attribute in common and at least one on which they differ. Thus, in the studies described by Halford (1993), children had to choose a pair of toys, which shared a certain dimension but differed in another, from a series of pairs that were of three kinds: inclusive, identical (no different attribute) and disjoint (no common attribute).

In one of the studies, children aged 3 -- 6 years received four such problems that required mapping the set of objects into a schema consisting of one common attribute and two distinct attributes. Children less than 5 years old showed performance at a chance level, which supported the hypothesis that the difficulties of young children with the inclusion schema might be at least partly due to the complexity of the mapping. The second study tested the capacity nature of these difficulties using the easy-to-hard paradigm. Subjects were children aged 3 -- 8. The hard primary task required the recognition of a pair of stimuli that formed a minimal inclusion hierarchy (similar to the pairs from the first study). The easy task required children to recognize whether two stimuli were the same or different along only one dimension. Probe reaction time to a tone was used as an index of performance on the secondary task. Performance on the easy primary task, performed jointly with the secondary task, predicted performance on the inclusion isomorph task, with performance on the two predictors performed separately partialled out. That is, positive evidence of capacity-limited performance was obtained.

Through the use of the dual-task methodology Halford has addressed two of the questions that are of interest in the present work. First, the method of selective distraction has been applied (Study 1 in: Halford, Maybery and Bain, 1986) to the problem for the nature of the capacity constraints associated with reasoning. Additional capacity load had been introduced through the secondary tasks for rehearsal and articulation. The results confirmed the expectations that active processing (rather than passive retention) is more disrupting. Second, the easy-to-hard paradigm was applied to provide evidence for the capacity character of the observed difficulties (Study 2 in: Halford, Maybery & Bain, 1986 and Halford & Leitch, 1989). The two questions, however, had been addressed separately. Thus, there is only indirect evidence that the observed interference at rehearsal and articulation is capacity-based. In addition, the developmental aspect (i.e., whether capacity changes with age) of the specific hypothesis for the relation between reasoning and capacity had not been addressed.

In conclusion, it can be said that the problem for the relation between reasoning and capacity, although shared by neo-Piagetian theories, has been approached from quite different perspectives and with different means. These differences concern both the treatment of the capacity construct and the understanding of the processes involved in reasoning.

The capacity constructs, as defined in the discussed theories range from global nonspecific architectural constraints of the cognitive system to characteristics of particular specifiable structures or processes. It can be argued that these are, in fact, different levels of generality in specifying one and the same reality. There is, however, a well-known trade-off between scope and precision in theory building. From this point of view, theories and models that treat capacity as characterizing particular processes are more promising when the goal is to specify the "bottleneck" of a particular function. Thus, in considering the capacity constraints on reasoning the important task is to define the processes that are central to the reasoning activity.

Such an attempt has been undertaken in Halford's theory and, to a certain degree, in Chapman's structural-functional model. Both models concentrate on processes that operate at the executive phase of the solution of a task and deal with task demands imposed by the formal properties of the reasoning demanded by the task. In both cases the process of solving the problem is presented as an application of an available inferential scheme to the material of the task. Capacity characterizes the "read off" process, or the quantitative features of the inferential scheme. Although important, and probably sufficient for explaining the solution in certain cases, these processes are only a part of the reasoning activity. In novel situations, for example, the inferential scheme has to be constructed and it can be argued that it is the success of this constructive process that determines the exhibited higher form of reasoning. From this point of view, the "bottleneck" is situated at the input phase of the solution and characterizes the active construction of the problem space.

Both Chapman's model and Halford's theory allow for interpreting the quantitative dimensions at this phase in terms of the same processes applied to the executive phase. That is, according to Halford, the ability of integrating the task dimensions in unitary representation would determine the complexity of the mental model that is built. In Chapman's terms, the number of operatory variables that can be assigned values simultaneously determines the level of understanding with which the problem would be approached. The influence upon the solution of the capacity limits at the two phases, however, has not been studied empirically as yet, partially due to the difficulties associated with the separation of these phases in the experimental situation.

In summary, as outlined in Table 1, the empirical evidence for the relation between reasoning and capacity postulated in the neo-Piagetian theories can be divided as coming from three sources. The first group comprises correlational studies that provide data for an association between performance on reasoning tasks and on tasks designed to measure capacity. There are two main problems with this type of evidence. The first one is the validity of the measurement tasks. The second is that the relation between performance on the two types of tasks might be explained as an artifact of correlated age changes. In other words, establishing the association does not prove that it is capacity-based and does not test the proposal that capacity acts as a necessary but not a sufficient condition for the development of reasoning.

The second source of evidence comprises studies that use statistical techniques for testing in a specific way the necessity-but-not-sufficiency condition. Chapman and Lindenberger (1989), for example, used the procedure of prediction analysis that allowed for generating and testing specific predictions for the distribution of subjects in the cells of the contingency table according to the type of relation between the measured factors. Although this method has certain advantages in comparison with the correlational analysis, it does not avoid the problem of measuring capacity with "direct" measures and does not allow for a more detailed analysis of the nature of the capacity constraints of the different processes involved in reasoning.

The studies that comprise the third source of evidence for the relation allow for separating the influence on performance of the capacity limits of the compartment processes involved in reasoning through the use of dual-task methodology for an experimental disruption of the solution process at different points. In addition, part of the problems associated with the independent measurement of capacity are avoided through the use of design-based procedures for detecting capacity-limited performance.

So far, the variance in performance attributable to particular compartment processes has been studied separately from the question of whether performance on a particular kind of task is capacity- or data-limited. In addition, by avoiding the problem of measuring individuals' capacity, the experimental approach has lost the means for testing the hypothesis for the necessity-but-not-sufficiency character of the relation by means of prediction analysis.

The study reported below is aimed at providing evidence for the operation of capacity as a necessary but not sufficient condition for the development of reasoning, by combining the advantages and avoiding the shortcomings of the approaches outlined so far.

Table 1
Summary of the empirical evidence.

Type of evidence	Studies	Advantages	Weaknesses
Correlational	Johnson & Pascual-Leone, 1989; Morra, Moiso & Scopesi, 1988; deRibaupierre & Pascual-Leone, 1979; Case, 1985 (Ch. 4-11)	Testing the (weak) prediction: if capacity were a developmental constraint, then there should be an association between levels of capacity and reasoning	1. Based on "direct" measures of capacity. 2. Does not explore causal relation. 3. The magnitude of correlations does not exceed the one that is usually found between cognitive tasks.
Prediction analysis/ Binomial test	Morra, Moiso & Scopesi, 1988; deRibaupierre & Pascual-Leone, 1979 Stewart & Pascual-Leone, 1992; Chapman, 1987; Chapman & Lindenberger, 1989; Pachev et al., 1993.	Test of the specific (strong) hypothesis: Capacity is a necessary but not sufficient condition for the development of reasoning.	1. Based on "direct" capacity measures, which have questionable validity; 2. Does not allow testing for multiple constraints.
Experimental	Case, 1985; Maybery, Bain and Halford 1986; Halford, Maybery and Bain 1986; Leitch, 1989; Halford & Leitch, 1989.	1. Avoids the use of "direct" capacity measures. 2. Allows for testing multiple constraints.	1. Does not test the "specific" (strong) hypothesis directly.

Chapter 3

Hypotheses and Method

Objectives and Hypotheses

The review of the literature identified two sets of problems associated with the hypothesis that capacity is a necessary but not sufficient condition for development of reasoning. The first set concerns the empirical approach to the claim that there is an age-related increase in the capacity characteristics of the processes involved in reasoning. The second set of problems has to do with the nature of the constraints. There is no agreement whether the difficulties associated with solving a reasoning problem are due to limits in some general capacity of the system, to inability for memorizing the premises, or to inability for coordinating all the necessary information in an inferential scheme. The present study was designed to address questions relevant to both problem areas.

With regard to the empirical approach to the relation between reasoning and capacity, it was established that most of the studies relied on procedures involving "direct" measures of capacity and task-analytical schemes with questionable validity. This resulted in rather rough estimates for the influence of capacity limits on reasoning, and the procedures very often involved averaging across performance or capacity measures in order to obtain statistically significant outcomes. In addition, the studies using the alternate approach to measuring capacity by means of a dual-task procedure focused on obtaining capacity-limited performance of subjects from a single group and did not address the questions about age-differences in capacity.

Thus, the first objective pursued with the study was to apply Hunt and Lansman's (1982) dual-task performance model to the question of age-related differences in resources.

As a reminder, the model required at least two levels of difficulty of the primary (reasoning, in the present case) task and yielded an index of capacity-limited performance for the more difficult version. The need to distinguish between different age-groups poses the requirement of more than two difficulty levels of the task. In addition, these levels should differ only in their demands of resources. To meet these requirements, a matrices-completion reasoning task was chosen for the study. The task had the logical structure of addition or multiplication of classes and relations. The level of difficulty was manipulated by increasing the number of attributes that defined the classification criterion. This increase in the number of attributes did not affect the way of presenting the task and the way the task was to be approached.

Assuming that there is an age-related increase in processing resources, a particular prediction can be stated about the pattern of results from the application of Hunt and Lansman's procedure. If age-groups of subjects are presented with versions of a reasoning task, which vary in their quantitative characteristics only, then one should expect to detect capacity-limited performance at different points of the "task scale" for the different groups. In particular, it was expected that younger subjects would exhibit capacity-limited performance at the easier levels of the task.

The second objective that was pursued with this study was to explore the constraints imposed by the capacity characteristics of multiple processes involved in the reasoning activity.

The review of the capacity concepts proposed in the four theories established considerable differences in the understanding of capacity constraints on reasoning. These discrepancies were partially due to the different conceptualizations of the critical processes involved in the reasoning activity. The authors had focused on specifying the limitations of processes associated with maintenance of information in short-term (or working) memory, and processes at the executive part of the solution. Despite these differences, the theories are similar in concentrating on a single process.

The view of capacity constraints proposed in the present work differs from these theories in assuming that reasoning, as a complex activity, will depend on the capacity characteristics of the different processes that comprise it. The exact processes to be included in the solution depends upon the material and way of presenting a task. Thus, it is possible to have tasks equivalent in logical structure that need to be solved by applying different processes. It is also possible to have tasks, differing in logical structure, the solution of which is carried out by the same processes. The latter is the case with the task that was chosen for the present study. The constant format of presenting the task allowed for including both additive and multiplicative classification by varying only the attributes that determined the classification criterion. It was assumed that subjects had to carry out the following processing steps in order to solve each problem: (1) identification of the relevant to the task attributes; (2) construction of a model for the correct answer;

(3) comparison of the model to the proposed items and choice of an item. Thus, performance on the task could depend upon the capacity characteristics of each of these processes.

One method for studying the influence of a particular process on the overall performance, as discussed in the first chapter, is the dual-task method of selective distraction. In brief, the method consists in disrupting the operation of particular processes by administering a specific concurrent task (e.g., additional memory load in order to disrupt the operation of the short-term store). A similar approach was used in the present study. Instead of applying a specific secondary task, however, the compartment processes were disrupted by administering one and the same secondary task at different time-points of a trial. In particular, the administration of the secondary task at the beginning of a trial was supposed to introduce additional load and to disrupt the process of identifying the relevant task attributes; the administration of the secondary task in the second half of a trial was aimed at disrupting the processes at the executive part of the solution.

This modification allowed for combining Hunt and Lansman's procedure with the task of selectively disrupting the operation of the separate processes involved in the solution of the reasoning problem. In addition, the subsequent test of the easy-to-hard prediction allowed for inferences about the capacity characteristics of the disrupted process: if performance on the secondary task predicted performance on the hard reasoning task, then the disrupted process would be, most likely, approaching its capacity limits. The consequences of insufficient capacity could be then estimated by comparing performance on the primary task alone with performance of the primary task in dual-task conditions at the same level of difficulty.

The results of these comparisons bear direct relevance to the second objective, stated above. If the limits in the capacity characteristics of the separate processes did constrain reasoning, then performance of the reasoning task in the dual-task conditions should be worse than performance of the reasoning task alone at the same level of difficulty. These were the expectations about the results from this part of the study.

Method

The measurements in the study combined the requirements of two dual-task paradigms. These are: the easy-to-hard version of the dual-task paradigm (Hunt & Lansman, 1982) for detecting capacity-limited performance, and the selective distraction paradigm for assessing the effect of an interfering task on reasoning performance. The two methods were combined in a single procedure and shared several measures but involved different analyses.

Participants

Participants were recruited from three public elementary schools in Abbotsford School District, British Columbia, and from the "Kids Club" at U.B.C. Child Care Services. Only students for whom a parental consent was obtained were approached.

Overall, eighty-six subjects between 6 and 14 years participated in the study. Three groups were formed with approximately equal number of boys and girls in each group. Group 1 included 31 children from the age range 6 -- 8 (*M* = 83.45 months; *SD* = 5.94). There were 15 girls (*M* = 82.33 months; *SD* = 5.60) and 16 boys (*M* = 84.50 months; *SD* = 6.24) in this group. Group 2 included 28 children from the age range 9 -- 11 (*M* = 117.71 months; *SD* = 9.82), with 16 girls (*M* = 120.63 months; *SD* = 10.66) and 12 boys (*M* = 113.83 months; *SD* = 7.27). Group 3 included 27 subjects from the age range 12 -- 14 (*M* = 149.74 months; *SD* = 6.63). There were 14 girls (*M* = 151.00 months; *SD* = 6.58) and 13 boys (*M* = 148.38 months; *SD* = 6.68) in this group.

The following criteria were taken into account when selecting the age groups for the study: (1) subjects had to be able to solve the easiest versions of the reasoning task; (2) the selection of subjects and groups had to ensure sufficient variance with regard to (the theoretically assumed) available resources within and between groups. With respect to the first criterion, additive and multiplicative classifications are characteristic for the stage of concrete operations and beyond (cf. Inhelder & Piaget, 1964). Thus, the children targeted in the study included those at the stage of concrete operations and formal operations. The second criterion was met by sampling children from the whole range around the ages of 7, 10, and 13.

Equipment

The presentation of stimuli and the recording of most of the responses were under the control of a NEC/ProSpeed/SX 20 portable computer. The experiment was programmed using the MEL 1.0 Integrated system (Schneider, 1990).

Tasks

The *primary task* was a matrices completion classification task. In this type of task, subjects are given a set of elements in a multiplicative table layout, where all spaces but one are filled. Subjects have to identify the criterion (or criteria) according to which the elements are divided into two (or more) subclasses and complete the matrix by filling in the last space.

The logical structure of the task corresponds to addition (in the case of one classification criterion) and multiplication (in the case of more than one classification criteria) of classes and relations. The psychological analysis, however, is complicated by difficulty in discriminating between different ways of solving the problem. Piaget, in his extensive study of classification and seriation in children (Inhelder & Piaget, 1964), points out that several conditions of the operational solution are met by the perceptual configuration of the matrix for the elements already given. Thus, it is possible to find the correct result not through operational reasoning but by extending the graphic properties of the given elements through following the vertical and horizontal symmetries in the matrix arrangement (pp. 153-154). In the present experiment, the attempts to ensure uniform solutions of the tasks included: 1) targeting age groups for which the operational solution is characteristic; 2) manipulation of the way of presenting the elements of the task, so that the influence of perceptual factors is reduced in certain experimental conditions.

For each trial of the present experiment, subjects were first shown on the top left part of the computer screen an incomplete matrix in the form of a square (8 cm x 8 cm) with four cells (4 cm x 4 cm each). The lower right cell was empty and the other three cells contained geometrical shapes that formed a pattern. Subjects were instructed to analyze the pattern and to try to figure out what the contents of the empty cell should be, in order to complete the pattern. The duration of the exposure of the incomplete matrix was set to 5 seconds (Phase 1). Next, four 4 cm x 4 cm squares, approximately 1 cm apart, appeared on the top right part of the screen. These squares were labeled A, B, C, and D, and contained geometrical shapes that represented possible elements for the empty cell of the square from Phase 1. Subjects were instructed to choose the element that corresponded to their solution and to indicate the choice by saying aloud the letter label of the cell (Phase 2). The choice was entered on the keyboard by the experimenter. The display of this second screen was terminated upon receiving a response or after 30 seconds if no response was entered. In fact, none of the subjects exceeded the 30 seconds time-limit of this phase. Subjects' choices, the accuracy of the choices, and their latency were recorded automatically by the computer. Finally, the termination of Phase 2 initiated the display of a third screen that contained the material from Phase 1 and Phase 2. Subjects were asked to explain their choice. The experimenter recorded on paper the number of relevant attributes referred to in the explanation for each trial. This number was later used to calculate a performance index, which was intended to capture eventual difficulties that subjects might have with the identification of the relevant task variables. Subjects were given 30 seconds for a justification of their answer. The next trial started after the 30 seconds had elapsed or after the experimenter entered a command signifying the end of the trial.

This basic structure of a trial appeared in modified versions across the different conditions of the experiment. The level of task difficulty was manipulated by varying the number of attributes defining the classification criterion. Tasks were designed at four levels of difficulty, that is, one-, two-, three-, and four-attribute tasks. The following attributes were used: shape (square, rectangle, circle, ellipse, or triangle), color (black or white), size (small or large), number (the maximum number of shapes was 3), and orientation (only rectangles, ellipses and triangles were combined with the orientation attribute; they were flipped 90 degrees to the right or to the left).

There were two conditions of each trial, differing in the display of the four possible answers at Phase 2. In one of the conditions, referred to as the "Perception" condition here, subjects made their choice in the presence of the incomplete matrix. In the second, "Memory" condition, the four possible solutions were presented alone. This manipulation was aimed at ensuring that subjects would engage in active analysis during Phase 1. Examples of the three events of a trial for the two conditions are given on Figures 4 and 5.

Concurrent performance on the primary task included the trial events described above and a secondary task administered at different points of a primary task trial.

The *secondary task* was a manual reaction (pressing a key) to a tone signal. A series of 1000 Hz tones, of 100 ms duration each, were administered at different points of the task and the subjects had to respond to each by pressing the "Z" key on the keyboard. For all trials that involved an administration of a tone, as well as for the series, which was designed to measure subjects' criterion reaction time, subjects were instructed to keep a left-hand finger on the "Z" key.

Criterion reaction time was measured in a separate series, at the beginning of the experimental procedure. The series consisted of 25 four-second trials. Within each trial the tone was administered at one of four time points: 500, 1000, 1500, or 2000 ms. (There were 7 signals for the 1000 ms tone-onset point and 6 for each of the other tone-onset points). The order of trials was random and controlled by the computer. The secondary task in the concurrent performance conditions of the experiment involved the administration of the tone signal at the beginning of Phase 1 (400 ms after the initial display of the primary task), or at the beginning of Phase 2 (500 ms after the presentation of the second screen), or at the beginning of both screens (500 ms and 500 ms respectively for Phase 1 and Phase 2).

Measures and Design

The purpose of the study was two-fold: 1) to provide evidence for capacity-limited performance at different levels of difficulty of the task for the different age groups, and 2) to provide a test for capacity limitations at different points of the reasoning process, thus allowing inferences about possible multiple quantitative constraints on reasoning performance. These two aims, coordinated in the study, pose certain requirements to the design and to the measures necessary for the tests.

Hunt and Lansman's (1982) procedure for detecting capacity-limited performance requires repeated measures of primary task performance at at least two levels of difficulty,

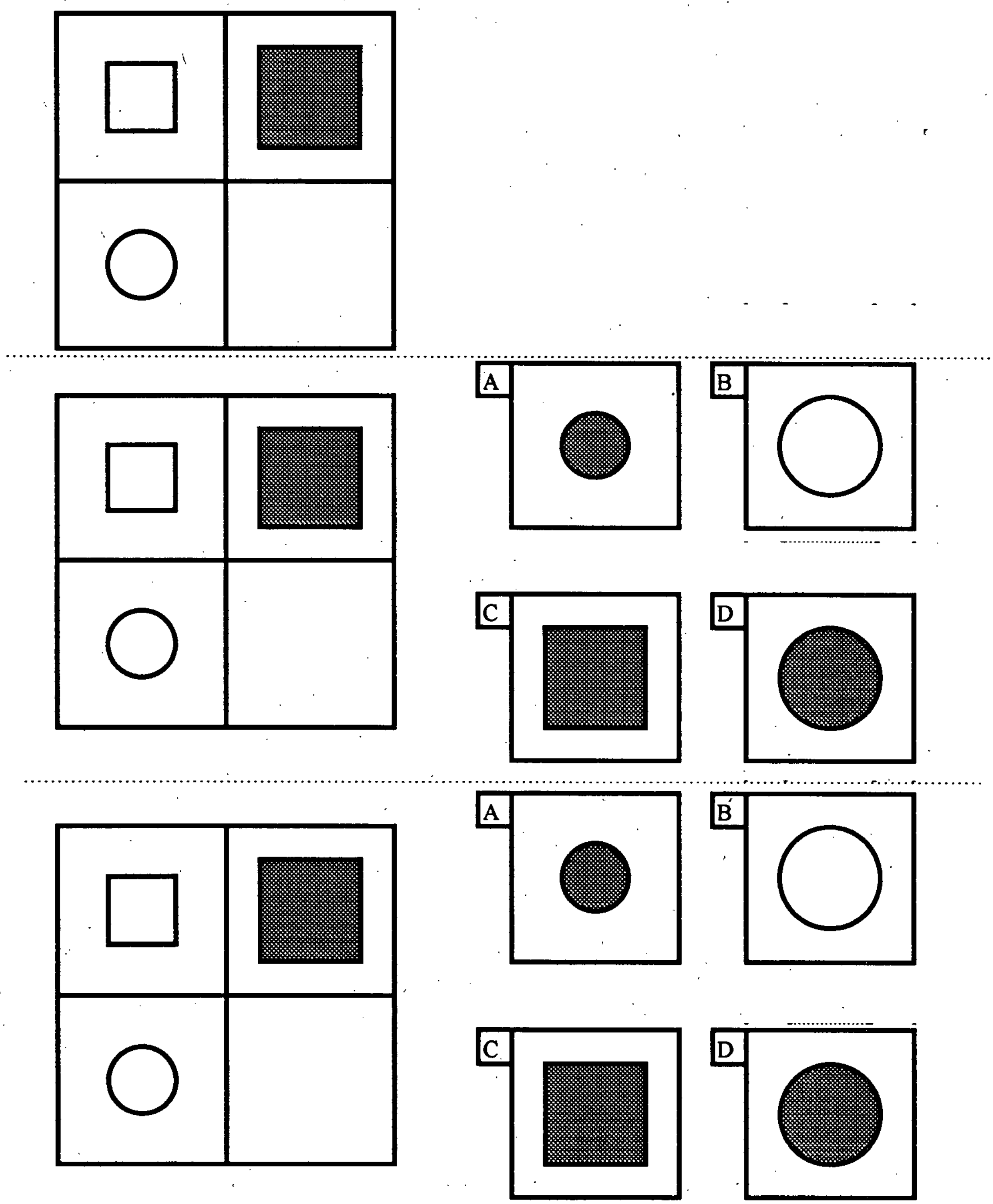


Figure 4. An example of a trial for the "perception" condition, three-attribute task. The attributes that define the classification criterion are: shape (square or circle), color (black or white), and size (smaller vs. larger).

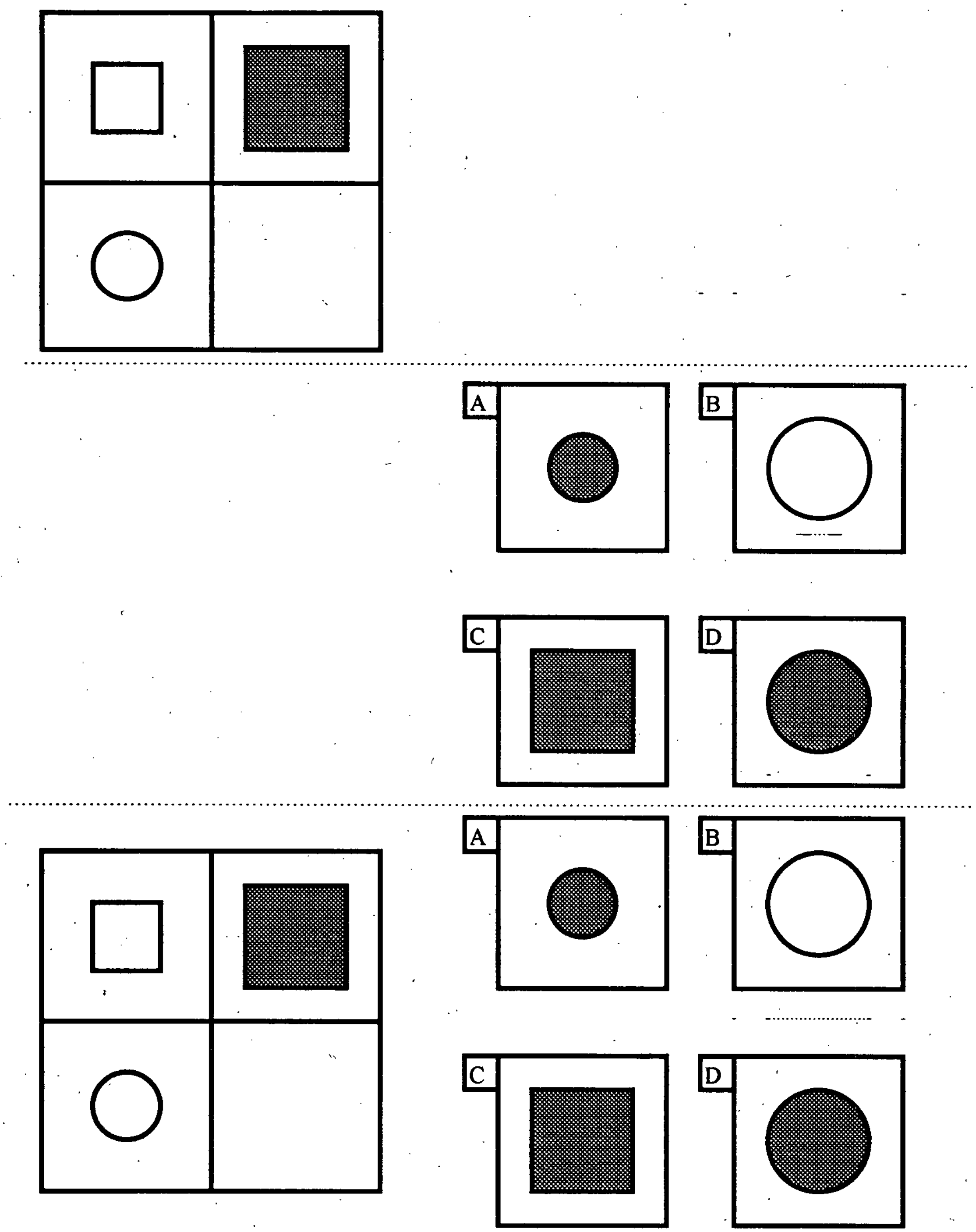


Figure 5. An example of a trial for the "memory" presentation condition, three-attribute task. The attributes defining the classification criterion are: shape (square or circle), color (black or white), and size (smaller vs. larger)

for dual-task conditions and alone. Similarly, secondary task measures include performance alone and in dual-task conditions. The correlational character of the test poses the requirement of a fixed order of experimental events for each subject. Further, the second objective requires that there are several dual-task conditions within each level of difficulty. Finally, both methods are more reliable when there are multiple measurements for each data point.

In the present study, subjects from each age group were given the reasoning (primary) task at three *levels of difficulty*. Pilot testing revealed that the youngest subjects could not solve correctly the four-attribute problem, thus, Group 1 was given the task at one-, two-, and three-attribute level of difficulty. A ceiling effect characterized the performance of subjects from the other two age-groups at the one-attribute level of the task. That is why, Group 2 and Group 3 were administered the task at the two-, three-, and four-attribute level of difficulty.

There were two *task conditions* for the administration of the trials within each level of difficulty: the Perception and Memory conditions, mentioned in the description of the task. The first one resembles more closely the standard format for presenting matrices tests (e.g., Raven's "Progressive Matrices"). The Memory condition was introduced here to force and control for a more uniform way of solving the problem. In brief, the presentation of the four possible answers without the initial incomplete matrix gives subjects no alternative but to form a model of the answer and to try to memorize it at Phase 1. At the same time, if subjects are consistent in their approach to the solution one should expect parallel (though, not necessarily equal) results, as indexed by the secondary and primary task measures of performance, in both conditions.

Each Level of difficulty x Task condition block contained four conditions, designed to meet the requirements for comparing the effects of distraction at different points of the reasoning process. The trials of the "No-distraction" condition presented subjects with the primary (reasoning) task only. The results obtained from these trials provided the measure of performance on the primary task alone, which was necessary for the Hunt and Lansman's test, and also served as a base for estimating the effect of distraction on reasoning performance. The other three conditions contained concurrent performance of the primary and secondary task trials. "Phase 1 distraction" involved administering the tone signal at the beginning of the first screen; in "Phase 2 distraction" trials the tone was administered at the beginning of the second screen; in the "Full distraction" trials it was administered at the beginning of both. Performance indices from these trials provided measures of concurrent performance on the primary and secondary tasks.

To meet the requirement for multiple measurements for each data point, there were three trials for each "level of difficulty" by "task condition" by "distraction condition" cell. Thus, subjects from Group 1 received 24 one-attribute task trials (12 trials in the Perception condition and 12 trials in the Memory condition), 24 two-attribute task trials, and 6 three-attribute task trials (the tasks at the highest level of difficulty were given in the No-distraction condition only). Subjects from Groups 2 and 3 were given 24 two-attribute task trials, 24 three-attribute task trials, and 6 four-attribute task trials.

Since the main interest was in the quantitative characteristics of the solution process, subjects were instructed about the type of problem they would be given. Thus, the presentation of the trials was blocked according to the level of difficulty and the condition of presenting the task (Perception or Memory condition). Within each block, the trials of the No-distraction condition were presented first, and then the nine trials of the three dual-task conditions followed in a random sequence. The order was determined by a random draw before the experiment and was the same for all participants who received these blocks.

Further, to minimize the differences in the impact that the attributes have on the difficulty of the task (for example, shape may be more readily utilized as a classification criterion than orientation), the number of tasks with a particular attribute or a combination of attributes was set to be approximately equal within a block. The same approach was adopted in determining the geometrical shapes for a trial: approximately equal number of squares, rectangles, triangles, circles, and ellipses, appeared within a level of difficulty block. The trials for the four-attribute level of difficulty task were the exception here. The inclusion of the relative orientation as part of the classification criterion restricted the choice of geometrical shapes to triangles, rectangles and ellipses.. Finally, an equal number of A, B, C, or D correct responses appeared within each level of difficulty block. Subjects were explicitly warned not to expect that the correct answer would always be one and the same. The order of attributes, shapes, and responses, as they appeared across the trials of a block was determined by means of a random draw before the experiment and was fixed for all participants that received this block. Tables 2 to 5 show

Table 2
Trials for the one-attribute level of the reasoning task: conditions, attributes, shapes, and correct responses.

Trial number	Task condition block	Distraction condition	Criterial Attribute	Shapes	Correct response
1	"Perception"	No distraction	shape	circle, square	D
2	"Perception"	No distraction	color	triangle	B
3	"Perception"	No distraction	orientation	ellipse	A
4	"Perception"	Phase 2.	number	circle	C
5	"Perception"	Full	orientation	rectangle	A
6	"Perception"	Phase 1.	size	square	B
7	"Perception"	Full	shape	circle, triangle	D
8	"Perception"	Phase 1.	color	rectangle	A
9	"Perception"	Phase 2.	orientation	ellipse	D
10	"Perception"	Phase 2.	size	square	C
11	"Perception"	Phase 1.	number	triangle	C
12	"Perception"	Full	shape	circle, square	B
13	"Memory"	No distraction	size	square	C
14	"Memory"	No distraction	shape	triangle, ellipse	D
15	"Memory"	No distraction	number	circle	C
16	"Memory"	Full	color	square	B
17	"Memory"	Phase 2.	number	ellipse	A
18	"Memory"	Phase 2.	orientation	triangle	C
19	"Memory"	Phase 1.	size	rectangle	A
20	"Memory"	Full	shape	ellipse, triangle	B
21	"Memory"	Phase 2.	color	square	D
22	"Memory"	Phase 1.	number	rectangle	A
23	"Memory"	Full	orientation	ellipse	D
24	"Memory"	Phase 1.	shape	triangle, circle	B

Table 3
Trials for the two-attribute level of the reasoning task: conditions, attributes, shapes, and correct responses.

Trial number	Task condition block	Distraction condition	Criterial Attributes	Shapes	Correct response
25	"Perception"	No distraction	shape, color	square, circle	C
26	"Perception"	No distraction	color, orientation	rectangle	A
27	"Perception"	No distraction	number, color	circle	D
28	"Perception"	Phase 1.	number, size	circle	B
29	"Perception"	Phase 2.	size, orientation	triangle	D
30	"Perception"	Full	number,orientation	ellipse	B
31	"Perception"	Full	size, color	square	C
32	"Perception"	Phase 2.	color, orientation	triangle	B
33	"Perception"	Phase 1.	number, shape	circle, square	A
34	"Perception"	Phase 1.	shape, orientation	rectangle, ellipse	A
35	"Perception"	Full	size, shape	circle, triangle	C
36	"Perception"	Phase 2.	number, size	rectangle	D
37	"Memory"	No distraction	size, color	circle	C
38	"Memory"	No distraction	shape, orientation	triangle, rectangle	B
39	"Memory"	No distraction	number, shape	ellipse, circle	B
40	"Memory"	Phase 2.	color, orientation	rectangle	A
41	"Memory"	Phase 1.	size, shape	square, rectangle	D
42	"Memory"	Full	number, color	triangle	D
43	"Memory"	Full	size, orientation	circle	B
44	"Memory"	Phase 2.	number,orientation	triangle	C
45	"Memory"	Phase 1.	shape, color	square, rectangle	A
46	"Memory"	Full	number, size	ellipse	C
47	"Memory"	Phase 2.	shape,orientation	ellipse, triangle	D
48	"Memory"	Phase 1.	size, shape	square, circle	A

Table 4
Trials for the three-attribute level of the reasoning task: conditions, attributes, shapes, and correct responses.

Trial number	Task condition block	Distraction condition	Criterial Attributes	Shapes	Correct response
49	"Perception"	No distraction	number, shape, color	square, circle	D
50	"Perception"	No distraction	color, orientation, shape	triangle, ellipse	C
51	"Perception"	No distraction	size, color, shape	ellipse, rectangle	B
52	"Perception"	Phase 1.	number, color, size	circle	C
53	"Perception"	Phase 2.	number, color, orientation	rectangle	A
54	"Perception"	Full	color, size, orientation	ellipse	B
55	"Perception"	Full	number, shape, orientation	rectangle, ellipse	C
56	"Perception"	Phase 2.	color, shape, orientation	triangle, ellipse	D
57	"Perception"	Phase 1.	number, shape, size	circle, square	A
58	"Perception"	Phase 1.	shape, size, orientation	rectangle, triangle	B
59	"Perception"	Full	number, shape, size	ellipse, rectangle	D
60	"Perception"	Phase 2.	number, size, orientation	rectangle	A
61	"Memory"	No distraction	size, color, shape	square, circle	A
62	"Memory"	No distraction	number, shape, color	circle, rectangle	D
63	"Memory"	No distraction	number, color, size	triangle	C
64	"Memory"	Phase 2.	color, size, orientation	rectangle	B
65	"Memory"	Phase 1.	color, shape, orientation	rectangle, ellipse	A
66	"Memory"	Full	shape, size, orientation	ellipse, triangle	C
67	"Memory"	Phase 2.	number, size, orientation	rectangle	D
68	"Memory"	Full	number, color, size	square	A
69	"Memory"	Phase 1.	number, shape, orientation	triangle, ellipse	B
70	"Memory"	Full	number, shape, color	circle, triangle	B
71	"Memory"	Phase 2.	number, color, orientation	rectangle, triangle	D
72	"Memory"	Phase 1.	size, color, shape	square, circle	C

Table 5
Trials for the four-attribute level of the reasoning task: conditions, attributes, shapes, and correct responses.

Trial number	Task condition block	Distraction condition	Criterial Attributes	Shapes	Correct response
73	"Perception"	No distraction	number, shape, orientation, color	ellipse, rectangle	B
74	"Perception"	No distraction	number, shape, orientation, size	ellipse, rectangle	A
75	"Perception"	No distraction	orientation, size, shape, color	rectangle, triangle	C
76	"Memory"	No distraction	number, color, orientation, size	triangle	D
77	"Memory"	No distraction	number, shape, orientation, color	ellipse, triangle	B
78	"Memory"	No distraction	orientation, shape, color, size	triangle, rectangle	D

the characteristics of the trials for each level of difficulty, with respect to order, task condition, distraction condition, attribute, geometrical shape, and correct response.

Two dependent measures characterized performance of the reasoning (primary) task. One measure was the *accuracy* of the solution as recorded at Phase 2 of each trial. The measure presented the percentage correct (accurate) of choices from the three trials of each condition cell. Thus, failure to provide a correct answer to any one of the three tasks was assigned a score of 0%; one correct answer out of three was assigned 33%; two correct solutions -- 67%; all three -- 100%.

The second measure, *attribute identification*, was derived from children's justification of their answers. After making their choice at Phase 2, subjects were asked to explain their choice. The explanations were then scored by assigning one point for each attribute that was correctly included in the justification. The sum of the points from the three trials of a particular condition cell was obtained. This sum was then converted into the percentage of the total number of attributes that formed the classification criteria for the trials of the particular condition cell. The different levels of difficulty of the task required the coordination of a different number of attributes for achieving a correct response. Thus, presenting the scores as a percentage of the correctly identified attributes allowed comparisons of performance across the different levels of difficulty.

Both primary task measures were treated as indices of concurrent performance or performance alone according to the conditions of the trials from which they were extracted. Secondary task measures, however, were derived from different tasks.

The measure of performance on the secondary task alone (*criterion reaction time*) was derived from a twenty-five trial series, that was administered as a separate task at the beginning of the procedure. The measure was calculated as the mean reaction time for the trials after the tenth trial of the series. Misses and reactions under 200 ms. were excluded from the calculation.

Concurrent performance of the secondary task was estimated on the basis of the dual-task trials of the three distraction conditions. All measures were calculated as mean of the reaction time of the three trials that formed a condition cell. The mean of the Phase 1 and Phase 2 distraction was used as performance measure in the Full distraction condition, where the secondary task was introduced twice -- at the beginning of both phases. In cases where subjects did not respond to a tone signal, the mean was calculated on the basis of the remaining trials.

Procedure

Subjects were tested individually in one session, approximately half an hour long, in a quiet room. Subjects sat at a table and faced the computer directly. The front end of the computer was set at about 20 cm from the edge of the table, so that there were about 60 cm from the subject's face to the screen, and the wrist of the subject's left hand could rest comfortably on the table when a left-hand finger was set on the "Z" key of the keyboard. The experimenter sat to the subject's right and controlled the events of the experiment through the keyboard with his left hand.

Subjects were briefly explained the purpose and the procedure for the study. The study was presented as being similar to a computer game and aimed at revealing "how well people manage to do several things together"; the procedure was presented as consisting of two simple tasks that the subjects had to perform either separately or together. At the end of this preliminary phase, subjects were reminded that they may discontinue their participation at any moment and were invited to try the tasks.

The trials, designed to provide a measure of the criterion reaction time, were administered first. Subjects were instructed to position their finger on the "Z" key, to listen for the tone signal and to press the key in response, without trying to anticipate the signal. After completing the series, subjects were acquainted with the reasoning task and given training trials. The structure of a trial was explained, using printed copies of four trials, one for each level of difficulty. Subjects from Group 1 were shown examples of one-, two-, and three-attribute tasks, while Group 2 and Group 3 were given examples of tasks from all four levels of difficulty. Subjects were explicitly told what they were supposed to do at each phase of a trial, what questions would be asked, and how they were supposed to answer these questions. Subjects were then asked to solve each one of the four tasks from the appropriate level of difficulty and to justify their choices. When there were mistakes, the tasks were explained again.

Training consisted of eight trials from the one-attribute level of difficulty administered through the computer. The first set of four presented the Perception condition, the second set -- the Memory condition. The duration of the displays for the first trials from both sets was controlled by the experimenter. These trials were used to explain to subjects the task one more time and to illustrate the difference between the Perception and Memory conditions. The other three trials from both sets followed the way of presentation of the test trials. None of the tasks used for explaining the procedure and for training were repeated in the test series.

After the training, subjects were warned that the testing trials were to begin and the No-distraction trials from the Perception condition of the lowest level of difficulty (one-attribute tasks for Group 1, and two-attribute tasks for Groups 2 and 3) were administered first. The nine dual-task trials for the Perception condition at the same level of difficulty followed. Then the two blocks of the Memory condition at the same level of difficulty were administered. This order was followed for the blocks at the other levels of task difficulty. Subjects were explicitly warned what kind of tasks to expect before each block. For the dual-task trials, the importance of solving correctly the reasoning task was emphasized. Subjects were told to respond to the tone "as soon as they heard it", but that the important thing was "to be accurate in finding the correct shape".

All subjects understood the tasks and participated willingly. The pauses between the blocks provided the necessary time for relaxation and rest. The whole procedure took between 30 and 40 minutes per subject.

Chapter 4

Results

Association between Concurrent Performance on the Secondary Task and Primary Task Performance

According to Hunt and Lansman's model, capacity-limited performance on a primary task (and the individual differences in resources, which determine different levels of success on this task) is reflected in the association between performance on the primary task, performed alone, and the secondary task, performed concurrently with an easier version of the primary task. When there is a significant correlation between the levels of performance on the two tasks, this is an index that individuals are performing near-the-limit in the dual task situation. It also means that a decrement in performance on the primary task is due to limits in their available resources. Further, statistical significance should hold even when eventual variance associated with measures of performance on the easy version of the primary task, performed alone, and the secondary task, performed alone, are partialled out of the correlation. Thus, the appropriate index for capacity- limited performance on the primary task will be the partial correlation between performance on the hard version of the task, performed alone, and performance on the secondary task, performed concurrently with an easier version of the primary task, when performance on the easy-primary and secondary tasks (both performed alone), is partialled out.

This model and the statistical test related to it were used here to provide evidence for age related changes in available resources, utilized for the solution of a reasoning problem. It was argued, that if there are age-related differences in capacity then the capacity-limited performance on the primary task would be detected at different levels of task difficulty for the different age-groups. The index of capacity-limited performance was the significant partial correlation between performance on the primary task (alone) and performance on the secondary task (performed together with an easier version of the primary task), with the relevant predictors[1] partialled out.

Two such kinds of tests were performed on the data for each age-group of subjects. The first one explored the relation between two successive levels. For each age-group there were two successive levels of the test: the correlations and partial correlations between the measures at the one-attribute and two-attribute levels, and between the two-attribute and three-attribute levels, for Group 1; the correlations and partial correlations between the measures at the two-attribute and three-attribute levels, and, three-attribute and four-attribute levels, for Groups 2 and 3. The second kind of test seeks to detect the association between two non-consecutive levels of difficulty (non-successive levels tests). One such test was performed on the data for each age-group: the correlations and partial correlations between performance on the two measures at the one-attribute and three-attribute levels of the task for Group 1; the correlations and partial correlations between the measures at the two-attribute and four attribute level for Groups 2 and 3.

[1]One of the predictors is always the secondary task, performed alone. The second predictor varies according to the dependent measure used in the test. Thus, when accuracy is the index of primary task performance, the second relevant predictor is the accuracy score on the easier version of the primary task. When the number of identified attributes is the index of primary task performance, the second predictor is the number of identified attributes on the easier version of the primary task.

These two kinds of tests were applied to the data for each age group and each task condition (Perception vs. Memory) separately. In addition, the relation was explored separately for each dependent measure of primary task performance and for each distraction condition (i. e., Phase 1 distraction, Phase 2 distraction, Full distraction), which yielded a secondary task concurrent performance score.

The statistical analysis was carried out in two steps. First, the correlations between the target variables were obtained and their significance was evaluated according to a more conservative (2-tailed, α=.01) criterion. Second, only relations that satisfied this criterion were explored further. Partial correlations for the variables were obtained, with the relevant predictors partialled out.

Group 1

The results of the procedure for the data from Group 1 are shown on tables 6 and 7 (for the Perception and Memory conditions, respectively). The results for the Perception condition (Table 6), are quite straightforward. First, consider the character of the established relations. All significant correlations (and, in fact, most of those that did not reach the significance criterion) between performance measures on the secondary and primary task are negative. This indicates that greater latencies on the reaction-time task (that is, less spare resources for the execution of the secondary task) are associated with lower performance scores (decrement in performance) on the primary task. This result is consistent with the assumption that the amount of available resources constrains reasoning.

Table 6
Capacity-limited performance tests for Group 1, perception condition.

	Test Measure	Correlations		Partial correlations	
		accuracy	attributes identification	accuracy	attributes identification
Successive levels tests:	Secondary (1-attribute,concurrent) to Primary task (2-attribute., alone)				
Distraction	phase 1.	-.0756	.2478	-	-
	phase 2.	-.1977	.1185	-	-
	full	-.4811**	-.1964	-.4907**	-
Successive levels tests:	Secondary (2-attribute, concurrent) to Primary task (3-attribute, alone)				
Distraction	phase 1.	-.7258**	-.5759**	-.6960**	-.5448**
	phase 2.	-.5459**	-.5581**	-.4231*	-.5443**
	full	-.5875**	-.4413	-.5529**	-
Non-successive levels tests:	Secondary (1-attribute, concurrent) to Primary task (3-attribute, alone)				
Distraction	phase 1.	-.1492	-.0984	-	-
	phase 2.	-.0925	-.0513	-	-
	full	-.2730	-.3663	-	-

Note, that only the correlations that exceeded the criterion value (.4556) were explored further.
$N = 31$
* $p < .05$
** $p < .01$

Table 7
Capacity-limited performance tests for Group 1, memory condition.

Test Measure		Correlations: accuracy	Correlations: attributes identification	Partial correlations: accuracy	Partial correlations: attributes identification
Successive levels tests:		Secondary (1-attribute, concurrent)	to	Primary task (2-attributes, alone)	
Distraction	phase 1.	-.1121	-.0452	-	-
	phase 2.	-.5466**	-.7142**	-.4548*	-.6882**
	full	-.5622**	-.3667	-.5129**	-
Successive levels tests:		Secondary (2-attributes, concurrent)	to	Primary task (3-attributes, alone)	
Distraction	phase 1.	-.4979**	-.5072**	-.3722*	-.3814*
	phase 2.	-.5707**	-.5525**	-.5162**	-.4496*
	full	-.6981**	-.5507**	-.6175**	-.4578*
Non-successive levels tests:		Secondary (1-attribute, concurrent)	to	Primary task (3-attributes, alone)	
Distraction	phase 1.	-.7120**	-.6315**	-.7018**	-.6178**
	phase 2.	-.4708**	-.4606**	-.4870**	-.4437*
	full	-.3869	-.3214	-	-

Note, that only the correlations that exceeded the criterion value (.4556) were explored further.
$N = 31$
* $p < .05$
** $p < .01$

Second, all but one of the correlations between the two primary task measures and the measures of performance on the secondary task in the three dual-task conditions were significant for the test that compared the two-attribute and three-attribute levels of the task. At the same time, the other successive levels test (i.e., the correlation between the measures at the one-attribute and two-attribute levels of the task) failed to yield significant correlations, except for the correlation between performance on the secondary task in the Full distraction condition at the one-attribute level, and primary task performance (accuracy index) at the two-attribute level. The non-successive levels tests failed to detect any significant relations.

It can be concluded from these results that subjects' reasoning performance is capacity-limited at the three-attribute level of this version of the task. The failure to detect significant relations between performance on the other two tests is also a valuable result. In general, such a failure in these kind of tests can be due to a ceiling (or floor) effect on one of the measures, or to the fact that the two measures do not vary together. The latter case, in the context of Hunt and Lansman's model, indicates that the joint demands of the tasks in the dual-task condition do not exceed the available resources. Thus, even if performance on the hard version of the primary task were data limited, the secondary task score would not be sensitive to the limits in resources and the two measures would not correlate. It can be hypothesized that this is the case in the present study. This supposition is supported by the fact that only the secondary task from the Full distraction condition for the one-attribute level correlated with reasoning performance at the two-attribute level: the administration of the tone twice was supposed to increase capacity demands of the secondary task, and the joint demands of the two tasks. The examination of the results on the secondary and the primary tasks can provide further data pertinent to the question of how to interpret the obtained pattern of correlations. The discussion of these results, however, follows the review of the data from the correlational tests.

Finally, the statistical significance of all correlations was retained after the variance associated with measures of performance on the easy version of the primary task, performed alone, and the secondary task, performed alone, were partialled out. The values of the partial correlations are displayed in the two rightmost columns of Table 6.

The results from the tests on the data for the Memory condition (Table 7) are more complicated. The expected negative relation between the variable pairs was maintained. Similar to the Perception condition, all correlations between the target measures reached or exceeded the criterion value for the relation between the two-attribute and three-attribute tasks. Unlike the Perception condition, however, a number of correlations from the other two tests were significant. For the test that related secondary task performance at the one-attribute level to primary task performance at the two attribute level, significant correlations were established between secondary task measures for Phase 2 distraction condition and both measures of primary task performance. Secondary task performance in the Full distraction condition correlated significantly with the accuracy measure of primary task performance. For the tests that related secondary task performance at the one-attribute level to primary task performance at the three-attribute level, secondary task measures from Phase 1 and Phase 2 distraction conditions correlated significantly with both primary task measures. Finally, all correlations that reached the significance criterion at the first step of the analysis, retained their significance after the other two predictors were partialled out.

The significant first order and partial correlations for the Memory condition tests relating the two- and three-attribute levels of the task indicate that for the subjects from Group 1 performance on the three-attribute level is capacity-limited. This claim is supported by the success of the tests relating secondary task performance at the one-attribute level with primary task performance at the three attribute level. In addition, limits in the available resources constrained the quality of performance at the two-attribute task, too, as indicated by the success of several tests relating the one-attribute and two-attribute levels of the task.

The difference in the pattern of results between Perception and Memory condition can be explained by the increased demands of the task to subjects' resources in the Memory condition. One should remember that the manipulation that distinguished the two conditions was the presentation of the stimulus material at Phase 2. The difference was that in the Memory condition the four possible answers were shown to the subjects without the support of the incomplete matrix. Subjects had to complete the analysis and form a model of the answer to the end of Phase 1, to remember, and then to compare their model to the answers from the set. In the Perception condition, by contrast, the availability of the incomplete matrix allowed subjects to avoid some of these tasks and to distribute them more regularly through the two phases. Thus, it is logical to assume that the task in the Memory condition is more difficult and imposes greater demands to subjects' processing resources.

If the task was more demanding in the Memory condition, then one should expect that performance of the primary task would become capacity-limited at an earlier point of the task scale of difficulty than in the Perception condition. Thus, the increased demands of the task explain the detection of capacity-limited performance of the primary task at the two-attribute level of difficulty. This explanation is supported by the fact that the secondary task from the Phase 2 distraction condition yielded the significant correlations for the first successive levels test: it is one of the conditions that should be directly affected by the experimental manipulation. The increased difficulty of the primary task, and the related increased "sensitivity" of the secondary task to limits in resources, can also account for the significance of the non-successive levels tests.

The analyses of primary and secondary task performance with more traditional methods can provide information about the overall reliability of the procedure for detecting capacity-limited performance, as well as information about the particular hypotheses and questions that were posed when considering the pattern of correlations. Thus, the analysis of primary task performance should answer the questions:

Is the expected overall decline of performance across the levels of the task preserved in the particular experimental conditions? Is there an indication of ceiling effects at certain levels of the task? How is the assumed increased difficulty of the task in the Memory condition reflected in the performance of the reasoning task?

The analysis of secondary task performance should answer the following questions:

What is the difference between performing the secondary task alone and in the different distraction conditions? Is the expected increase in reaction time found across the different levels of task difficulty? What is the differential effect of the distraction conditions on the magnitude of response-time latencies? Is there an indication that subjects used different strategies for solving the reasoning task in the Perception and Memory condition? Is the assumed increased difficulty of the reasoning task in the Memory condition reflected in secondary task performance?

Table 8 contains the results from subject's performance on the primary task for both dependent measures. Overall, there was a steady decline of performance with the increase in the number of attributes, defining the classification criterion. The data sets for the two dependent measures were analyzed using repeated-measures ANOVAs in which the within-subjects factors were task condition (Perception vs. Memory) and difficulty level of the task (one-, two-, and three-attribute level of difficulty). For the accuracy measure, only the main effect for difficulty level was significant, $F(2,60) = 100.11$, $MSe = 607.64$, $p < .001$. The differences between the successive levels in both task conditions were evaluated against Dunn-Bonferroni $t'_{\alpha=.01}(4,30) = 3.30$. The performance of each level of the task was significantly worse than the performance of previous: $t(30) = 10.20$, $MSe = 470.04$, and $t(30) = 6.77$, $MSe = 288.79$, for Perception condition; $t(30) = 9.99$, $MSe = 543.51$, and $t(30) = 3.59$, $MSe = 402.04$, for Memory condition. The data for the second dependent measure yielded significant main effects for level of difficulty, $F(2,60) = 96.88$, $MSe = 230.07$, $p < .001$, and task condition, $F(1,30) = 4.83$, $MSe =$

Table 8
Group 1: Means and standard deviations for the two measures of primary task performance (no-distraction trials) across task conditions and levels of task difficulty.

	One-attribute task level	Two-attribute task level	Three-attribute task level
"Perception" condition			
Accuracy (% correct)	82.94 (17.04)	48.42 (27.17)	19.19 (18.70)
Attributes identification (% correct)	92.52 (16.53)	69.84 (16.25)	52.74 (14.04)
" Memory" condition			
Accuracy (% correct)	83.97 (24.09)	41.84 (31.09)	23.55 (23.10)
Attributes identification (% correct)	88.26 (18.27)	66.07 (18.47)	52.65 (16.99)

N = 31

70.72, $p < .05$. Within the two task conditions, the differences between the adjacent levels were also significant: $t(30) = 10.56$, $MSe = 180.62$, and $t(30) = 7.18$, $MSe = 87.95$, for the Perception condition; $t(30) = 8.01$, $MSe = 269.36$, and $t(30) = 3.84$, $MSe = 189.26$, for Memory condition.

Thus, subjects' performance declined across the levels of task difficulty for both conditions and for the two dependent measures. The main effect for task condition in the second ANOVA provides evidence in support of the proposition that the task is more difficult in the Memory condition.

There are two important consequences of the fact that this difference in difficulty was reflected in the results of the attributes identification measure. First, one should expect greater latencies on the concurrent performance of the reaction-time task in the Memory condition as compared to the Perception condition, for the distraction condition where this process is (presumably) located. In the context of the present experiment, this is the Phase 1 distraction condition. Second, the fact that the increased difficulty is not reflected in the accuracy measure indicates that subjects manage to compensate for the difficulties they encounter in identifying the relevant attributes in some cases and to build a precise model of the answer. The first of these propositions will be evaluated when performance of the secondary task is analyzed below; the second proposition will be discussed in the second part of the analyses, dealing with the effect of the distraction conditions on reasoning performance.

Table 9
Group 1: Means and standard deviations of the secondary task performance measure for the different conditions of the experiment.

	Secondary task concurrent performance at the one-attribute level of the primary task		Secondary task concurrent performance at the two-attribute level of the primary task	
"Perception" condition				
Phase 1. distraction	1179	(524.63)	1360	(446.11)
Phase 2. distraction	1287	(625.87)	1364	(491.53)
Full distraction	1536	(574.99)	1578	(541.12)
"Memory" condition				
Phase 1. distraction	1511	(513.11)	1353	(476.66)
Phase 2. distraction	1122	(447.85)	1241	(388.38)
Full distraction	1462	(496.67)	1554	(562.79)
Criterion R. T.		550	(193.59)	

Note. The measures are in milliseconds.
$N = 31$

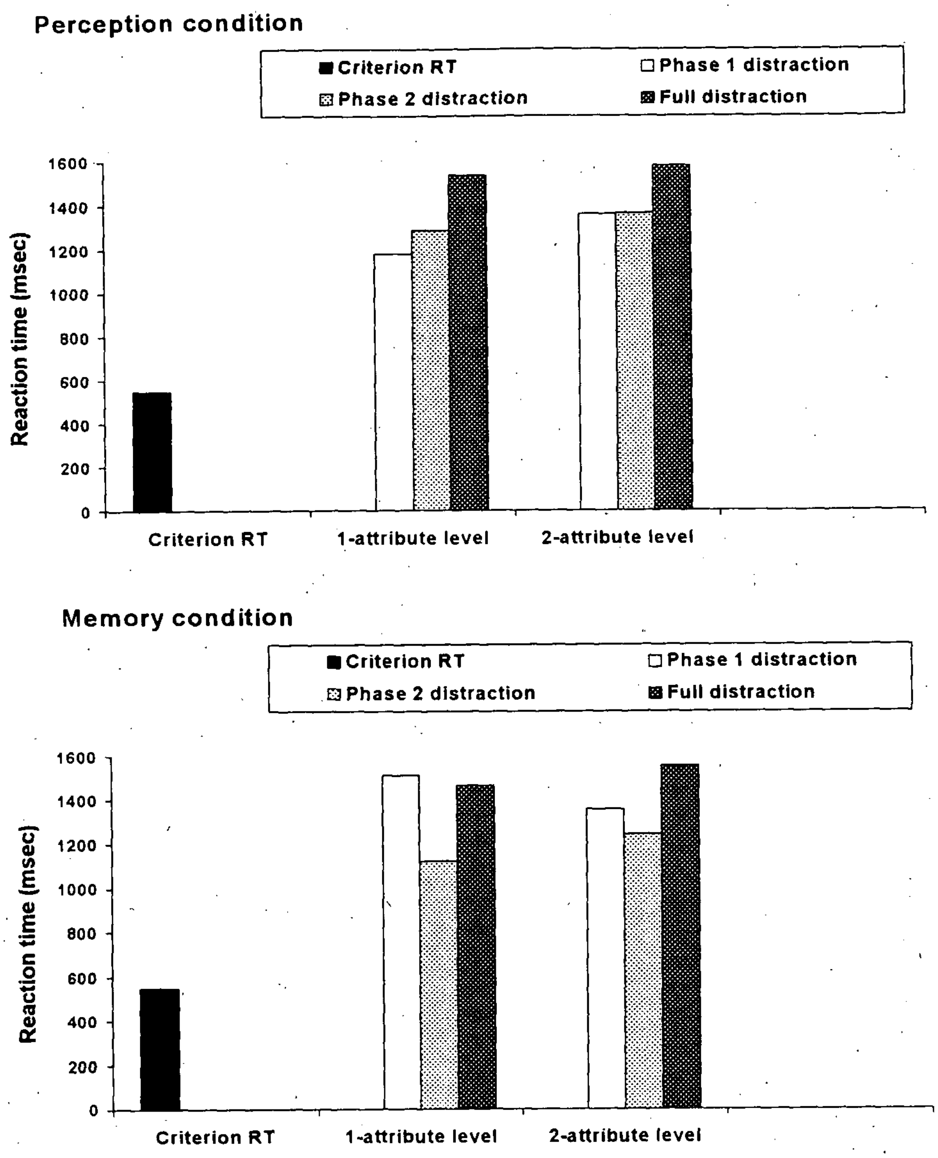


Figure 6. Group 1: Secondary task performance by task condition, level of difficulty, and distraction condition.

The results of secondary task performance are summarized in Table 9. Figure 6 gives a graphical presentation of the results across the different conditions of the experiment.

The two task conditions were first analyzed separately using a repeated-measures analysis of variance procedure in which the within-subjects factors were the six distraction conditions and the control condition for the secondary task. Performance of the secondary task given alone differed significantly from performance of the secondary task in the different distraction conditions, as estimated with Tukey's HSD, $a(7,30) = 335.96$.

A repeated-measures ANOVA with three within-subjects factors (difficulty level, task condition, distraction condition) was used in order to explore the changes in secondary task performance across the different conditions of the experiment. The analysis revealed a main effect for distraction condition, $F(2,60) = 28.85$, $MSe = 86389.16$, $p < .001$, qualified by a significant task condition by distraction condition interaction, $F(2,60) = 7.63$, $MSe = 99625.24$, $p < .001$. In addition, the three way interaction approached significance ($p = .068$). In order to clarify the nature of the differences among the distraction conditions, the data were further analyzed for the Perception and Memory condition separately by means of 2 x 3 (Difficulty level x Distraction condition) repeated-measures ANOVAs and comparisons of Phase 1 distraction and Phase 2 distraction conditions to Full distraction condition.

For the Perception condition, only the simple main effect of distraction was significant, $F(2,60) = 12.72$, $MSe = 113312.96$, $p < .001$. The follow-up analysis by means of Dunn-Bonferroni $t'_{\alpha=.01}(4,30) = 3.30$, revealed a significant difference for the Phase 1 vs. Full distraction contrast at the one-attribute level of difficulty, $t(30) = 4.30$. None of the contrasts at the two-attribute level of the task reached significance.

For the Memory condition, the simple main effect of distraction was significant $F(2,60) = 24.91$, $MSe = 72701.45$, $p < .001$, as was the difficulty level by distraction condition interaction, $F(2,60) = 3.28$, $MSe = 110423.22$, $p < .05$. Unlike the Perception condition, the paired comparisons by means of the Dunn-Bonferroni test at the two levels of the task revealed significant differences between Phase 2 distraction and Full distraction conditions: $t(30) = 4.97$, at the one-attribute level; $t(30) = 4.57$, at the two-attribute level.

These results provide the necessary information for answering the questions posed at the beginning of this part of the analysis. The first question concerned the differences in performance of the secondary task both alone and in dual-task conditions. If concurrent performance of the task is an index of spare resources, then one should expect larger latencies on the secondary task in the distraction conditions as compared to performance of the secondary task alone. In the present experiment, this increase of reaction time in the distraction conditions was consistent throughout all levels of difficulty and task conditions.

A result that further supports the reliability of the secondary task as an index of spare resources is the detection of larger latencies for the secondary task in the Full distraction condition in several cases. This condition involved the repeated administration of the tone signal. According to the model of dual-task performance, the repeated administration of the distraction increased the overall demands of the task to subjects' resources.

The analysis failed to detect the expected increase in reaction time with the increase in the primary task's difficulty. This failure, however, does not invalidate the use of the secondary task as an index of spare capacity. Such a failure is not a rare result in other studies using the dual task methodology (see Lansman & Hunt, 1982, p. 14). In this particular case, the failure to detect differences between levels of difficulty (but note, such differences are detected within a difficulty level) can be explained by practice effects which were due to the requirements for multiple measurements for each data point and for a fixed order of events for each subject.

Finally, the inspection of the Phase 1 and Phase 2 distraction results reveals consistent patterns in the two task conditions. In the Perception condition, the latencies for the Phase 1 distraction responses are of a lesser or nearly equal value to the latencies for the Phase 2 distraction responses. In the Memory condition, the latencies for the Phase 1 distraction responses are larger than the latencies for the Phase 2 distraction responses. This change of the pattern can be interpreted as a change in the strategy that has been used in solving the reasoning problem. Subjects obviously used the display of the incomplete matrix at Phase 2 of the Perception condition to complete the construction of the response model at this stage of the task (or to perceptually match the answers to the graphical pattern of the initial matrix). Consequently, the demands to capacity were lowered at Phase 1 and increased at Phase 2. The presentation of the task in Memory condition precluded the possibility for such a "shift" of the activities. Thus, the demands for resources at Phase 1 increased. This explanation provides further support for the hypothesis raised at the inspection of the results from the correlation tests, that the Memory condition is more demanding in terms of resources than the Perception condition. In addition, it is consistent with the expectation, discussed in the analysis of primary task performance, for longer latencies on the concurrent performance of the reaction-time task at Phase 1 in the Memory condition as compared to Phase 1 distraction in the Perception condition.

Group 2

The results from the application of Hunt and Lansman's procedure on the data from Group 2 are summarized in Table 10 (Perception condition) and Table 11 (Memory condition). The direction of the significant correlations was negative, as expected. For the Perception condition (Table 10), significant relations between secondary task performance and performance of the primary task at a higher level of difficulty were established for the three sets of tests. Secondary task performance at Phase 1 distraction condition of the two-attribute level of difficulty correlated significantly with primary task performance at the three-attribute level (accuracy measure), suggesting that performance of the latter task was capacity-limited. The secondary task at the two-attribute level (Phase 1 and Phase 2 distraction condition) predicted performance on the four-attribute level of the primary task (accuracy measure). One more correlation for the non-successsive levels test was significant but failed to retain significance in the partial correlations tests. For the tests relating the three- and four-attribute levels of the task,

Table 10
Capacity-limited performance tests for Group 2, perception condition.

	Test Measure	Correlations		Partial correlations	
		accuracy	attributes identification	accuracy	attributes identification
Successive levels tests:	Secondary (2-attributes, concurrent) to Primary task (3-attributes, alone)				
Distraction	phase 1.	-.5054**	-.4452	-.5504**	-
	phase 2.	-.3630	-.3722	-	-
	full	-.2436	-.2295	-	-
Successive levels tests:	Secondary (3-attributes, concurrent) to Primary task (4-attributes, alone)				
Distraction	phase 1.	-.4285	-.5755**	-	-.4722*
	phase 2.	-.5974**	-.5676**	-.5146**	-.5755**
	full	-.4389	-.4186	-	-
Non-successive levels tests:	Secondary (2-attributes, concurrent) to Primary task (4-attributes, alone)				
Distraction	phase 1.	-.5371**	-.2576	-.4560*	-
	phase 2.	-.5701**	-.5889**	-.4581*	-.3792
	full	-.0444	-.0222	-	-

Note, that only the correlations that exceeded the criterion value (.4785) were explored further.
$N = 28$
* $p < .05$
** $p < .01$

Table 11
Capacity-limited performance tests for Group 2, memory condition.

	Test Measure	Correlations		Partial correlations	
		accuracy	attributes identification	accuracy	attributes identification
Successive levels tests:		Secondary (level 2., concurrent) to		Primary task (level 3., alone)	
Distraction	phase 1.	-.2186	-.6965**	-	-.6955**
	phase 2.	-.0118	-.5438**	-	-.5707**
	full	-.2451	-.2938	-	-
Successive levels tests:		Secondary (level 3., concurrent) to		Primary task (level 4., alone)	
Distraction	phase 1.	-.4728	-.3827	-	-
	phase 2.	-.7458**	-.4566	-.7386**	-
	full	-.6307**	-.6327**	-.6224**	-.5304**
Non-successive levels tests:		Secondary (level 2., concurrent) to		Primary task (level 4., alone)	
Distraction	phase 1.	-.1965	-.5706**	-	-.5975**
	phase 2.	-.2395	-.3913	-	-
	full	-.4763	-.3051	-	-

Note, that only the correlations that exceeded the criterion value (.4785) were explored further.
$N = 28$
* $p < .05$
** $p \leq .01$

secondary task performance at the Phase 2 distraction condition correlated significantly with both measures of primary task performance. In addition, secondary task performance in Phase 1 distraction condition yielded a significant correlation with the attributes identification measure of primary task performance.

These results indicate that performance at the four-attribute level of the task is capacity-limited for the subjects from Group 2. Only one test was significant for the three-attribute level of the task.

The results for the Memory condition (Table 11) follow a similar pattern. Secondary task performance in the Phase 1 and Phase 2 distraction conditions at the two-attribute level of the task correlated significantly with the attributes identification measure of primary task performance at the three-attribute level. Both measures of primary task performance at the four-attribute level were predicted by secondary task performance at the three-attribute level in the Full distraction condition; the accuracy measure was predicted by the secondary task performance in the Phase 2 distraction condition. For the non-successive levels tests, secondary task performance in Phase 1 distraction condition correlated significantly with primary task performance at the four-attribute level, attributes identification measure. All correlations that exceeded the criterion value at the first stage of the analysis remained significant after the relevant predictors were partialled out.

Taken together, the results for the Perception and Memory condition clearly indicate that performance of the primary task at the four-attribute level is capacity-limited for the subjects from Group 2. There was also an indication of a capacity-limited performance at the three-attribute level of the task for certain conditions of the experiment but most of the tests were significant for the four-attribute level. Unlike the results from Group 1, the pattern of correlations did not suggest that subjects from Group 2 experienced a different degree of task difficulty in the Memory and Perception conditions.

Primary task performance of Group 2 subjects for the No-distraction trials was analyzed by means of repeated measures ANOVAs for each of the dependent measures, with task condition (Perception vs. Memory) and level of difficulty (two-, three-, and four-attribute level) as within-subjects factors. The analysis of accuracy data yielded main effect for level of difficulty $F(2,54) = 19.99$, $MSe = 625.50$, $p < .001$, qualified by a significant task condition by level of difficulty interaction, $F(2,54) = 5.98$, $MSe = 435.07$, $p < .01$, suggesting that differences in performance for the different levels of difficulty were not consistent across the two task conditions. That is why the follow up analysis was carried out separately for the two task conditions and compared the adjacent levels of difficulty within each task condition.

For the Perception condition, subjects' performance at the two-attribute level differed significantly from the performance at the three-attribute level, $t(27) = 5.42$, but the second comparison (three-attribute vs. four-attribute level) failed to reach significance, as revealed by the Dunn-Bonferroni test, $t'_{\alpha=.01}(4,27) = 3.33$.

For the Memory condition, two-attribute task performance was superior to three-attribute task performance, $t(1,27) = 4.40$, Dunn-Bonferroni $t'_{\alpha=.01}(4,27) = 3.33$. The

Table 12
Group 2: Means and standard deviations for the two measures of primary task performance (no-distraction trials) across task conditions and levels of task difficulty.

	Two-attribute task level		Three-attribute task level		Four-attribute task level	
"Perception" condition						
Accuracy (% correct)	79.82	(24.61)	46.39	(26.38)	55.96	(28.91)
Attributes identification (% correct)	94.00	(10.37)	83.50	(11.40)	77.64	(9.37)
" Memory" condition						
Accuracy (% correct)	73.96	(21.00)	59.64	(29.30)	42.82	(22.18)
Attributes identification (% correct)	90.96	(10.67)	87.43	(8.84)	82.36	(11.03)

N = 28

difference between three-attribute task performance and four-attribute task performance failed to reach significance.

The repeated measures ANOVA with task condition and level of difficulty as within-subjects factors yielded a significant main effect for level of difficulty $F(2,54) = 26.95$, $MSe = 81.35$, $p < .001$, and a significant task condition by level of difficulty interaction, $F(2,54) = 4.57$, $MSe = 55.75$, $p < .05$. It should be noted, that the source of interaction was in the differences between the two-attribute and three-attribute levels for the two task conditions. At the same time, performance at the two-attribute level of the task was near the ceiling (over 90%) and the differences in this region should be treated with caution. Nevertheless, the follow-up analysis was carried out for the two task conditions separately. As with Group 1, the analysis included comparisons among the adjacent levels of difficulty, evaluated against Dunn-Bonferroni $t'_{\alpha=.01}(4,27) = 3.33$. For the Perception condition, two-attribute task performance was significantly better than performance at the three-attribute level, $t(27) = 7.42$. The comparisons between adjacent levels of difficulty within the Memory condition failed to yield significant differences. Performance at the two-attribute level of the task, however, was significantly better than performance at the four-attribute level of the task, $t(27) = 3.57$.

The pattern of results for the two primary-task measures suggests that the subjects from Group 2 managed to overcome to a great extent the difficulties associated with the sub-task of identifying the relevant task variables and constructing the problem space, as evidenced by the nearly ceiling performance on the attributes identification measure. At the same time, Group 2 accuracy performance was far from being perfect. This indicates that the difficulties for this age-group are associated primarily with the processes of coordinating the available (and correctly identified) information into the model of the answer and the application of this model to the solution.

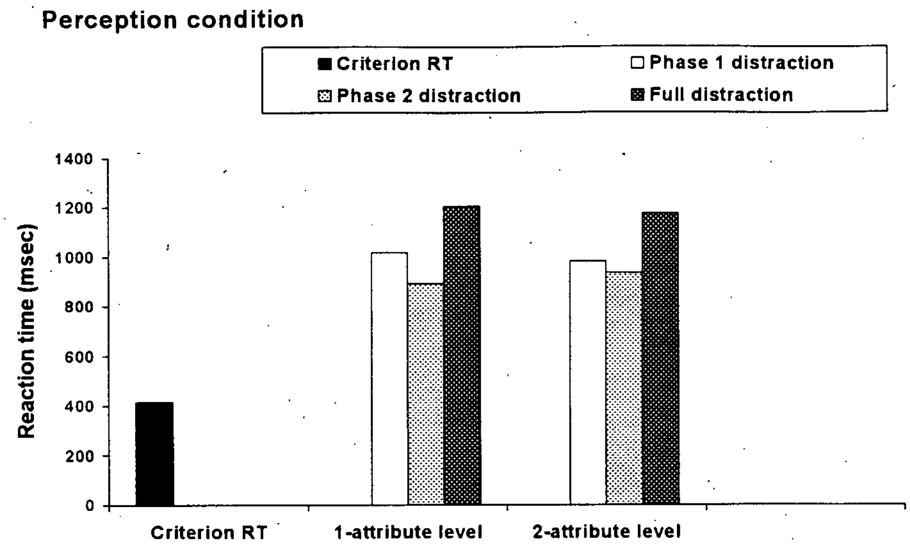
Means and standard deviations of the secondary task performance measures for Group 2 are displayed in Table 13 and Figure 7. As was the case with Group 1, performance on the secondary task alone was significantly different from performance on the secondary task in all distraction conditions, according to Tukey's HSD, $a(7,27) = 322.85$. The overall Task condition x Level of difficulty x Distraction condition repeated-measures ANOVA revealed main effect for distraction condition only, $F(2,42)$[1] $= 16.35$, $MSe = 132463.65$, $p < .01$. The follow-up analysis compared by means of paired contrasts (Dunn-Bonferroni $t'_{\alpha=.01}(4,27) = 3.33$) secondary task performance in Phase 1 and Phase 2 distraction conditions to Full distraction condition for each task condition and level of difficulty cell. For the Perception condition, the latency of the Full distraction trials was significantly greater than the latency of the Phase 2 distraction trials at the three-attribute level, $t(27) = 3.47$, $MSe = 66806.63$. For the Memory condition, performance of both Phase 1 and Phase 2 trials differed significantly from the Full distraction trials at the two-attribute level of the task, $t(27) = 3.63$, $MSe = 59732.55$, $t(1,27) = 3.94$, $MSe = 70254.81$, respectively.

[1]Because of violation of the sphericity assumption, the actual degrees of freedom were multiplied by Greenhouse-Geisser's Epsilon (.76881, in the particular case). The reported degrees of freedom and significance level are the corrected ones. The actual values are as follows: $F(2,54) = 16.35$, $MSe = 132463.65$, $p < .001$.

Table 13
Group 2: Means and standard deviations of the secondary task performance measures

	Secondary task concurrent performance at the two-attribute level of the primary task		Secondary task concurrent performance at the three-attribute level of the primary task	
"Perception" condition				
Phase 1. distraction	1018	(351.05)	981	(384.46)
Phase 2. distraction	892	(356.70)	935	(351.71)
Full distraction	1203	(555.72)	1174	(530.53)
"Memory" condition				
Phase 1. distraction	1009	(449.07)	923	(398.29)
Phase 2. distraction	925	(298.99)	942	(324.21)
Full distraction	1155	(555.09)	1220	(501.05)
Criterion R. T.		413	(159.41)	

The measures are in milliseconds.
$N = 28$



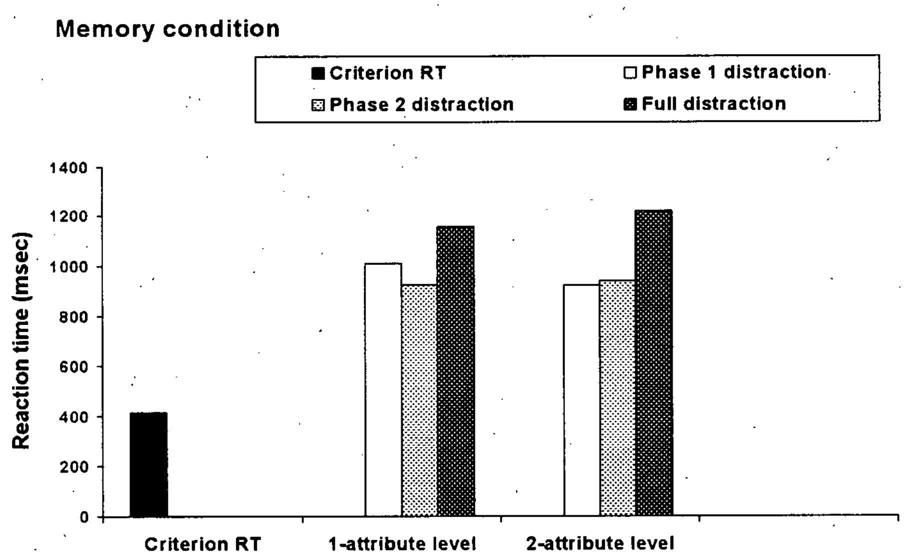


Figure 7. Group 2: Secondary task performance by task condition, level of difficulty, and distraction condition.

Thus, the results of the secondary task performance analysis confirmed the expectations for larger latencies for the Full distraction condition trials. The differences were significant at the level of the task where secondary task performance predicted performance of the primary task at the next difficulty level for this age-group. This result is consistent with the assumption that performance of the secondary task is sensitive to limits in capacity and that it can be a valid index of capacity-limited performance.

Group 3

The application of Hunt and Lansman's procedure to the data from Group 3 failed to detect capacity-limited performance on the primary task in the Perception condition. For the Memory condition, two significant correlations were obtained in the tests that related secondary task performance at the three-attribute level with primary task performance at the four-attribute level. Secondary task performance in Phase 2 distraction condition correlated with the attributes identification measure of primary task performance (r_{01}= -.6174, $p < .01$; $r_{01.23}$ = -.5677, $p < .01$), and secondary task performance in the Full distraction condition correlated with the same measure of primary task performance at the four-attribute level (r_{01} = -.6003, $p < .01$; $r_{01.23}$ = -.4534, $p < .05$).

Such results could be expected given the assumption of differences between the groups in the available resources, and in the context of the results from the other two groups. Group 1 capacity-limited performance for the primary task was predominantly at the three-attribute level; Group 2 displayed such performance at the four attribute level of the task. If subjects from Group 3 were characterized by a larger processing resource pool, one could expect to detect capacity limited performance for this group at a level beyond the four-attribute task level, or for only a few conditions of the four-attribute task. The obtained correlations can be interpreted as meeting the second proposition above.

Primary task performance of Group 3 subjects (see Table 14 for a summary of the means and standard deviations) was analyzed by means of repeated-measure ANOVAs for the two dependent measures separately, with task condition and level of difficulty as within-subjects factors. For the accuracy measure, there was a main effect for level of difficulty only, $F(2,52) = 13.21$, $MSe = 603.76$, $p < .001$. The follow-up analysis consisted of paired comparisons between the successive levels of difficulty in the two task conditions. The significance of the obtained differences was exaluated by comparing the obtained values to Dunn-Bonferroni's $t'_{\alpha=.01}(4,26) = 3.35$. Only the three-attribute vs. four-attribute task contrast for the Perception condition was significant, $t(1,26) = 3.70$.

For the attributes identification measure, performance at the two-attribute level for both Perception and Memory condition was excluded from the analysis due to ceiling performance. The repeated measures ANOVA with task condition and level of difficulty as independent variables revealed a significant main effect for level, $F(1,26) = 60.47$, $MSe = 46.15$, $p < .001$. Performance of the three-attribute level task was significantly better than performance of the four-attribute level task in both task conditions, $t(26) = 5.87$, for the Perception condition, and $t(26) = 5.11$, for the Memory condition.

As was the case with Group 1 and Group 2, performance of the secondary task alone (see Table 15 and Figure 9) differed significantly from performance of the secondary task in the different distraction conditions, as revealed by Tukey's HSD, a (7,26) = 233.59. The overall Task condition x Level of difficulty x Distraction condition

Table 14 Group 3: Means and standard deviations for the two measures of primary task performance (no-distraction trials) across task conditions and levels of task difficulty

	Two-attribute task level	Three-attribute task level	Four-attribute task level
"Perception" condition			
Accuracy (% correct)	84.07 (19.24)	77.89 (22.65)	56.89 (26.00)
Attributes identification (% correct)	97.48 (6.15)	87.78 (7.05)	75.96 (10.86)
" Memory" condition			
Accuracy (% correct)	84.07 (19.24)	71.78 (27.26)	63.07 (28.34)
Attributes identification (% correct)	96.26 (8.47)	90.22 (9.81)	81.70 (10.16)

N = 27

repeated-measures ANOVA revealed main effect for distraction, $F(2,52) = 10.99$, $MSe = 117362.84$, $p < .001$, and level of difficulty, $F(1,26) = 8.67$, $MSe = 102370$, $p < .01$. In addition, the interaction between task condition and level of difficulty was significant, $F(1,21)$[3]$= 5.76$, $MSe = 77824.64$, $p < .05$.

It should be noted, before proceeding with the follow-up analysis, that Group 3 was the only group for which a main effect for level of difficulty was obtained in the analysis of secondary task performance. This is in support of the validity of the dual-task approach to capacity and the assumption that performance on the secondary task in dual-task conditions reflects the spare resources. The obtained interaction, however, indicates that this effect was not constant across the two task conditions. The comparisons of the difficulty levels within the two task conditions revealed that the interaction effect was due to the significant increase of reaction time at the three-attribute level in the Memory condition, $F(2,52) = 7.47$, $MSe = 90097.32$, $p < .01$, that is, the only task condition, which predicted performance of the primary task at the four-attribute level.

The effect of the different distraction conditions was explored by means of paired comparisons that contrasted Phase 1 and Phase 2 distraction condition performance to Full distraction condition performance. The significance of the comparisons was evaluated by Dunn-Bonferroni's $t'_{\alpha=.01}(4,26) = 3.35$. In the Perception condition, Full distraction yielded significantly greater latencies than both Phase 1 distraction and Phase 2 distraction performance at the three-attribute level of the task, $t(26) = 4.87$, $MSe =$

[3]The actual degrees of freedom (i.e., 1, 26) were multiplied by Greenhouse-Geisser Epsilon (.802) because of violation of the sphericity assumption.

Table 15
Group 3: Means and standard deviations of the secondary task performance measures.

	Secondary task concurrent performance at the two-attribute level of the primary task		Secondary task concurrent performance at the three-attribute level of the primary task	
"Perception" condition				
Phase 1. distraction	758	(264.89)	804	(225.37)
Phase 2. distraction	832	(327.79)	817	(260.46)
Full distraction	958	(359.48)	1020	(405.19)
"Memory" condition				
Phase 1. distraction	749	(297.88)	834	(411.16)
Phase 2. distraction	787	(362.19)	952	(465.14)
Full distraction	863	(423.86)	1154	(637.35)
Criterion R. T.		406	(97.23)	

Note. The measures are in milliseconds.
N = 27

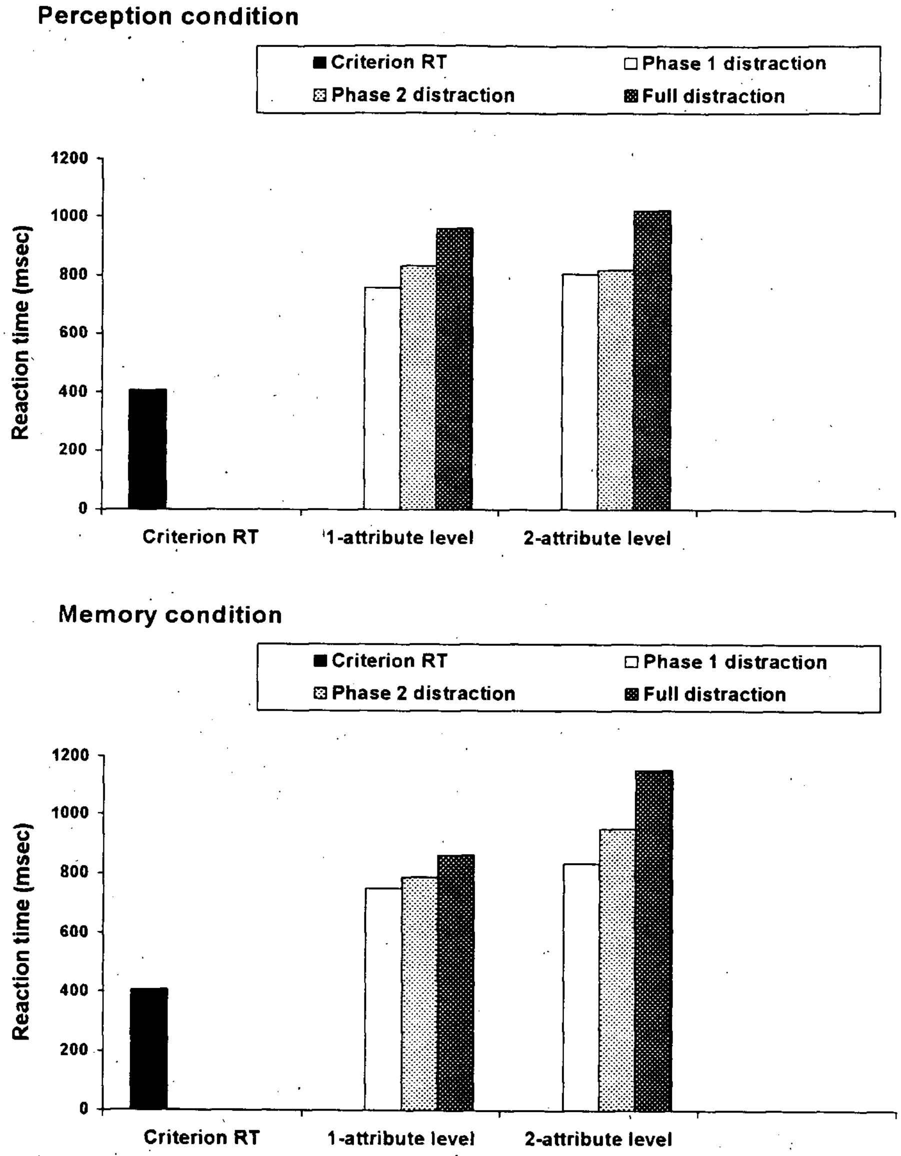


Figure 8. Group 3: Secondary task performance by task condition, level of difficulty, and distraction condition.

26423.30, and $F(26) = 3.56$, $MSe = 43449.91$. For the Memory condition, only the comparison between Phase 1 distraction and Full distraction performance at the three-attribute level of the task was significant, $t(26) = 8.94$, $MSe = 149111.89$.

Taken together, the results from the analysis of primary and secondary task performance can help in clarifying the results from Hunt and Lansman's procedure. For the accuracy measure, the deterioration of performance at the four-attribute level of difficulty in the Perception condition was preceded by a relatively constant level of performance of the secondary task, indicating that the joint demands of the tasks did not stress the limits of the individuals' resources, and thus, the decline in performance cannot be attributed to capacity limitations. Similar is the case for the Memory condition, where no significant difference between adjacent levels of the task was found, and the latencies from the concurrent performance of the secondary task at the two-attribute level did not indicate that resource limits have been stressed. The increase in reaction time at the three-attribute level was not followed by a significant decline in performance of the primary task at the four-attribute level, which explains the lack of significant correlations between the two tasks at these levels.. For the number of identified attributes measure, the ceiling effects at the two-attribute level of the task for both conditions, and the nearly ceiling performance at the three-attribute-level in Memory condition account for the lack of significant correlations between the measures at these levels of the task. The significant deterioration of performance of that measure at the four-attribute level for both task conditions was preceded by an increase in the reaction time for the secondary task in the Memory condition only, where, in fact, the only significant correlations were found.

Compared to the other two groups, subjects from Group 3 exhibited the most consistent pattern of performance on the secondary task through the different experimental conditions. There was no indication of change of strategies or inconsistent strategy use in the different task conditions.

Effects of Distraction on Reasoning Performance

The second objective of the study was to provide information about the character of the influence, exerted by limits in capacity, on reasoning performance. Such information can be obtained by comparing the performance in trials, where the capacity limits are not stressed, to performance under conditions of limited capacity. In the context of the present experiment, these are comparisons between performance in the No-distraction condition and performance in the different distraction conditions, within levels of difficulty, at which secondary task performance predicts successfully primary task performance at the next difficulty level. As a reminder, the secondary task in the distraction condition trials was supposed to use resources that were not allocated to the performance of the primary task. In cases where the joint demands of the two tasks stressed the limits of individuals' resource pools (i.e., at levels where the secondary task is expected to predict performance on a harder version of the primary task), performance of the primary task in the distraction conditions is, in fact, under conditions of insufficient capacity. At these levels of difficulty, information about the effect of limited resources on performance can be obtained by comparing performance of the no-distraction trials to performance in the different distraction conditions.

To illustrate, consider the hypothetical pattern of results displayed in Figure 9. In the example, a group of subjects has been given three levels of the primary task and capacity-limited performance has been detected at the highest level of difficulty (Level 3). Primary task performance has been predicted at this level by performance on a secondary task, administered at the second level of difficulty. Thus, the target level where one can compare performance under conditions of insufficient resources to performance that is backed up by enough resources, is the second level of the task. Consistent with the hypothesis that insufficient resources would prevent individuals from performing at a higher level, performance in the dual-task conditions of Level 2 was displayed as worse than performance in the No-distraction condition.

The analyses that follow, were carried out for each group and dependent measure separately because of the differences in task difficulty levels that the groups were exposed to, and because subjects from Group 1 were probably inconsistent in their approach to the task in the different task conditions, as suggested in the preceding section.

Group 1

Performance on the primary task, accuracy measure, for Group 1 subjects was first analyzed by means of a repeated-measures ANOVA with task condition (Perception vs. Memory), level of difficulty (one-attribute vs. two-attribute level), and distraction condition (No-distraction, Phase 1., Phase 2., and Full distraction), as within-subjects

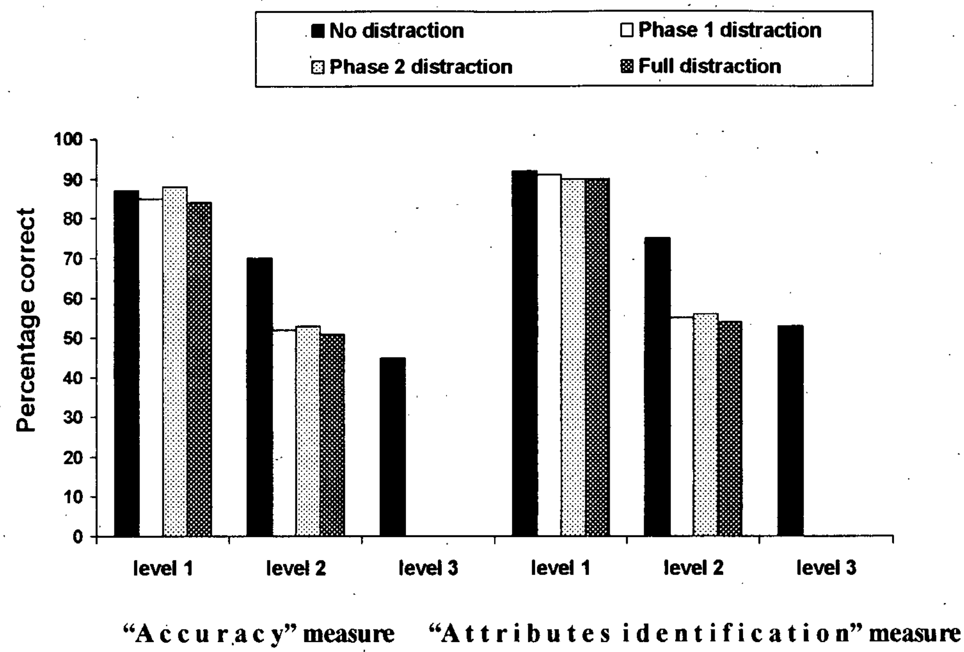


Figure 9. Hypothetical data pattern.

factors. The analysis revealed main effects for level of difficulty, $F(1,30) = 92.80$, $MSe = 1144.82$, $p < .001$ and distraction condition, $F(3,90) = 7.67$, $MSe = 457.69$, $p < .001$). In addition, the level of difficulty by distraction condition interaction was significant $F(3,90) = 2.80$, $MSe = 608.79$, $p < .05$, and the three-way interaction was significant, $F(2,63)$[4] $= 9.80$, $MSe = 415.46$, $p < .01$. To disentangle this complex relationship, the further analyses were carried out for each task condition separately.

For the Perception condition, the repeated measures ANOVA with level of difficulty and distraction condition as within-subjects factors preserved the established pattern of the effects: there were significant main effects for level of difficulty, $F(1,30) = 58.44$, $MSe = 804.71$, $p < .001$, distraction condition, $F(3,90) = 3.48$, $MSe = 458.23$, $p < .05$), and significant level of difficulty by distraction condition interaction, $F(3,90) = 7.92$, $MSe = 523.56$, $p < .001$. The first main effect reflects the decline in performance across difficulty levels (see the analysis of primary task performance for Group 1 in the previous section). The interaction indicates that the effect of distraction was not constant across the two levels: a result that allows us to attribute eventual differences at the two-attribute level to limits in resources.

Within-level comparisons involved contrasts between the No-distraction condition and each of the distraction conditions of each particular level of difficulty. The significance of the comparisons was evaluated against Dunn's $t'_{\alpha=.01}(4,30) = 3.30$. The analysis failed to detect significant differences at the one-attribute level of the task.

[4]The actual degrees of freedom (i.e., 3, 90) were multiplied by Greenhouse-Geisser Epsilon (.6997) because of violation of the sphericity assumption.

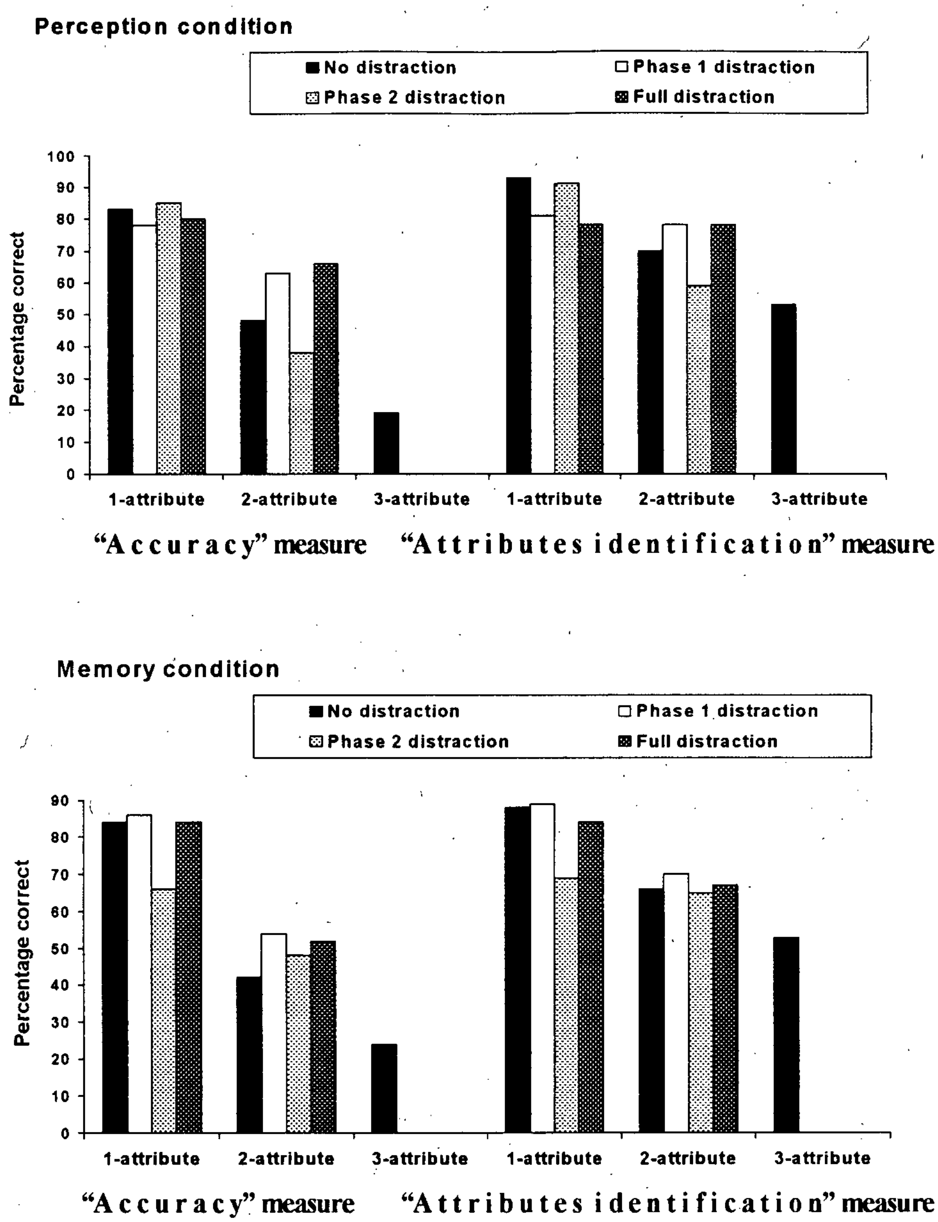


Figure 10. Group 1: Primary task performance (percentage correct) by performance measure, task condition, level of difficulty, and distraction condition.

Comparisons at the two-attribute level of the task (i.e., the level at which secondary task performance successfully predicted primary task performance at the three-attribute level), revealed significant differences between the No-distraction condition and each of the distraction conditions: $t(1,30) = 3.33$, $MSe = 401.16$, $t(30) = 3.68$, $MSe = 615.38$, $t(30) = 3.97$, $MSe = 293.37$, for the comparisons of No-distraction condition with Phase 1, Phase 2, and Full distraction, respectively[5]. Contrary to expectations, however, a significant decline was observed only for the Phase 2 distraction condition; the introduction of a secondary task in Phase 1 distraction condition and in Full distraction condition significantly improved performance.

For the Memory condition, the results followed a similar pattern. The Level of difficulty x Distraction condition repeated-measures ANOVA preserved the main effects for level of difficulty, $F(1,30) = 91.30$, $MSe = 652.62$, $p < .001$, and distraction condition, $F(3,90) = 3.52$, $MSe = 579.76$, $p < .05$, as well as the interaction between the two factors, $F(2,70)$[6] $= 3.25$, $MSe = 500.70$, $p < .05$. The comparisons within a level of difficulty indicated that only performance in the Phase 2 distraction condition of the one-attribute level was significantly lower than performance in the No-distraction condition, $t(30) = 3.39$, $MSe = 598.81$.

Several of the results, obtained from the analysis of the accuracy measure data, are worth more attention here. For the Perception condition, the task level at which

[5]The following order of the conditions within a level of difficulty was consistently used throughout the analyses: No-distraction, Phase 1. distraction, Phase 2. distraction, Full distraction. The comparisons within a difficulty level are presented as follows: No-distraction vs. Phase 1. distraction; No-distraction vs. Phase 2. distraction; No-distraction vs. Full distraction.

[6]The actual degrees of freedom (i.e., 3,90) were multiplied by Greenhouse-Geisser Epsilon (.7820) because of violation of the sphericity assumption.

performance in the distraction trials was under conditions of insufficient resources was the two-attribute level of the task. The comparisons between the no-distraction condition and each of the distraction conditions revealed significant effects for each contrast. That these effects were due to capacity constraints is supported by the pattern from the same comparisons at the one-attribute level, where no differences between no-distraction and distraction trials were found.

The direction of the effects at the two-attribute level, however, was quite surprising: the expectation for a decline in performance was confirmed only for Phase 2 distraction trials, while Phase 1 and Full distraction improved performance. These results have a plausible explanation in light of the different strategies that subjects from Group 1 probably used in the two task conditions, as suggested in the previous section. In the Perception condition, subjects tended to "shift" the solution towards the second part of the trial, spontaneously using a less demanding strategy of perceptual completion of the pattern to a "good form", instead of abstracting the relevant attributes and mentally constructing a model of the correct response. Naturally, in these conditions Phase 2 distraction would be disruptive, as far as it interferes with the major part of the solution process. Phase 1 distraction would facilitate performance because it disrupts the process of "perceptual" solution or the impulsive choice of an answer that "seems" appropriate, and turns subjects to the task of successively identifying the relevant attributes. Performance in the Full distraction condition would be improved, as well, following the improvement of Phase 1 distraction.

For the Memory condition, performance at both one-attribute and two-attribute levels was under conditions of insufficient capacity in the distraction trials. In addition, the analysis of secondary task performance in the previous section suggested that subjects used a strategy, different from the one used in the Perception condition, and allocated more efforts to the first part of the trials. Thus, one should expect that while the effect of the Phase 1 distraction condition is maintained at both levels of the task, the improvement of performance in the Phase 1 distraction and the Full distraction conditions would not be observed, because of the consistent application of the non-perceptual strategy across the conditions. The results fitted these expectations quite well. For the Memory condition, the deterioration of performance in the Phase 2 distraction condition appeared at the one-attribute level of the task, and no facilitating effect was observed for the Phase 1 and Full distraction conditions.

The analyses of the data for the second dependent measure followed the same sequence. The repeated-measures ANOVA with task condition, level of difficulty, and distraction condition as within-subjects factors, yielded significant main effects for all three factors: $F(1,30) = 7.57$, $MSe = 202.56$, $p < .01$, for task condition; $F(1,30) = 54.90$, $MSe = 501.63$, $p < .01$, for level of difficulty; $F(3,90) = 6.95$, $MSe = 268.06$, $p < .001$, for distraction condition. In addition, the level of difficulty by distraction condition interaction was significant, $F(3,90) = 4.25$, $MSe = 307.63$, $p < .01$, as was the three-way interaction, $F(3,90) = 11.37$, $MSe = 310.67$, $p < .001$. The further analyses were carried out for each task condition separately.

For the Perception condition, the repeated-measures ANOVA with level of difficulty and distraction condition as within-subjects factors, revealed significant effect for level of difficulty, $F(1,30) = 36.65$, $MSe = 347.82$, $p < .001$, and a significant level of difficulty by distraction condition interaction, $F(3,90) = 11.82$, $MSe = 323.77$, $p < .001$. The contrasts within a level of difficulty revealed significant decline of performance in Phase 2 distraction condition, as compared to performance in the No distraction condition, $t(30) = 4.59$, $MSe = 210.73$, at the two-attribute level of difficulty.

The analysis of the attributes identification measure in the Memory condition, revealed significant effects for level, $F(1,30) = 48.61$, $MSe = 305.17$, $p < .001$, and distraction, $F(3,90) = 4.96$, $MSe = 384.24$, $p < .01$, qualified by a significant Level of difficulty x Distraction condition interaction, $F(3,90) = 3.44$, $MSe = 294.54$, $p < .05$. The contrasts within a level of difficulty failed to reveal significant differences at the two-attribute level. At the one-attribute level, performance in the Phase 2 distraction condition was significantly worse than performance in the No-distraction condition, $t(30) = 10.82$, $MSe = 564.67$.

Group 2

In considering the results from Group 2 (see Figure 11) one should keep in mind that distraction trials at which performance is under conditions of insufficient resources are found at bouth levels of difficulty and for both task conditions. The analyses of these data was carried out following the same sequence as with Group 1. First, the data from

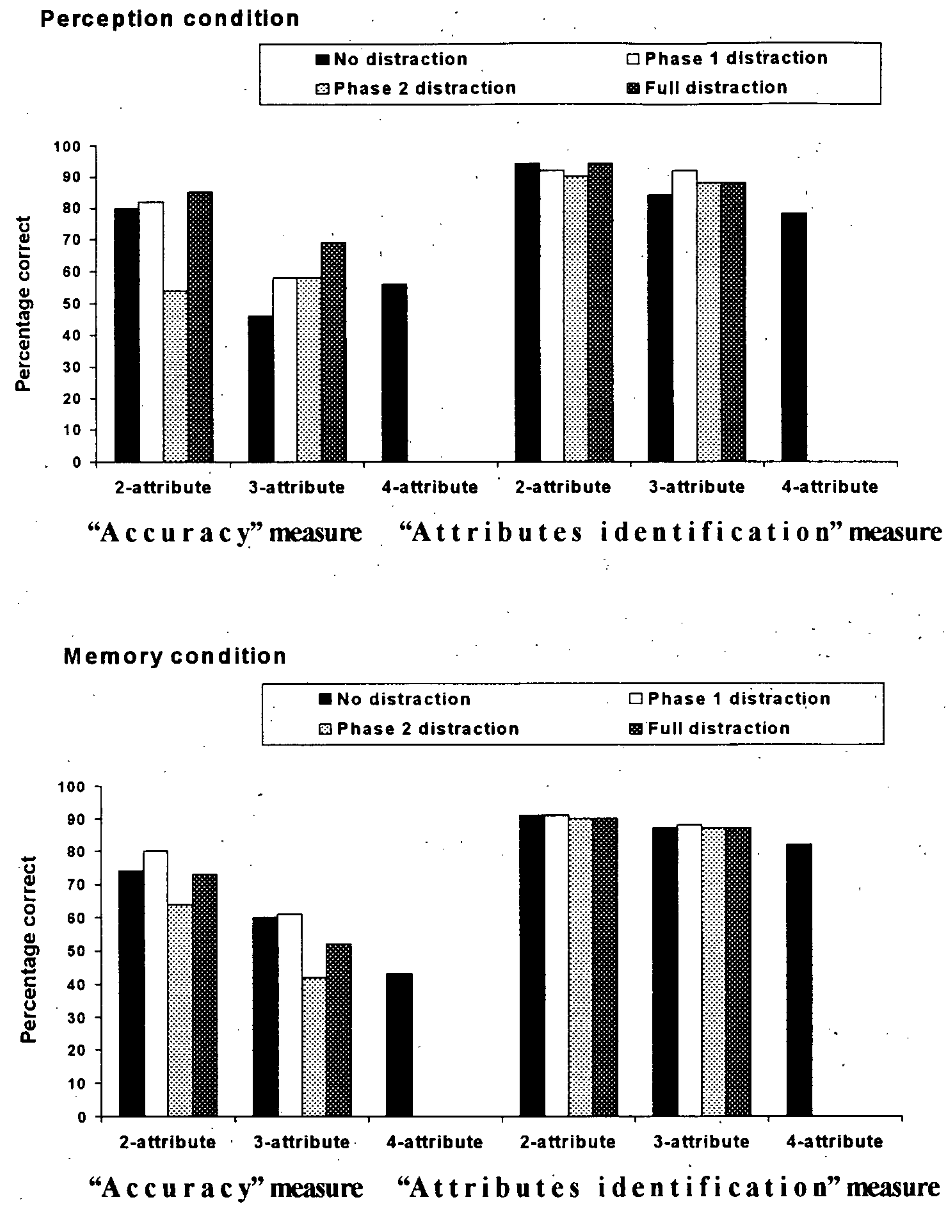


Figure 11. Group 2: Primary task performance (percentage correct) by performance measure, task condition, level of difficulty, and distraction condition.

the accuracy measure was analyzed by means of a repeated-measures ANOVA with task condition, level of difficulty, and distraction condition as within-subjects factors. The analysis revealed significant main effects for level of difficulty, $F(1,27) = 93.77$, $MSe = 391.43$, $p < .001$, and distraction condition, $F(3,81) = 9.14$, $MSe = 649.49$, $p < .001$, qualified by a significant Level of difficulty x Distraction condition interaction, $F(3,81) = 3.97$, $MSe = 428.00$, $p < .05$), and a significant three-way interaction, $F(3,81) = 4.99$, $MSe = 535.57$, $p < .01$. The further analyses were carried out for each task condition separately.

For the Perception condition, the repeated-measures ANOVA with level of difficulty and distraction condition as within-subjects factors, revealed significant effects for level of difficulty, $F(1,27) = 32.63$, $MSe = 499.07$, $p < .001$, distraction condition, $F(3,81) = 7.74$, $MSe = 589.40$, $p < .001$, and a significant Level of difficulty x Distraction condition interaction, $F(3,81) = 5.03$, $MSe = 730.49$, $p < .01$. The contrasts within a level of difficulty revealed significant decline in performance in the Phase 2 distraction condition at the two-attribute level of the task, $t(27) = 4.91$, and improvement in performance for the Full distraction condition at the three-attribute level of the task, $t(1,27) = 3.39$. The significance of the comparisons was evaluated against Dunn-Bonferroni's $t'_{\alpha=.01}(4,27) = 3.33$.

For the Memory condition, the repeated-measures ANOVA with level of difficulty and distraction condition as within-subjects factors, revealed only main effects for the two variables: level of difficulty, $F(1,22)$[7] = 41.46, $MSe = 495.39$, $p < .001$, and distraction condition, $F(3,81) = 6.43$, $MSe = 488.09$, $p < .01$. The comparisons of the No-distraction condition with the distraction conditions within a level of difficulty, showed significant deterioration of performance in the Phase 2 distraction condition at the three-attribute level, $t(27) = 3.29$.

The Task condition x Level of difficulty x Distraction condition repeated-measures ANOVA on the data for the second dependent measure yielded a significant main effect for level of difficulty only, $F(1,27) = 16.91$, $MSe = 93.36$, $p < .001$. This indicates that performance in the different distraction conditions and across the task conditions of the experiment followed a uniform pattern. Nevertheless, in order to be able to compare performance on this dependent measure with the pattern of results from the accuracy measure, the comparisons between the No-distraction condition and the different distraction conditions within a level of difficulty were carried out for the two task conditions separately. For the Perception condition, the planned contrasts failed to yield significant results at the two-attribute level of the task, and performance was better in the Phase 1 distraction condition than in the no-distraction condition at the three-attribute level $t(27) = 3.76$. No significant differences were found within the two levels of difficulty in the Memory condition.

These results reveal effects, similar to those that were detected for Group 1. Deterioration of performance in the Phase 2 distraction condition was found at both levels of the task: at the two-attribute level for the Perception condition and at the three attribute

[7]The actual degrees of freedom (i.e., 1, 27) were multiplied by Greenhouse-Geisser Epsilon (.80224) because of violation of the spherisity assumption.

level for the Memory condition. Improvement of performance was detected in the Full distraction condition (accuracy measure) and in the Phase 1 distraction condition (attributes identification measure). Note, that the latter two results were for the Perception condition. It is the task condition where similar effects were found in the analysis of the data for Group 1.

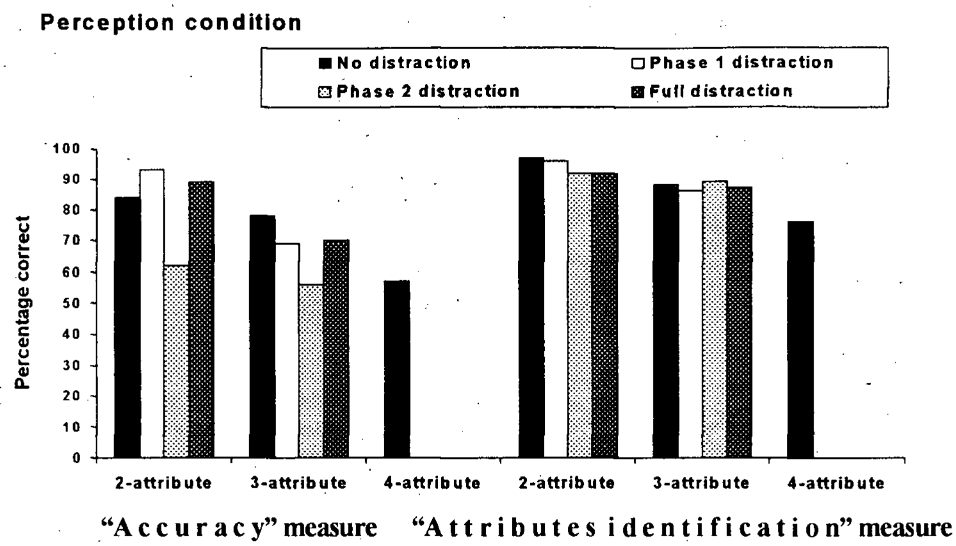
Group 3

Secondary task performance predicted performance of the harder version of the primary task for Group 3 in only two distraction conditions: Phase 2 and Full distraction conditions at the three-attribute level of the Memory task (see Figure 12). These are the conditions where performance is under conditions of insufficient capacity and where eventual capacity effects can be expected.

The repeated-measures ANOVA with task condition, level of difficulty, and distraction condition as within-subjects factors, revealed main effects for level of difficulty, $F(1,21)$[8] = 59.32, $MSe = 595.94$, $p < .001$), and distraction condition, $F(3,78) = 17.69$, $MSe = 430.31$, $p < .001$, qualified by a significant level of difficulty by distraction condition interaction, $F(1,26) = 5.28$, $MSe = 470.49$, $p < .05$. Although there was no indication of task condition effects, the follow-up analysis was carried out for the two task conditions separately.

The contrasts within a level of difficulty for the Perception condition revealed significant decline of performance in the Phase 2 distraction condition for both levels of

[8]The actual degrees of freedom (i.e., 1,26) were multiplied by Greenhouse-Geisser Epsilon (.79918) because of violation of the sphericity assumption.



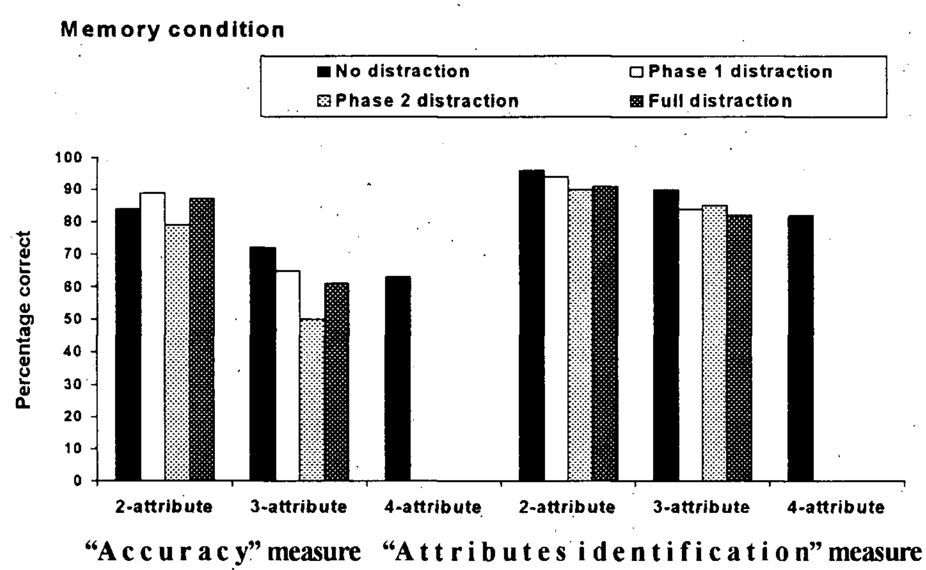


Figure 12. Group 3: Primary task performance (percentage correct) by performance measure, task condition, level of difficulty, and distraction condition.

difficulty: two-attribute level, $t(26) = 5.27$, and three-attribute level, $t(26) = 5.69$. For the Memory condition, quality of performance was significantly worse than performance in the No-distraction condition only for Phase 2 distraction condition at the three-attribute level, $t(26) = 5.64$. All contrasts were evaluated against Dunn-Bonferroni's $t'_{\alpha=.01}(4,26) = 3.35$.

The Task condition x Level of difficulty x Distraction condition repeated-measures ANOVA on the data for the second dependent measure yielded significant main effects for all three factors: task condition, $F(1,26) = 5.68$, $MSe = 56.06$, $p < .05$, level of difficulty $F(1,26) = 39.83$, $MSe = 140.17$, $p < .001$, and distraction condition, $F(3,78) = 6.66$, $MSe = 73.07$, $p < .001$. No significant differences were found between the No distraction condition and any of the distraction conditions.

These results are consistent with the expectations for a deteriorating effect of Phase 2 distraction on reasoning performance. Similar to the other two groups, no such effect was observed for the Phase 1 distraction condition.

Chapter 5

Discussion

The results from the study, as outlined above, demonstrated that in the context of a multiplication of classes and relations reasoning task, there is an increase with age in the quantity of information that can be processed at the different stages of the solution. Further, the analysis of performance under conditions of insufficient resources, allowed for specifying the impact that capacity limits have on the quality of the processes involved in these different stages of the solution process. In particular, the results indicated that: (1) a capacity "bottleneck" may exist at different points of the solution process and is associated with the quantitative characteristics of different processes; (2) the impact of insufficient resources may be quite different, depending on the character of the processes that are constrained; (3) there was an indication that younger subjects readily adopt a less demanding strategy for solving the problem, where the conditions of task presentation allowed such an approach. Taken together, these three groups of findings provide further evidence in support of the hypothesis, maintained by several theorists, that capacity is a necessary but not sufficient condition for the development of reasoning. It is necessary, however, to delineate the scope and validity of the present findings in comparison with the results generated by research under other models of the relation between reasoning and capacity.

Age Differences in Capacity

Age differences in processing resources were studied in the present work by means of a dual-task approach. Compared to the theories reviewed in Chapter 2, this approach is closest to the one used by Halford and his colleagues. These studies (Halford 1993; Halford, Maybery, & Bain 1986; Maybery, Bain, & Halford 1986), conducted in the context of transitivity and class inclusion tasks, used the model of dual-task performance proposed by Hunt and Lansman (1982), to prove that the reasoning performance of subjects (as young as 3 years for the class-inclusion task, and 5 years for the transitivity task) is capacity-limited.

The same model of dual-task performance was used in the present study to provide evidence for age differences in capacity. It was argued that if there is an age-related increase in processing resources, then capacity-limited performance will be detected at different points of the task scale for different age-groups of subjects. To test this hypothesis, four versions of a matrix completion task, with a logical structure corresponding to addition and multiplication of classes and relations, were developed. The versions, or levels of difficulty of the task, differed only in the number of attributes that defined the classification criterion. Three age-groups of subjects (6--8, 9 -- 11, 12 -- 14 years old) were given tasks from three consecutive levels of difficulty each, the first two levels in single and concurrent performance conditions, and the third -- in single performance only.

The results supported the predictions of the hypothesis. Capacity-limited performance was detected predominantly at the three-attribute level of the task for the youngest group of subjects. Subjects from the next oldest group exhibited capacity-limited performance mainly at the four-attribute level of the task. Subjects from the oldest group were obviously "out of scale" for the levels of the task they were given: capacity limited performance for these subjects was detected for only two of the tests at the four-attribute level of the task, that is, the last level at which they were tested. Assuming that there is an age-related increase in processing resources and given the results from the younger age-groups, the expectations are that performance of the oldest age group will be predominantly beyond the four-attribute level of the task.

An important question that should be dealt with here, concerns the validity of the established age trend. That is, to what extent the different findings for each age group are mediated by differences in subjects' processing capacity?

Several measures were taken at the stage of constructing the experiment to ensure valid results. The task scale, as mentioned above, was based on the number of attributes that defined the classification criterion. That is, the different levels of the task differed in the quantity of units that were to be processed. (This "metric" of the scale, and the uniform way of presenting the task levels, allowed for using two different in logical form tasks: logical addition and multiplication of classes and relations).

To ensure that differences in performance reflected the differences in the capacity demands of the task, subjects were warned as to what level of difficulty and kind of distraction to expect before each level of difficulty block and distraction conditions block. The process of the solution was explained in detail and a substantial amount of training was provided before the test trials began. To eliminate eventual effects due to the materials, to attributes that were used in a task, or to biases towards a particular answer, these aspects of the task varied randomly across trials, though care was taken that there were approximately equal number of each geometrical shape, attribute, or type of answer within each block of difficulty. In addition, the order of the trials within the concurrent performance blocks was random.

Finally, in order to control for different approaches to the task, two formats for presenting the task were designed. One of the conditions followed the traditional procedure for matrices completion tests, but allowed for a "perceptual" solution, that is, finding the missing shape through completing the pattern to a "good form". The other format precluded the possibility for such a solution and ensured that subjects used a strategy that was close to the instructions.

This manipulation proved to be quite useful. It was established that subjects from Group 1 readily applied the "perceptual" strategy when the conditions allowed it. This was reflected in their performance on the secondary task. Two distinct patterns of relative latencies were obtained in the two task conditions: in the Perception condition, efforts were concentrated towards the second phase of the trial (as evidenced by the longer reaction time for the probe in the Phase 2 distraction condition), while in the Memory condition, relatively more efforts were allocated to the first part of the trial (relatively longer latencies for the Phase 1 distraction condition) where the process of identifying the relevant attributes was supposed to take place. Performance of the same subjects on the primary task confirmed the expectation that the Memory condition would be more difficult for them. The fact that the effect of difficulty was found only for the attributes identification measure, indicated the process with which these difficulties were associated: the process of abstracting and identifying the relevant task variables.

These results allowed for identifying Memory condition as the condition where primary task performance, and consequently, the results from the application of Hunt and Lansman's procedure, were comparable across all three groups. Differences in performance for the Perception condition trials would tell more about the effectiveness of the different strategies, than about an eventual increase of processing resources with age. Taken separately, the results of Hunt and Lansman's procedure for the Memory condition reveal the same trend of detecting capacity limited performance at increasingly higher levels of task difficulty. The only obvious difference was that performance of the secondary task successfully predicted performance on a harder version of the primary task at a lower level of difficulty for a number of experimental conditions. This was explained by the overall higher demands of the task in the Memory condition due to the additional requirement to memorize the model of the answer at Phase 1 and the inability to shift some of the processes towards the next phase.

All in all, the comparison of performance in the two task conditions did not invalidate the observed trend for an increasingly larger pool of processing resources with age. In addition, the results indicated that capacity is not the only factor shaping performance: the present study detected the importance of the availability and accessibility of the different strategies for approaching a reasoning problem. The question of whether there is a relation between these two types of factors cannot be answered on the basis of the results presented so far, but it seems a promising path for future exploration.

Another set of results, in support of the validity of the observed trend, concerns the method for detecting capacity-limited performance. The present study used a version of the dual-task approach to capacity and relied on indirect indices of resources, instead of applying more or less "direct" measures of capacity. The question here is: to what extent was secondary task performance in the concurrent performance conditions a reliable index of spare capacity? If performance on the secondary task was an index of spare resources, then one should expect marked differences between performance of this task alone and in dual task conditions. This expectation was confirmed: for the three age groups, secondary task reaction time was significantly longer in all concurrent performance conditions, than when subjects responded to the tone only, without being engaged in solving a reasoning problem. Further, if secondary task performance was a reliable index of spare resources, then one should expect that the different difficulty of the primary task would be reflected in greater latencies for the secondary task when it is performed concurrently with a more difficult primary task. This pattern was found for Group 3 only. The failure to detect it in the results from Group 1 and 2 was attributed to the expected practice effects. Although not invalidating the overall results, this failure indicates the need for improvement of the experimental procedure. Finally, the Full distraction condition involved a repeated administration of the tone signal, and presumably, this increased the overall demands of the dual-task situation. If secondary task performance was a reliable index of spare resources, then one should expect longer reaction time in this condition as compared to performance on the secondary task in the other distraction conditions. This expectation was also met throughout the different experimental conditions: secondary task performance in the Full distraction condition was characterized by greater latencies although the differences did not always reach significance. These exceptions, however, could be explained by the excessive allocation of efforts to the processes in either the first or the second phase of the trial in some conditions.

The considerations, listed so far, support the validity of the observed trend for an increase of processing resources with age. In addition, they provide evidence for the validity of the dual-task approach to capacity and for its potential in the study of age-related changes in resources. It should be noted, however, that the results were obtained for processes that are specific to a particular task and a particular format of presenting the task. In this respect, one possible and necessary development of the present approach is the study of other forms of reasoning and of different formats of presenting the tasks.

Effects of Capacity Limits on Reasoning Performance

The hypothesis that capacity is a necessary but not sufficient condition for the development of reasoning involves two propositions. First, there is the assumption that processing resources increase with age. Second, there is the assumption that insufficient resources would prevent subjects from performing at a higher level.

The results relevant to the first proposition were discussed in the section above. The second one required analysis of the results at levels of the primary task where performance was identified as capacity-limited. Both groups, however, for which such levels were identified exhibited capacity-limited performance for the last level of difficulty at which the task was administered. As a reminder, there were only No-distraction condition trials at the last level of difficulty. This precluded the possibility for informative comparisons among the different conditions at these levels.

It was argued, however, that performance under conditions of insufficient resources could be observed for the distraction condition trials at levels where secondary task performance predicted successfully performance of the primary task at the next levels of difficulty. That is why the analyses were conducted on the data from the distraction trials at these levels.

Two groups of results require more attention. First, all significant effects of Phase 2 distraction were associated with declines in performance, when compared to performance of the primary task in the No-distraction condition. These effects were consistent for all three groups, for the two measures of primary task performance, and for both task conditions. Processes, related to the coordination of the identified attributes into a model of the answer, retaining this model in memory, and/or matching it to the given answers, are supposed to take place at this phase of the solution, according to the assumptions about the processes involved in performing the task. Thus, it can be concluded that the quantitative characteristics of these processes constrain the overall solution of the problem. In the present situation, the experimentally induced (by means of a secondary task, performed concurrently) deficit of processing resources prevented subjects from performing at a higher level. This result is in support of the hypothesis that a certain amount of resources is a necessary but not a sufficient condition for the development of reasoning.

The other results, however, were not in accord with the expectations that can be derived from this hypothesis. All significant effects of Phase 1 distraction were associated with improvement of primary task performance, when compared to No-distraction condition. In addition, such an effect was observed for the Full distraction condition, though only for the accuracy measure.

One explanation, discussed in the previous chapter, was based on the conditions under which this effect occurred. First, improvement of performance occurred in the Perception task condition only. Second, the effect was observed only in performance of Group 1 and Group 2. Third, for both groups it occurred at the level of difficulty, where the secondary task predicted successfully primary task performance at the next level of difficulty (i. e., at the two-attribute level for Group 1, and at the three-attribute level for Group 2). This last condition suggests that the effect was specific for performance under insufficient resources. The first two conditions suggest that the effect might be related to the different strategies that were applied to the solution by the younger subjects in the Perception version of the task. The analysis of secondary task performance of Group 1 subjects suggested that these subjects were attempting a solution, different from the one required by the instruction and most probably based on perceptual extension of the graphic properties of the stimuli to a "good form". The case might be similar with the younger subjects in Group 2. Thus, the following explanation can be proposed: the secondary task, applied shortly after the onset of Phase 1, disrupts the process of "perceptual" solution and subjects adopt a strategy that is close to the one required by the instruction. This latter strategy, being the more effective one in the context of the particular tasks, results in improvement of performance on the reasoning task in Phase 1 distraction condition trials as compared to performance on the No-distraction condition trials, where the less effective strategy is applied.

If this explanation is correct, then the important question here is whether the results disconfirm the hypothesis that insufficient resources prevent individuals from performing at a higher level. Several considerations should be taken into account in order to answer this question. It should be noted that the influence of distraction, according to the proposed explanation, was not direct. Performance improved due to overcoming the impulsive approach to the problem and undertaking a more efficient one. Whether this would happen if the insufficiency of resources was induced in some other way (for example, by a shorter time for this phase), is an open question that can be tested empirically. However, the obtained decline in performance of the primary task in the No-distraction condition at the next level of difficulty (i.e., where performance was capacity-limited) suggests that most probably improvement would not be observed. Thus, it can be accepted that the effect is specific to the type of distraction, in addition to being specific to insufficient resources.

Another way to look at the discrepant results for Phase 1 and Phase 2 distraction, is to consider the type of processes that were operating at the different stages of the solution. The identification of the relevant to the task attributes requires *sequential* processing of the material in the rows and columns of the matrix. The construction of the model of the answer and the matching of this model to the proposed responses requires that subjects *simultaneously* take into account the attributes identified at Phase 1. It could be that the two types of processing are affected differently by insufficient resources. In particular, subjects might be able to compensate for an imperfect performance on the subtask at Phase 1 at later stages of the solution.

This explanation is consistent with the claims of several theorists (e.g., Chapman, 1987; Halford, 1993) that the complexity of a problem is associated with the simultaneous coordination of a number of dimensions. It is also consistent with the pattern of results obtained for the Memory condition of the present study: administering a secondary task at Phase 2 consistently resulted in deterioration of performance on the reasoning task; disrupting the processes at Phase 1 had no effect on reasoning performance. From this point of view, the role of capacity as a developmental constraint would be associated with processes involving the simultaneous coordination of elements in an inferential scheme. It should be noted, however, that both explanations discussed so far are speculative and that specially designed studies are necessary for evaluating their adequacy.

The conclusion that can be drawn on the basis of the results from the present study is that the capacity limits of processes associated with the executive part of the solution constrained the quality of performance on a reasoning task. Thus, the capacity of these processes can be considered to be a developmental constraint as well because the insufficient resources for carrying out these processes prevented individuals from performing at a higher level.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory.* Vol. 2. New York: Academic press.

Atkinson, R. C., & Shiffrin, R. M. (August, 1971). The control of short-term memory. *Scientific American*, 82-90.

Baddeley, A. (1983). Working memory. *Philosophical Transactions of the Royal Society of London,* B 302, 314-324.

Baddeley, A. (1986). *Working memory.* Oxford: Oxford University Press.

Baddeley, A. (1992). Working memory. *Science*, 255 (5044), 556-559.

Baddeley, A., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory (pp. 47-89).* New York: Academic Press.

Case, R. (1972). Validation of a neo-Piagetian capacity construct. *Journal of Experimental Child Psychology, 14*, 287-302.

Case, R. (1974). Structures and strictures: Some functional limitations on the course of cognitive growth. *Cognitive Psychology, 6*, 544-573.

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

Chapman, M. (1981). Pascual-Leone's theory of constructive operators: An introduction. *Human Development, 24*, 145-155.

Chapman, M. (1987). Piaget, attentional capacity, and the functional implications of formal structure. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 20, pp. 289-334). San Diego: Academic Press.

Chapman, M. (1990). Cognitive development and the growth of capacity: Issues in neo-Piagetian theory. In James T. Enns (Ed.), *The development of attention: Research and theory,* Elsevier Science Publishers D. V. (North-Holland).

Chapman, M. & Lindenberger, U. (1988). Functions, operations and decalage in the development of transitivity. *Developmental Psychology, 24,* 542-551.

Chapman, M. & Lindenberger, U. (1989). Concrete operations and attentional capacity. *Journal of Experimental Child Psychology, 47,* 236-158.

Gelman, R. & Baillargeon, R. (1983). A review of some Piagetian concepts. In P. H. Mussen (Ed.), *Handbook of Child Psychology, Vol. 3: Cognitive Development* (pp.167-230). New York: Wiley.

Halford, G. S. (1982). *The development of thought.* Hillsdale, NJ: Erlbaum.

Halford, G. S. (1987). A structure-mapping approach to cognitive development. *International Journal of Psychology,* 22, 609-642.

Halford, G. S. (1993). *Children's understanding: The development of mental models,* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Halford, G. S., Maybery, M. T., & Bain, J. D. (1986). Capacity limitations in children's reasoning: A dual task approach. *Child Development, 57,* 616-627.

Halford, G. S., & Wilson, W. H. (1980). A category theory approach to cognitive development. *Cognitive Psychology, 12,* 356-411.

Hildebrand, D. K., Laing, J. D. & Rosenthal, H. (1977). *Prediction analysis of cross classifications.* New York: Wiley.

Hitch, G. J. & Halliday, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society of London*, B 302, 325-340.

Howe, M. L. & Rabinowitz, F. M. (1990). Resource panacea? Or just another day in the developmental forest. *Developmental Review, 10*, 125-154.

Hunt, E., & Lansman, M. (1982). Individual differences in attention. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 207-254), Hillsdale, NJ: Lawrence Erlbaum Associates.

Inhelder, B. & Piaget, J. (1969). *The early growth of logic in the child: Classification and seriation*. New York: W. W. Norton & Company. Inc.

Johnson, J., & Pascual-Leone, J. (1989). Developmental levels of processing in metaphor interpretation. *Journal of Experimental Child Psychology, 48*, 1-31.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Kail, R. (1988). Developmental functions for speeds of cognitive processes. *Journal of Experimental Child Psychology, 45*, 339-364.

Kuhn, D. (1983). On the dual executive and its significance in the development of developmental psychology. In D. Kuhn and J. A. Meacham (Eds.) On the Development of Developmental Psychology. *Contributions to Human Development, 8*, 81-110.

MacLane, S. (1972). *Categories for the working mathematician*. New York: Springer-Verlag.

Maybery, M. T., Bain, J. D. & Halford, G. S. (1986). Information processing demands of transitive inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 600-613.

Morra, S., Moiso, C., & Scopesi, A. (1988). Working memory (or the M-operator) and the planning of children's drawings. *Journal of Experimental Child Psychology, 46*, 41-73.

Navon, D., & Gopher, D. (1979). On the economy of the human processing system. *Psychological Review, 86,* 214-255.

Neisser, U. (1976). *Cognition and Reality.* San Francisco: W. H. Freeman.

Norman D. A., & Bobrow, D. G. (1975). On data limited and resource-limited processes. *Cognitive Psychology, 7,* 44-64.

Pachev, G., McBride, M., Carpendale, J., & Chapman, M. (1993). Effects of experimental conditions on the reasoning-capacity relation. *Poster presented at the SRCD Conference*, New Orleans, Louisiana.

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

Pascual-Leone, J. (1978). Computational models for metasubjective processes. *Behavioral and Brain Sciences, 1*, 112-113.

Pascual-Leone, J. (1980). Constructive problems for constructive theories: The current relevance of Piaget's work and a critique of information-processing simulation psychology. In R. H. Kluwe & H. Spada (Eds.), *Developmental models of thinking* (pp. 263-296). New York: Academic Press.

Pascual-Leone, J. (1984). Attention, dialectic, and mental efford. Toward an organismic theory of life stages. In M. L. Commons, F. A. Richards, and C. Armon (Eds.) *Beyond formal operations*. New York: Plenum.

Pascual-Leone, J. (1987). Organismic processes for neo-Piagetian theories: A dialectical causal account of cognitive development. *International Journal of Psychology, 22*, 531-570.

Pascual-Leone, J. & Goodman, D. (1979). Intelligence and experience: A neo-Piagetian approach. *Instructional Science, 8*, 301-367.

Pascual-Leone, J. & Smith, J. (1969). The encodings and decodings of symbols by children. *Journal of Experimental Child Psychology, 8*, 328-355.

Pascual-Leone, J. and Sparkman, E. (1980). The dialectics of empiricism and rationalism. A last methodological reply to Trabasso. *Journal of Experimental Child Psychology, 29*, 88-101.

deRibaupierre, A. & Pascual-Leone, J. (1979). Formal operations and M power: A neo-Piagetian investigation. In D. Kuhn (Ed.), *Intellectual development beyond childhood. (Sourcebooks on new directions in child development)* (pp. 1-43). San Francisco: Jossey-Bass.

Scardamalia, M. (1977). Information processing capacity and the problem of horizontal decalage: A demonstration using combinatorial reasoning tasks. *Child Development, 48*, 28-37.

Schneider, W. (1988). Micro experimental laboratory: An integrated system for IBM PC compatibles. *Behavior Research Methods, Instruments, & Computers, 20*, 206-217.

Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84* (1), 1-66.

Shiffrin, R. M. (1976). Capacity limitations in information processing, attention and memory. In W. K. Estes (Ed.), *Handbook of Learning and Cognitive Processes, Vol. 4. Attention and Memory.* Hillsdale, New Jersey: Laurence Erlbaum Associates, Publishers.

Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84* (2), 127-190.

Stewart, L. & Pascual-Leone, J. (1992). Mental capacity constraints and the development of moral reasoning. *Journal of Experimental Child Psychology, 54*, 251-287.