DEVELOPING STRUCTURAL REPRESENTATIONS : THEIR ROLE IN ANALOGICAL REASONING

Sharon L Ding, B.A. (Hons.)

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy, March, 1995.

ACKNOWLEDGEMENTS

I would like to thank the following people:

My supervisor, Dr. Claire O'Malley, for her advice and guidance.

Professor David Wood and the staff of the E.S.R.C. Centre for Research in Development, Instruction and Training for providing an intellectual environment which revived my enthusiasm at a time when it was beginning to wane.

Dr. Peter Bibby and Professor David Wood for commenting on the draft of this thesis.

Shaaron Ainsworth for her time and patience spent in making me computer-literate and also for the loan of her machine.

The staff and pupils of Berry Hill First School, Mansfield and Coxmoor Primary School, Kirkby in Ashfield, for their time, their hospitality and their friendliness.

My friends, particularly Shaar, Pete and Susie, for listening, supporting, advising and arguing.

My parents, for their practical help which aided the progression of this thesis.

My son James, whose existence continues to stimulate my curiosity, increase my knowledge and enrich my life.

The research reported here was funded by the Engineering and Physical Sciences Research Council (formerly the Science and Engineering Research Council)

ABSTRACT

Recent research into the development of analogical reasoning has shown that young children are able to recognise and use relational similarity between situations, provided that they possess the necessary domain knowledge (Goswami, 1992). However, in most of the reported studies, the relational structure of the analogy has been made very salient. Circumstances where the relational structure of a problem has to be represented by the problem-solver themselves could result in differing performance. We do not know whether, or in what circumstances, children can correctly construct a representation of the relevant relational information.

This thesis reports a series of experiments which investigate the role and development of structural representations for the purposes of analogical reasoning. The first two experiments tested whether primary aged children are able to construct an integrated external task representation by combining separate pieces of relational knowledge. Using series problems as a domain, they provided evidence that performance was not affected by the actual relation used, i.e. either spatial or non-spatial (abstract). However, it was observed that the order in which the task information was presented had an effect. The next four studies explored this by using spatial series problems. They showed that tasks which required a novel item to be placed to the left of (that is, at the front of) a partially ordered array inhibited performance. A further three experiments found that the reason for the inhibition was that unless the different pieces of relational information were highlighted as distinct items, they would be incorrectly integrated by using simple 'add-to-end' ordering rules. The final set of studies, using abstract evaluative relations in series problems, found that relational-highlighting effects generalised to these types of tasks. Also, the results showed that some evaluative relations were tied to either horizontal or vertical spatial representations and that performance was affected by how consistent the representation was with the child's experience of every-day life.

The thesis showed that the ability to construct structural task representations is affected by features which are inherent in the presentation of specific tasks, and that incorrect structural representations in turn affect analogical mapping. These findings are **discussed in terms of the 'generalised schemas' used during analogical mapping. It is** suggested that these might be reconstructed using specific task information, rather than being retrieved intact from memory.

TABLE OF CONTENTS

CHAPTER 1 : THESIS OVERVIEW

This brief chapter gives an overview of the eight chapters which form the theoretical and experimental components of this thesis. The research described is concerned with investigating some of the circumstances in which primary aged children are able to construct a relational task representation for the purposes of analogical mapping. The contents of each chapter are as follows:

Chapter 2 reviews some of the past research with both adults and children within the domain of analogical reasoning. A conclusion is reached from the adult literature that whilst we have a good idea of the core process of analogical reasoning, that is, the mapping of relations from base to target domains, more research is needed which investigates the pragmatic task constraints which affect the application of this core process. The review of the developmental literature accepts that children are capable of recognising and using relational similarity at a much earlier age than that stated in Piagetian stage-based theories of analogical reasoning. The last part of the chapter considers the importance of the way in which different analogical tasks are structured. It is suggested that those tasks which require the re-structuring of relational information are difficult for children, because they have difficulties in representing a task at a structural level.

Chapter 3 describes the choice of task for the research undertaken in this thesis. Series problems are chosen because they involve the integration of separate pieces of relational information into a single relational representation (Trabasso, 1975). If this is successfully carried out, the problem can be solved by the mapping of relational information from a 'generalised ordering schema' (Halford, 1992). Some of the past research involving series problems is reviewed, and their suitability for addressing the questions raised as part of this thesis is discussed.

The first two studies are reported in Chapter 4. Experiment 1 is exploratory, using series problems which involve a horizontal spatial relationship, and looking at the effect of using abstract and concrete problems. It was found that there was no difference with respect to context. However, it appeared that both 5 and 7 year old children were unable to correctly integrate the relational information into an external relational representation. The second experiment replicated a previous study (Pears and Bryant, 1992) in which 5 year old children were successful when solving spatial series problems in a vertical dimension. However, whilst some improvement over Experiment 1 was demonstrated, the precocious performance demonstrated by Pears and Bryants' subjects was not replicated. Because of the results from these two studies, it was decided to further investigate factors which could be affecting children's construction of a structural task representation.

Chapters 5 and 6 are concerned with the role of the order in which the relational information is presented to the subjects. Results from the four experiments in these two chapters lead us to conclude that those orderings which require that a novel item be placed to the left of (that is, at the front of) a partially ordered array result in inhibited performance. This finding is replicated for 7 year olds working with both drawings in an unfamiliar task and also when ordering familiar toys in a well-known context. It is also shown that 9 year old children are susceptible to this effect, but only when the novel item which needs to be placed to the left of an array is the last item required to complete the array.

Chapter 7 describes the findings of Experiment 7. This was carried out to test the hypothesis that children are unable to correctly integrate separate pieces of relational information because they are unaware that this information contains redundant information. However, there was no evidence to support this hypothesis.

Chapter 8 outlines the results of three studies which were designed to investigate the effect of task and information presentation relationships. Still using spatial series problems, the actual task relationship (vertical or horizontal) was varied together with the dimension in which the relational information was presented to the subjects. The results showed that when these two relationships were incongruent (different) performance was significantly better than when they were congruent (the same). This finding was replicated when the subjects worked with toys in a familiar context, and also when 5 year olds were integrating fewer items of information. These results are interpreted with regard to the salience of the 'gap' between different pieces of relational information.

Chapter 9 exploits the results obtained in earlier chapters and goes on to investigate 7 year olds performance when ordering abstract evaluative relations in series problems. The first experiment described in this chapter replicates the 'relational congruence' effect described in Chapter 8, but when using an abstract relationship. The next study explores the effect of using either horizontal or vertical base relationships when ordering evaluative non-spatial relationships. The results show that some relationships are facilitated by the use of one of these dimensions, whereas others appear to be unaffected. The final study begins to consider the extent to which young children's ordering schemas are tied to everyday experience. The results show that children are more successful in integrating relational information when the base domain is closely tied to familiar, plausible situations.

Chapter 10 begins by providing an overview of the research reported in this thesis and discusses some of the main findings from the studies. These findings are integrated into the following conclusions:

1. The development of children's structural task representations are adversely affected by features inherent in the task.

2. These features prevent children from correctly using relational information.

3. The resultant incorrect structural representations adversely affect analogical mapping.

It is also suggested that 'generalised schemas' for performing analogical mapping are not retrieved intact from memory, but might be reconstructed using information from specific tasks. The second part of the chapter discusses the implications of this, together with some suggestions for further research.

CHAPTER 2 : LITERATURE REVIEW - THE NATURE OF ANALOGICAL REASONING

The ability to perceive analogies is pervasive and is central for human cognition (Polya, 1957; Halford, 1992), being utilised in learning, problem solving and scientific discovery. As we go about our everyday lives, we are required to reason about new situations and experiences. This task would be much more difficult if we could not utilise some of our pre-existing knowledge in order to make sense of novel data. Analogical reasoning is a means by which we can apply our available knowledge to new situations. There is a general consensus that this involves the transfer of relational information from a domain which already exists in memory (known as the source or base domain) to the new (or target) domain (Vosniadou and Ortony, 1989). This emphasis on relational information is best illustrated by the use of an example. A common analogy is that the hydrogen atom is like the solar system (Gentner and Gentner, 1983). What is important is that the relations which exist between objects are common to both domains, but not that the attributes of equivalent objects are common. Thus the relations between the sun and the planets (e.g. internal forces) also exist between the hydrogen nucleus and electrons, but the sun's attributes (e.g. yellow and hot) are not attributes of the hydrogen atom.

Bearing in mind the centrality of analogy and its particular utility for learning, the development of this type of reasoning in children has considerable significance for both psychological and educational knowledge. This thesis is therefore concerned with young children's ability to reason analogically, and particularly with the development of the facility to represent a task in terms of its relations.

2.1 OVERVIEW OF LITERATURE REVIEW

This literature review will begin by describing examples of the different uses of analogy. Some of the psychological theories of analogical reasoning will then be discussed. These have built mainly on empirical work with adults, and as such do not emphasise the **development** of the ability. Nonetheless, they have added to our understanding of the various components which make up the analogical reasoning process. A review of these suggest some candidates which might be affecting children's ability to reason analogically. We then continue by discussing current developmental approaches, particularly the debate between structural (Piagetian) and knowledge-based theories. It will be suggested that a currently under-researched area, that of the development of the facility to represent tasks such that relational information is salient, may add a new element to our understanding of the domain.

2.2 ANALOGY IN USE

There are many examples in the literature of scientists using already understood areas as a source of information when working in novel and often problematic new domains (Holland, Holyoak, Nisbett and Thagard, 1986). The theory that sound consists of waves is extremely old, probably dating back to the Stoics (Samburski, 1973). The first discussion of the wave theory of sound comes from the Roman architect Vitruvius. He was aiming to describe the principles by which the greek amphitheatre was able to amplify sound. He concluded that the voice moved like the increasing circular waves which appear when a stone is thrown into water, and which continues to spread until prevented by an obstruction. Holland *et al* claim that the discovery of the wave theory of sound depended on the crucial analogy with water waves.

Perhaps the most famous scientific analogy is that made by Archimedes in the 3rd century BC. His task was to determine whether a crown made for the king was constructed from gold or base metal. Because the crown was so highly decorated, it was impossible to calculate its volume. Archimedes was stepping into a bath of water when he thought of the solution. He realised that as he got into the bath, an equal volume of water was displaced by his body. He then transferred the relation between the volume of his body and the volume of water displaced to the problem concerning the crown. In order to determine whether the crown was made of gold, he had to immerse it in water. He would then know the crown's volume, and could therefore calculate its weight.

It has to be pointed out that there is some argument in the literature concerning the centrality of analogy in scientific discovery. It would be wrong to suggest that these analogical insights occur without much prior painstaking work in the unknown area. However, it does seem that analogical insight can provide a new way of thinking about a domain which is novel and not fully understood.

Analogy has had a place in teaching for a considerable length of time. Secondary school children, learning new scientific concepts, are often reminded of other systems which they are familiar with. A frequently used example is the explanation of electrical current by use of the water flow analogy (Gentner and Gentner, 1983). The base domain is a plumbing system, and as water flows through the pipes of this system, so electricity flows through the wires of an electrical system. The reservoir of water is equivalent to the electrical battery, electrical current is similar to flowrate and electrical voltage is like water pressure. This analogy enables the student to differentiate between current and volts, as an increase in voltage occurs when two batteries are connected in series, just as water pressure is increased when two reservoirs at different heights are connected. Providing that the student understands the water flow system, new concepts about electrical current will be learnt.

In the psychological literature, there are many studies which look at people's abilities to solve novel problems using information which has already been presented to them in the form of a story. For example, Gick and Holyoak (1980) asked adults to solve Duncker's 'radiation problem'. The problem was posed as follows:

> *"Suppose you are a doctor faced with a patient who has a malignant tumour in his stomach. It is impossible to operate on the patient but unless the tumour is destroyed the patient will die. There is a kind of ray that in a sufficiently high density can destroy the tumour. Unfortunately, at this intensity the healthy tissue that the rays pass through on the way to the tumour will also be destroyed. At lower intensities the rays are harmless to healthy tissue, but will not affect the tumour either. How can the rays be used to destroy the tumour without injuring the healthy tissue?"*

In order to provide subjects with a potential base analog, the experimenters read out a story about a general who wanted to attack a fortress. Many roads converged on the fortress, but they all had mines placed on them so that a large group of people, for example an army, would detonate a mine and be blown up. A small group of people would be able to pass over safely, however. In order to launch a successful attack, the general needed to get his army, a large number of people, over to the fortress.

Different versions of the story described different solutions to the general's dilemma. For example, in one version the army was divided up into small groups and each group was sent down one of the roads so that all the groups met up at the fortress at the same time. The alternative version described how the general came across an unmined road and sent the whole army down that. All the subjects were then given the problem described above and asked to give solutions, using the story about the attack on the fortress to help them. Those subjects who had previously read the former story were more likely to suggest sending many weak rays from different directions which would converge on the tumour. On the other hand, those who had read the latter story were very likely to suggest sending rays down an open passage, such as the esophagus, thus hitting the tumour without coming into contact with any healthy tissue.

The common features shared between these three examples of analogy form the basis for its definition. In order for the use of the analogy to be productive, attention must be focused on the relations between the salient items in each of the initial situations (these are termed source or base domains) When the relations have been extracted, they can be transferred over to the new situation (termed the target domain) The point of the analogy is that the relations between items in the base domain can be applied to the items in the target domain. The 'convergence and summation of weak forces' relation which exists between the smaller armies and the solution in the first example can be mapped onto the target domain, and the 'avoidance of obstacles' between the unmined road and the solution in the second example can also be similarly mapped.

It seems then, that the use of analogies can facilitate scientific discovery, learning and problem solving. Since the late 1970's, considerable work has been carried out to investigate analogical thinking. The following section will review some of the theories which have arisen from the experimental performance of adults, and which have informed some of the developmental theories.

SOME THEORETICAL APPROACHES TO ANALOGICAL REASONING

Researchers working in the information processing paradigm have been concerned with identifying the different subprocesses that are used when reasoning by analogy. Spearman (1923) has been credited with introducing the first information-processing model of analogy solution. However, he was concerned only with the solving of classical analogies.

The term 'classical analogy' dates back to Aristotle, who defined analogy as 'an equality of proportions......involving at least four terms......where the second is related to the first as the fourth is to the third'. This definition concentrates on the equivalence of relations which join the pairs of terms. This equality of relations is conventionally represented in the format A:B::C:D (for example, up:down::big:small) The relation which links the C and D terms should be the same as that linking the A and B terms.

Spearman outlined three important processes in the solving of classical analogies: the apprehension of experience, the eduction of relations and the eduction of correlates. In simple terms, apprehension of experience is Spearman's term for the encoding of separate terms in the analogy, eduction of relations is the search for a relation between the first two terms of the analogy, and the eduction of correlates is the process whereby the relation between the A and B terms is applied to the C term. However, this model was neither tested or elaborated, probably due to a lack of appropriate experimental techniques. The terminology used in the original model was updated by Sternberg to bring it in line with modern cognitive psychology, and he also elaborated some of the processes into smaller components. The resultant (1977) model is probably the most well known and has inspired many further studies which have looked in more detail at some of the component processes involved. A brief description of Sternberg's model is given below.

The model describes the processes which Sternberg believed people went through in order to generate solutions to classical analogies. An example of such an analogy is:

Hospitals : Virginia :: Schools : 1. Peter 2. Paul

A : B :: C : D1 D2

<u>Note</u> The subjects had to select the correct D term

Sternberg divided the solution process into six components as follows:

Stage 1 : Encoding

When first presented with the analogy, each term is perceived, and the relevant attributes for each of these terms have to be accessed from memory.

Stage 2 : Inferencing

During this process the relationship between the A and B terms (Hospitals and Virginia) is determined and encoded in working memory.

Stage 3 : Mapping

Here the relationship between the A and C terms (Hospitals and Schools) is determined, thus linking the two halves of the analogy.

Stage 4 : Application

In this subprocess, a similar relationship to the one inferred at stage 2 (Inference) is applied to the C term and both D terms (1 and 2) in turn.

Stage 5 : Justification

When problems are presented in a forced choice format, there will be cases when neither of the D terms correspond exactly to the relationship inferred between the A and B terms. If this is the case, then one of the two options needs to be justified as the closest match.

Stage 6 : Response

Finally the chosen answer is given as a solution to the problem.

Since Sternberg first put forward his componential model, several researchers (e.g. Goswami, 1992) have commented that the mapping subprocess is not always a necessary step in the solving of analogies. In order to determine that the correct answer to the analogy above is 'Peter', we do not need to determine the relationship between 'Hospitals' and 'Schools' (the A and C terms). We can simply decide that the A-B relation is 'has government responsibility for' and then apply the relationship to the C term. Sternberg himself has also stated that he now views this process as optional (Sternberg and Rifkin, 1979) and so his original componential model has been slightly modified.

It seems however, rather than just being unnecessary, that for many analogies the mapping process described above would be at best inefficient. If one considers the A and C terms in isolation, then there are many ways in which they are related. Some examples might be that they are both buildings, both publicly funded or both employ many people. It is suggested that to generate these relationships without considering the B and potential D terms would be a very inefficient, and potentially unsuccessful strategy. It seems that we should be considering the structure of the task as an integrative whole, rather than isolating two items and determining the relations between them. It is not however within the aims of this thesis to investigate this claim.

One of Sternberg's aims in identifying the component processes was to test his claim that there were four possible solution strategies for analogical reasoning in adults. These models involve the same six subprocesses but differ along three different dimensions.

1. Exhaustion

A subprocess (inference, mapping or application) is termed exhaustive if the process considers all of the information available from memory before initating another process

2. Self termination

A subprocess (inference, mapping or application) is termed self-terminating if the process uses only the first piece of information available from memory and then initiates the next process

3. Option scanning

This concerns stage 4 (application). In sequential option scanning, all the inferences which have been made about the relationship between the A and B terms are applied to the C and D1 term before they are applied to the C and D2 term. This is contrasted with alternating option scanning where the first inference which has been made between the A and B terms is applied to the C and D1 terms and then to the C and D2 terms. This continues for each inference in turn.

Sternberg used a pre-cueing methodology to determine the time his subjects spent on each of the six sub-processes. This involved presenting either 0, 1 or 2 parts of a classical analogy to the subjects in the first half of an experiment, and then presenting the full analogy in the second half. The difference in problem solving times which resulted enabled Sternberg to hypothesise about which of the four strategies listed above was being used. He found that people with differing analogical abilities use the same strategy - exhaustive inference and self-terminating mapping and application but, more importantly, that better problem solvers spent more time on the encoding of the analogy, and less time on inference, mapping and application.

This conclusion was arrived at by firstly presenting one cue analogies. Thus, the subjects would be given the first part of the analogy:
Hospitals.

When they indicated they were ready, they were presented with the whole analogy:
Hospitals:Virginia::Schools:1. Peter 2. Paul

The subjects then responded as soon as they knew the answer. Sternberg treated the first score as the score for encoding the first term.

In a 2 cue condition, the subjects would be presented with the first part of the task:
Hospitals:Virginia.

When they indicated they were ready, they were presented with the whole analogy:
Hospitals:Virginia::Schools:1. Peter 2. Paul
The average encoding time times two (because there are two terms) was subtracted from the total time, and the remainder was treated as inferencing time (the second sub-process).

There are some problems with this approach, however. The subjects in Sternberg's experiments completed many trials, and can therefore be viewed as experienced problem solvers. Because the structure of all classical analogies are identical (the relation which exists between the A and B terms must be applied to the C term) then it seems very likely the subjects spent some time when being presented with the first term in retrieving likely relationships, rather than solely attributes, as Sternberg suggested. Also, it is unlikely that subjects would retrieve attributes for each of the four terms in isolation. As more terms are presented, they would begin to constrain their search to those relations which match with the terms they were already aware of.

Indeed, Sternberg himself argued that spending more time on encoding is a superior strategy as the later sub-processes can run more efficiently if a good problem representation has been achieved. This suggests that the subjects were already beginning to represent relationships, rather than with a potentially very large set of attributes for each term. This is interesting, especially if we consider that Sternberg's subjects were working only with classical analogies, where the structure of the task is inherent in the task itself. As long as we are familiar with the layout of these types of problems, we know that in order to arrive at the correct answer, we must apply the relation which exists between the A and B terms to the C term.

The types of analogies used in scientific understanding, learning and more 'open-ended' problem solving such as that studied by Gick and Holyoak are less well structured. This is because the domains are richer, resulting in several potential mappings which could be made between the base and target domains. Thus, as successful analogical mapping depends on the mapping of relations, the relational structure of these types of problems is much more important. The less salient these relations are, then the more difficulty subjects will have. It would seem therefore that the sub-process of encoding will also be crucial for 'ill-defined' analogical problems (those where the relational structure is not salient). However, Sternberg's definition of this sub-process as being concerned with perceiving each term and then accessing each term's attributes may well need amending. It is suggested that the sub-processes of encoding and inferencing in Sternberg's theory are not in fact separate components and should be amalgamated. This analysis of the sub-processes of analogical reasoning is consistent with one suggested very recently by Keane, Ledgeway and Duff (1994), which will be discussed in more detail later in this chapter.

As previously stated, Sternberg's theory has given rise to two later theories of analogical reasoning which are often placed in opposition to each other. These, both of which are situated within the information-processing perspective, will be discussed briefly below.

Gentner's structure mapping theory

The primary aim of the structure-mapping theory (SMT) is to define the crucial features which constitute an analogy and the computational operators which are necessary for processing analogies. Gentner describes the central idea in SMT as being the mapping of knowledge from one domain (the base) into another domain (the target) such that a system of relations which exists amongst objects in the base domain is carried over to objects in the target domain. Thus an analogy focuses on and uses relationships which exist between objects in a domain, without needing to take account of the actual objects. The principle of systemacity is central to the SMT, which holds that people prefer to map systems of relations rather than isolated predicates. The purpose of Gentner's theory is to describe what is specific about analogy. She accepts that the reasoner's plans and goals (the context of reasoning) can influence the reasoning process, and therefore must be taken into account. However, she does not accept that they are central to the analogical reasoning processor or indeed definitive of analogy. In view of this claim, she has proposed an architecture for analogical reasoning in which plans and goals influence reasoning before and after the analogical process, but not during it. Thus the problem context determines the representation of the current problem in working memory, which influences the retrieval of a potential analog from long term memory. Once this is accessed, the analogical process of mapping systems of relations from base to target begins. It ends by producing a candidate inference between these domains. This inference is then evaluated to determine whether the extent of the system match between domains is sufficiently extensive. If so, then the inference is further evaluated to see if it satisfies the reasoner's original goals.

In Gentner's model, therefore, the analogy processor is a separate cognitive module, which acts on the working memory representation of the target problem. It produces candidate inferences by mapping the relational structures of base and target domains. These results from the analogy processing module then interact with other processes, such as those which deal with contextual-pragmatic considerations.

Gentner claims that separating the context of the analogy from the analogy processor is justified because many cognitive processes need to take account of pragmatic considerations, whereas the matching process is unique to analogy and furthermore can occur in many different pragmatic contexts. She argues that in order to account for the use of analogy in scientific discovery, we must accept that the process can occur in the absence of initial problem solving goals. The advantage of a structure driven rather than a goal driven process, allows for the possibility of finding unexpected matches, or even matches which contradict initial problem solving goals. This is illustrated by the work of Poincaré, a mathematician, whose intention was to prove a particular theorem but who ended up by discovering a class of functions which proved the theorem to be incorrect. Thus, although Gentner does not deny the relevance of plans and goals (the pragmatic context) she firmly believes that the analogical reasoning process *per se* does not include these features, though it interacts with subsystems which deal with these considerations.

Holyoak's pragmatic theory

Holyoak (1985) proposed an alternative pragmatic account of analogical reasoning. He argued that Gentner's syntactic, or structure-based, approach cannot accurately predict the basis for analogical transfer because it fails to take account of goals. He stated that there are many different factors which affect the perceived structure of an analogy. Examples of these include which aspects of the source domain are normally considered to be important (Ortony, 1979), the goal in using the analogy and the central causal relationships which exist in the base domain (Winston, 1980). These factors will then in turn dictate which aspects of the base domain are transferred to the target. His account of analogical reasoning allows these features a central part in the process, and asserts that the structure of the analogy, far from being separable from pragmatic considerations, is actually closely tied to such considerations as the problem solver's goals and knowledge.

Holyoak also discusses the induction of generalised analogical schemas. He describes work by Gick and Holyoak (1983) in which groups of subjects received two stories which contained the 'convergence' theme discussed earlier . One involved a fire-fighting incident where converging sources of retardent foam were used to extinguish a large fire and the other was the army and fortress story described on page 8. Other groups received one convergence story and one other which was unconnected to the solution. All subjects were asked to write down accounts of how the stories were similar, i.e. to produce a 'problem schema'. This was intended to lead to the construction of a representation of the shared items. Subjects in the two analogy groups were more likely to produce the convergence solution, and this was interpreted as evidence that induction of a problem schema facilitated mapping between different contexts.

However, it has been pointed out that two analogs might be better than one because they offer two opportunities for similarities between base and target to be recognised. In order to control for this Catrambone and Holyoak (1985) replicated this experiment but varied whether or not subjects were asked to describe similarities between the two stories. They reasoned that if comparison is critical for schema induction, provision of two analogs will only be beneficial when subjects are asked to compare them. However, if the analogs are retrieved independently, then the comparison condition should have no effect.

The results showed that the provision of two analogs did not facilitate transfer unless the subjects were told to directly to compare them. Holyoak (1985) has interpreted these results in terms of 'problem schemas'. These are abstract categories which are linked to rules for categorising situations and constructing solutions. Once a person has induced a general schema, novel problems can be solved without retrieving examples of the original analogical situations, as the novel problems are categorised as instances of the schema. Holyoak differentiates this from 'strict' analogical reasoning as "it does not involve the direct transfer from a representation of a particular prior situation to a novel problem" (page 79).

In a similar situation, Halford (1992) argues that children use 'pragmatic reasoning schemas' as a basis for solving new problems. He suggests that this is carried out via the process of structure-mapping, i.e., the central feature in all current theories of analogical reasoning. Because his is a developmental theory, it is discussed at more length later in this chapter. However, it does seem that the construction of a problem schema does not preclude the claim that analogical problem solving occurs. People then map from the general schema to the target domain, just as when they map from a particular previous problem structure to the target. Indeed, it is possible to argue that the majority of analogical mapping involves the use of previously generalised schemas, and the provision of a base analogy serves mainly as a cue for retrieving the schema.

Comparison of structure mapping and pragmatic theories

A closer consideration of Gentner and Holyoak's theories would suggest that they are considering differing issues. Gentner is concerned with isolating the processes which she considers are **unique** to analogy, and as such has concentrated on the mapping of structural properties from base to target domains. Whilst she accepts that there is a clear need for explanations of pragmatic considerations, she sees no need to include these within a theory of analogical reasoning.

Holyoak, on the other hand, sees pragmatic considerations as being inexorably linked with the mapping process, such that a complete theory cannot exclude them. He feels that they have a direct effect on the structure of the base domain, which will in turn alter the nature of the analogical mapping. Thus to isolate processes which are unique to the mapping process of analogy is not warranted.

The two theories also differ in that Holyoak's encompasses the production of 'generalised problem schemas', which people use to focus on the relational structure of problems. Gentner does not deal with this, as she is concerned solely with structure-mapping *per se* as the central and unique process in analogical reasoning.

However, a central feature in both theories concerns the importance of domain structure. For Gentner, this is indeed the fundamental issue on which the structure mapping theory hangs. Whilst Holyoak argues that the pragmatic content of the reasoning task has a direct effect on the way the domains are structured, he is nonetheless accepting that the way in which this structure is represented will have a crucial effect on reasoning success. Thus these two theories differ from Sternberg's information processing account in that they explicitly address the issue of representational structure.

Sternberg's account, because it aimed solely to identify the subprocesses involved in the solving of classical analogies, did not need to consider the way in which the domain structure was represented. Classical analogies are of the form A:B::C:? . Here the domain structure in which the reasoning process needs to occur is inherent in the task presentation. As long as the reasoner is familiar with the format of these types of problems, they cannot fail to be aware that the relationship which exists between A and B has to be mapped onto the C term in order to produce a response. There are no other relationships or objects existing within this problem which are potential candidates for mapping.

As previously discussed, the situation is not as clearly defined if we consider the analogical reasoning process in problems such as those described at the beginning of this chapter, where the task structure is not explicit. It does seem that a full theory of analogical reasoning needs to consider seriously the issue of the problem representation. Whether or not it is encompassed within the analogical reasoning process *per se*, as in Holyoak's theory, or is a precursor to the analogy module as Gentner asserts, is not for the purposes of this thesis, particularly relevant. What is important is

1. That the way in which a problem is represented at the structural level will have a direct effect on the mapping of relationships between base and target domains.

2. That the use of common relational structures will induce the construction of generalised schemas.

A meta-framework for analogy

Keane (1994), drawing on a metatheoretical framework first proposed by Palmer in 1989, suggests that neither of the above theories are comprehensive, as they are both dealing with only the highest level of analysis. Palmer suggested that this level is concerned with characterising the informational constraints in a task situation, that is what an analogy is and what needs to be computed (cf. Marr's computational level, 1982). The midlevel concerns behavioural constraints (equivalent to Marr's algorithmic level), that is those which describe how people behave when solving different types of analogies and the different sorts of errors which are made. The lowest level in this analysis is concerned with hardware constraints, that is the physiological factors which affect analogical thought (Marr's hardware level).

Keane maintains that both Gentner and Holyoak, along with most other researchers in this area, are working at the highest level. He identifies three different types of informational constraints which exist at this level :structural, similarity and pragmatic.

Structural constraints

These are concerned with finding a one to one mapping between two domains. They exploit consistency between domains and favour systematic sets of relationships over single ones. Much of Gentner's work fits in this area.

Similarity constraints.

These can be used to decide between alternative candidate matches, in that only identical concepts will be matched between the two domains (Gentner, 1983).

Pragmatic constraints

These can also be used to disambiguate matches, as one match might be more goal relevant than another and so will be preferred (Holyoak, 1985) Keane also suggests that there are many other task demands, which have so far received little research attention, which might impose pragmatic constraints on analogical mapping. These include the specific instructions given or the way in which the materials are presented.

It is interesting to speculate on the way in which Keane feels these constraints might affect analogical reasoning. Keane himself provides no further details on this. It would seem that the way in which the task is presented might well affect the way the way in which it is represented. For example, different objects might be presented as more salient or the information might be presented in a certain sequence which could differentially affect the way in which the domain representation was constructed. This could be crucial when analogical mapping occurs, as we have already argued that the task representation can affect what relationships are mapped from base to target. This is an important point, and will be investigated during the empirical sections of this thesis.

The second level of Keane's meta-framework, that of behavioural constraints, contains fewer examples. He states that this is because much less work has been carried out in this area. Two constraints are briefly discussed which he asserts exist at this level.

Working memory limitations

Keane (1990) found that domains which were more conceptually rich resulted in more mapping errors. He suggests that this is due to parts of the problem representation being lost or forgotten as working memory becomes overloaded with information.

Background knowledge

Recent work by Keane (1991) showed that a mapping task can be performed faster if the mappings required are consistent with background knowledge than if they are inconsistent or neutral.

This recent discussion of the current analogical reasoning literature is calling for a much more comprehensive view of the domain, when compared to Sternberg's earlier theory. It asserts that it is important to consider the importance of initial task representations, their effects on the mapping process and what can alter the type of task representation which is constructed. This is aptly summed up in Keane's (1994) description of the stages of analogical thinking. He claims that there are five subprocesses which constitute analogical thought: representation, retrieval, mapping, adaptation and induction. These are described as follows.

Representation

In order to problem solve by analogy, the problem must be represented. Novick, 1988 has shown that problem representation affects analogy success.

Retrieval

The next stage is to retrieve an analogous problem solution from long term memory

Mapping

Corresponding elements or concepts in both domains are matched. Some or all of the relations in the retrieved base domain are then mapped across to the target domain on the basis of the matched items. Keane agrees with Gentner in that mapping is the crucial process in analogical reasoning, and is unique to the process. However, he also calls for research in the other sub-processes, stating that the vast majority of studies have been concerned with mapping.

Adaptation

In some circumstances, the first solution may not be totally correct in that it may need to be tested or adapted to fit the situation.

Induction

A higher order schema might be induced based on the two domains.

This description of the subprocesses differs from Sternberg's in several ways. These differences are mainly due to the fact that Sternberg's theory was explicitly developed to explain classical analogy, whereas Keane appears to be thinking about problem analogies. However, for the purposes of this thesis, the two important points which arise from Keane's account are:

1. It acknowledges the importance of the initial problem representation construction.

2. It allows for the induction of generalised schemas which can be held in long term memory.

These are two themes which have been taken up in our research, and as such will be returned to throughout this thesis.

The above conclusions concerning the importance of task representations and generalised schemas were arrived at by considering adult performance in analogical reasoning. This thesis is concerned with the way in which children develop structural task representations and the effect this has on analogical reasoning. We will now review developmental studies in the area, in order to investigate the current status of knowledge concerning children's ability to reason by analogy, and particularly their ability to represent a problem at the structural level.

2.4 DEVELOPMENTAL APPROACHES

Although a successful role for reasoning by analogy has always been presumed in educational circles, the topic has not received much attention from developmental psychologists until fairly recently. This seems surprising, considering the extensive literature which has been built up in the last two decades in the domain of cognitive psychology. Since childhood is a period in our lives when we have to deal with large amounts of novel data, it seems reasonable that the ability to make use of pre-existing knowledge would be of most use here. This was indeed the motivating force behind the research which is described in this thesis. If we can add to current knowledge concerning how children transfer knowledge to new domains, then we may be able to make suggestions concerning the use of analogy in educational contexts.

It is suggested that the two factors which emerged from Keane *et al's* review of the existing literature, namely the importance of the construction of the initial problem representation and the induction of generalised schemas, are also crucial when we consider the **development** of analogical reasoning ability. It will be argued that the first of these points has a significant effect on children's analogical performance and can account for some of the 'failures' which are reported in the literature.

Before we can discuss the evidence which leads us to this claim, it is necessary to describe the background to the current position concerning the development of analogical reasoning. This was reviewed by Goswami (1992) and much of the following draws heavily on her work.

Historically, the predominant research focus in this area has been based on Piagetian perspectives. These viewed the ability to reason analogically as a late developing cognitive skill, central to the 'formal operational' stage of logical thought. This period of reasoning typically emerged around the age of 11-12 years, only after children have become competent at thinking with progressively more abstract representations by passing through the preoperational and concrete operational stages. Because of these claims, primary aged children (those younger than 11 or 12 years) were thought to be incapable of reasoning using relational similarity and so, until relatively recently, very few studies were undertaken with this age group.

2.4.1 Piaget's stage based theory of analogical reasoning

Piaget's work on the development of analogical reasoning ability used a pictoral version of the classical analogy test, also used in psychometric testing. Most of these tests make the need to reason using relational similarity very clear. Children are told to select a D term from up to 4 or 5 alternatives which have been provided, such that the term bears the same relation to the C term as the B term does to the A term. This criterion of equality of relations is generally accepted to be the hallmark of analogical reasoning (Goswami 1992). According to Piaget, young children are unable to reason by analogy because they are incapable of reasoning using relational similarity. Instead they attempt to solve analogical problems by selecting D terms which are highly associated with the C term but do not share the same relationship as the relationship which exists between the A and B term. For example, a young child might well produce 'fat' as an answer to the problem 'up : down :: big : ?'. 'Fat' is a word which is highly associated with 'big', but the two terms are not antonyms, as are 'up' and 'down'. They can often explain their choice of response in terms of successive relations, such that they explain the relation between the A and B terms and also that between the C and D terms. However, they are unable to reason using relational similarity, that is they cannot equate the relations which they are applying to the C term with the relation which exists between the A and B terms.

Piaget's evidence for this claim came from a pictoral version of the classical analogy task which has since been adapted in further research (Goswami and Brown 1989, 1990) In Piaget's (1977) task, children were required first to pair pictures that went well together. This tested their ability to correctly link together the individual terms on separate halves of the analogy - Piaget named these lower order relations. They were then required to sort the pairs into foursomes that went well together and justify their selection. Piaget termed the relations between pairs higher order relations* . The majority of the children were able to pair up the pictures appropriately, but they found the latter part of the test much more difficult. Increasingly more helpful hints were therefore given to them when they were grouping the paired items into foursomes. If the children were initially unable to complete a foursome, they would be asked a question which highlighted the relationship between a pair. Following a continuing

* He maintained that the distinction between lower and higher order relations was psychologically significant. In other words, the ability to reason using higher order relations develops separately and at a later date from the ability to reason about lower order relations. Thus young children are unable to reason using relational similarity because this involves higher order relations which, it is claimed, are qualitatively different from lower order relations.

lack of success, the experimenter would select the appropriate C picture and tell the child that they should choose a picture which went with it in the same way that the A and B pictures went together. Finally, if the child was still unable to complete the analogy, they were shown three alternative pictures and asked to choose between them. This was to test Piaget's claim that in order to demonstrate full analogical reasoning ability a child should reject a 'counter suggestion' if it did not fulfil the relational similarity constraint that a full analogy entails.

Piaget used the children's performance on these tasks to identify 3 progressive stages in the development of reasoning by analogy. These correspond to his more general theory of mental development: preoperational at around age 5 to 6 years, concrete operational reasoning up to around 10-11 years and formal operational thereafter.

1. Preoperational stage

In this stage, the subject's pairings between pairs and the reasons for these pairings were extremely non-systematic. Some of the children were not even able to form the first pairings successfully (those of the A and B terms). Instead they linked terms using egocentric and idiosyncratic relations. However, as Goswami (1992) has pointed out, success in these task means choosing pairs of terms required by the experimenter. The different associations made by the children are no less 'correct' *per se* than those which were decided upon during the construction of the relational materials. Some of the more 'advanced' children were able to join the first and second terms appropriately but they could not go on to link together two pairs of terms by using equivalent relations for both pairings.

2. Concrete operational.

Here, the children occasionally showed successful performance. Some children however, could construct analogical foursomes by trial and error and were willing to accept false counter suggestions. Others were able to resist these false suggestions, but didn't do so consistently.

3. Formal Operational

This stage marked full successful performance in analogical reasoning. Here the children could easily and consistently extract and use higher order relations, could imagine alternative analogical relations and understand that mathematical notation can be used to represent analogical problems (e.g. glass/greenhouse=brick/house)

The children's performance in Piaget's tasks supported his three main claims about the development of analogical reasoning:

1. Children are unable to reason about relational similarity until the formal operational stage of development (around 11 to 12 years).

2. Reasoning about 'lower-order' relations (simple relations such as associative relations) develops before and is separate from reasoning about 'higher order' relations. This is needed in order to deduce relational similarity.

3. Younger children's lack of full understanding of the relational similarity constraint as shown by young children accepting 'counter-suggestion' solutions whereas older children and adults did not.

Evidence in support of Piaget and criticisms thereof

A large number of studies were carried out which followed in the Piagetian tradition, in that investigations were carried out into the three assertions referred to above. In most of the studies, success was measured by the ability to solve analogies correctly, and occasionally by the subjects being able to justify their choice by reference to relational similarity. The experimenters claimed that their results upheld the structural theory.

However this evidence, which is commonly cited in support of Piaget's claims, has been critically reviewed by Goswami (1992) She asserted that the studies which have claimed to produce results in support of Piaget's stage theory can be re-interpreted in terms of other constraints which could be affecting performance. The evidence, together with the relevant criticisms, is summarised below.

1. Analogical reasoning (reasoning about relational similarity) does not occur until the formal operational stage of development.

Gallagher and Wright (1979), working with 9 to 12 year old children, compared performance on 'concrete' and 'abstract' classical analogies. The former were those analogies which could be solved by "directly observable features of content". The latter were described as those where "solution seemed to be based upon a movement beyond observable features to a higher order rule" (page 116). It seems that the researchers were claiming that in order to solve 'abstract' analogies, the children needed to have conceptual knowledge of the problem domain. Their results showed that 9 to 10 year old children could solve the 'concrete' analogies, but that there was a large increase in performance on the 'abstract' analogies around the age of 12.

Although the younger children were correctly solving the 'concrete' analogies, Gallagher and Wright claimed that their success could be due to reasoning by association. Thus, although the children were giving the correct responses, they were not using the relational similarity constraint.

However, as Goswami has pointed put, it is impossible to tell whether the younger age group were using associative reasoning, because both associative and analogical reasoning would result in the same response. Also, the 'abstract' analogies required more conceptual knowledge than the 'concrete' ones. The poorer performance demonstrated by the 9 to 10 year olds on these types of problems could therefore be due to lack of the appropriate knowledge, rather than a change in ability to engage in reasoning using relational similarity. The children were not pre-tested for relevant conceptual knowledge before the experiment, neither were any tests carried out to determine which words were high associates of each other.

2. Reasoning about higher order relations (needed in order to deduce relational similarity) develops at a later stage than reasoning about simple, associative relations (termed lower order relations).

It seems that historically this claim was partly based on an assertion put forward by Inhelder and Piaget (1958). This stated that reasoning about analogies involved reasoning about proportions. Proportional reasoning was taken to be a corner-stone of the formal operational stage. Thus the conclusion was reached that analogical reasoning required formal operational skills. Following this argument, Piagetians such as Lunzer have expressed the view that the logical structure of classical analogies such as person:house::fish:pond is equivalent to proportional problems such as 3:4::15:20. However, Goswami, quoting the original writings of Piaget and Inhelder, states that their "actual beliefs about the connection between analogical and proportional reasoning are not very clear" (page 27).

Despite this lack of a claim of a strong link between the two types of reasoning, many researchers, assuming that measurement of proportional reasoning provides a measure of analogical reasoning, have investigated the former. Levinson and Carpenter (1974) looked at the extent to which an understanding of proportional reasoning affected the ability to reason analogically. They asked subjects to explain the reasons for their correct responses to classical analogies. Those responses which described relations within and between the two halves of the analogy were scored as demonstrating proportional understanding. It was demonstrated that the number of these responses increased significantly as the children got older. As there was also a relationship between the ability to explain responses using proportional reasoning and successful analogical reasoning, the researchers concluded that an understanding of proportionality may have contributed to successful analogical reasoning. Goswami points out that in these studies the measure of proportional reasoning was actually just another way of measuring whether the children could explain the relational similarity constraint. Thus the results of this study do not help in deciding whether analogical reasoning is related to proportional reasoning and might therefore be similarly classed as a formal operational skill.

Also, as Goswami has pointed out, some of the mathematical knowledge required to solve the series problems was quite complex and may well not have been present for the younger children. Thus they may have been unsuccessful because of a lack of domain knowledge, rather because they did not understand about proportional reasoning. It seems therefore that the claim for a link between proportional and analogical reasoning has not been proven. This in turn throws some doubt on the assertion made by Piagetian researchers that reasoning about higher order relations is equivalent to reasoning about proportions.

The claim that there exists an *a priori* distinction between higher and lower order relations, such that children can reason about lower order relations at a developmentally lower stage than higher order ones, has also been challenged (Goswami, 1992).

Lower order relations are those which exist between the A and B terms, and between the C and D terms in a classical analogy. For example, in the analogy horse:blanket::person:umbrella the lower order relations link together the paired items. Thus possible lower order relations for the A:B pair would be 'protects from rain' or 'keeps warm'. Higher order relations are those which link the two parts of the analogy i.e.. they link the lower order relations. In the example above this would be 'protects from rain' because both of the pairs of relations share that same lower order relation.

Traditionally, children below the age of around 11 years were said to be incapable of reasoning using higher order relations. They could reason only at the lower order level, so resulting in successive lower order relations being generated between the different pairs of items. Thus the younger children found it quite acceptable to use different relations liking the A and B terms to that linking the C and D terms. A change in reasoning competence was therefore brought about by the emergence of the ability to reason using the same relation.

Goswami claimed that the distinction made in the literature between the two types of relations was flawed. She points out that the classification into lower and higher order relations depends solely on the specific analogy used. So, in the example above, the higher order relation is 'protects from rain'. In another analogy, the same relation could be a lower order relation. Consider, for example, horse:blanket::person:hot water bottle. Here, the higher order relation is 'keeps warm', as it links both pairs of items. The relation 'protects from rain' is a lower order relation only, linking the items 'horse' and 'blanket'.

Thus, what is important is not the actual relations themselves, but the fact that the two halves of the analogy both share the same relations.

Therefore the status of relations cannot be defined independently as either lower or higher order. What is actually important is the child's knowledge about how paired items are related and their knowledge that the same relations must link both paired items. Goswami has termed this the 'relational similarity constraint' and states that the ability to recognise and use this constraint is the central issue which determines analogical competence. In order to solve analogies like yacht:wind::car:?, the child must apply the relational similarity constraint, such that the relation linking yacht to wind (makes move) must be applied to the choice of the 4th term.

3. Young children (below the formal operational stage) will accept non-analogical counter-suggestions, showing that they understand analogies only in terms of the successive relations between the two terms in each half of the problem. Older children reject such suggestions, as they are aware that a response must obey the relational similarity constraint.

Work by Gallagher and Wright (1977) found that 9 year old children accepted a non-analogical counter suggestion more often than the correct response. These responses were often justified by the children using different relations to link the C and D terms from the one which they had used to link the A and B terms. In other words, they were reasoning with successive relations, rather than relational similarity. However, again we are faced with the possible confound due to a lack of appropriate knowledge. This could either be relational domain knowledge, or knowledge about the nature of the task. If children do not know the relation on which the analogy is based, then they may select a counter suggestion which is highly associated with the C term. Also, if they think that the task requires the application of successive relations, then it is not surprising that this is what they actually do! This result therefore does not necessarily mean that the subjects were unable to use the relational similarity constraint *per se*, rather that they lacked the appropriate knowledge concerning either the nature of the task or the relationships between the task items, to enable them to apply this constraint.

It has also been suggested that in some cases the children might have chosen the counter suggestion because they felt that it did satisfy the relational similarity constraint. This could be due to either a lack of the appropriate linguistic knowledge in order to be able to make a distinction, or to not having a comprehensive understanding of the nature of the task. This is not the same as being unable to reason using the relational similarity constraint.

Goswami summarises her review of the evidence in support of Piagetian theory by concluding that:

1. The evidence concerning analogical reasoning as a formal operational ability is confounded by a possible lack of the relevant relational knowledge. A failure to test for the presence of this knowledge means that the children's apparent failure could be attributed to domain level issues, rather than to an inability to reason using the relational similarity constraint.

2. The research which investigated a link between higher order reasoning and an understanding of proportional equivalence is inconclusive. We have no evidence that there is a strong link between the two, as the studies have been confounded by a possible domain knowledge lack and a confusion between the measurement of proportional reasoning *per se* and an understanding of the relational similarity constraint.

3. Studies which claimed that children below the age of 11 years solved analogies by reasoning about both of the two halves of the analogy in succession, or by associative

reasoning, were flawed. The studies which cited an acceptance of false counter suggestions as evidence for these claims could be explained by a lack of relevant domain knowledge, or because of semantic difficulties or misunderstandings about the task.

2.4.2 Empirical evidence which challenges Piaget's structural theory

Because of these fundamental flaws in the Piagetian evidence summarised above, Goswami continues by considering more recent evidence which directly set out to test some of the theory's assumptions. This is reviewed below.

1. Do children possess the relevant relational knowledge?

Before the child can determine the relations used in an analogy, they must have the relevant domain knowledge. Without this knowledge, they will not understand why paired items are related, and so it follows that they will not be able to apply the relational similarity constraint. A child cannot select relations which are similar if he or she does not know or understand the individual relations.

Consider the analogy wigwam : hide :: igloo : ?
If a child doesn't know that a wigwam is made from hide, then he or she cannot apply the 'made from' relation to the second half of the analogy. Similarly, if the child knows the relation between 'wigwam' and 'hide', but doesn't have the 'made from' knowledge associated with 'igloo', then the task still will not be completed successfully.

Surprisingly, Piaget and his colleagues failed to pretest their subjects (e.g. Lunzer, 1965, Gallagher and Wright, 1979) Thus their apparent failure to demonstrate competence in applying the relational similarity constraint may have been due to their lack of relevant domain knowledge. If the relations involved in the task are too difficult for the subjects, then they will necessarily fail.

Recent research has used simple causal relations and has pretested subjects for possession of the relevant knowledge. Goswami and Brown (1989) gave 3, 4 and 6 year old children pictoral classical analogies based on relations such as cutting and melting. They were required to pick out the correct D term to complete the analogy from a choice of 5 pictures. For example in the 'cutting' analogy, subjects saw pictures of a tub of playdoh and cut playdoh (A and B terms) and then a picture of an apple (C term). Five alternatives were chosen to give children an associative choice (banana), a different change to the same object (bruised apple), a similar appearance (ball) and the same change with a different object (cut bread) In addition to the analogy condition, the children's relational knowledge was also tested. This was done by showing the child three pictoral examples of the action e.g. cut apple, cut cake and cut playdoh, and the children were asked to choose the picture of the object which had caused the action (again they had five alternatives) This study showed that the children who understood the causal relations could also solve the analogies. The older children were more successful with the analogy task, but they also knew more of the causal relations. Similar results were obtained when the objects in the analogy underwent perceptual change, for example a whole lemon and a slice of lemon. Thus it would seem from these studies that the apparent inability in the Piagetian studies might be due to a lack in the relevant relational domain knowledge.

2. Can children reason about proportions?

Piagetian research was based on the assumption that the skills required for proportional and analogical reasoning are very similar, in that both require reasoning about relations which are common to both halves of a problem. It was therefore reasoned that neither type of problem would be solvable by children below the formal operational stage. In order to test this, Goswami (1989) required children to reason about proportional classical analogies. She showed children pictures of proportions of shapes (for example 1/2 circle) where the relevant portion of each stimulus was shaded. An example would be 1/4 square:1/4 rectangle::1/2 square:?. A correct response to this problem required that the child had to reason about the proportions involved. Again five alternative responses were provided which were designed to test whether the subjects were reasoning using similarity matching, partial extraction of relations (that is, just shapes or shaded areas) or proportional reasoning (the correct response). The results showed that success in solving the analogies was related to relational knowledge, that is, the children's knowledge about proportions.

It would seem that we have more evidence that the earlier failure reported in the literature was due to a lack of relational knowledge. In Goswami's experiment, children as young as 4 years could reason about proportions. Thus some understanding of the equivalence of proportional, as well as analogical reasoning ability, is present well before the formal operational stage.

Nonetheless, Goswami has made the point that the subjects in her study could have been successful by using perceptually based proportional reasoning, whereas it seems likely that Piaget was more concerned with a logical understanding of proportions. However, this does not appear to have been clearly defined in the literature.

3. Do children reason about successive relations?

Piaget claimed that younger children try to solve analogies by reasoning successively about the two halves of an analogy, and this is why they are willing to accept incorrect counter-suggestions. Other researchers have claimed that children respond with words which are high associates of the C term, rather than those which obey the relational similarity constraint.

Goswami and Brown (1990) tested this claim by careful selection of the alternative responses provided for the subjects. They based their choice on category sorting tasks, which show that 3 to 5 year olds will sort items thematically rather than categorically (Markham and Hutchinson, 1984). They therefore decided to provide a choice of responses which were 'incorrect' but which were either highly related thematically or categorically to the C term, or were similar in appearance to the 'C' term. These were contrasted with the 'correct' analogical response which was less highly associated. An example of this would be Spider:Web::Bee:?, where Honey is the highly associated thematic relation, Ant is the categorical match, Fly is the appearance match and Hive is the correct analogical response. The correct response was chosen by 4 year olds for 59% of the time, showing that the children were able to reason using relational similarity rather than thematic, categorical or appearance matches.

Types of analogical tasks

In common with investigations into adult analogical reasoning, research into the development of the ability has been carried out using both classical and problem analogies. However, the above review was based mainly on the former, as problem analogies have only been investigated within the last decade or so, compared with classical analogies which were first used by Piaget and colleagues (Piaget and Inhelder, 1958: Piaget at al. 1977).

The problem analogies which have been used are similar to the ones used with adults (e.g. Gick and Holyoak, 1980), in that the solution can be reached by mapping the relational structure of a story (the base domain) into a novel problem (the target domain). Typically, the problems and analogous stories are about everyday concepts and relations with which children are familiar.

Research using problem analogies has demonstrated successful performance at a younger age than that of the Piagetian experiments which used classical analogies. Indeed, this discrepancy was used as an argument for the early acquisition of analogical reasoning skills (Goswami, 1992). Because the problems were concerned with knowledge which was easily available to children, they did not come up against inhibition of performance caused by a lack of the relevant relational knowledge.

Collins and Burstein (1989) suggest that problem analogies are simpler than classical analogies because of the number of elements which need to be mapped. This is therefore an alternative explanation for the discrepancy in performance referred to above. They describe classical analogies as 4 element problems (i.e. the four terms) which require both within and between group comparisons. Those between the A and B terms and the C and D terms are within group, whereas those between the A and C terms and the B and D terms are between group. In this system of analysis, problem analogies are two element comparisons as there is a single item (the base system) to be compared with another single item (the target).

However, it is also possible to represent problem analogies in a classical format (that is, as 4 element comparisons). Thus for the Gick and Holyoak Problem:

Small armies : fortress :: weak rays : tumour

Here the relation which links both halves of the analogy is 'converges on'. This analysis highlights the relationship between items as the salient aspect of the mapping. An alternative classical representation of the problem highlights the relationship between problem and solution as the salient aspect:

Problem (base):Solution (base)::Problem (target):Solution (target). (Holyoak, 1984)

This is consistent with Holyoak's claim that the two levels of analogy are different in that in classical analogies the nature of the task (that analogical mapping is required) is made clear and the goal is to solve the analogy, whereas in problem analogies the goal is to solve the problem and the problem solver needs to recognise that an analogical mapping is required. Thus there seems to be no *a priori* distinction between the number of element comparisons required in the different types of analogy. It appears that the initial discrepancy in the literature is due to the different types of knowledge often used in classical and problem analogies. When classical analogies use simpler relational knowledge children show successful performance at an early age, just as they do with problem analogies.

However, another way of describing the distinction is that in classical analogies the task structure is made clear, in that there can be only one base analogy (that between the A and B terms). Thus the goal is to identify the items which complete the pattern of relational similarity. On the other hand, in problem analogies, the initial goal is to identify and represent the correct base relationship, such that it is possible to complete the relational similarity pattern. This highlights again the importance of the initial construction of the relational task representation for many analogies. In classical analogies, this stage is often trivial, and the relational structure is explicit in the task itself. In problem analogies, and in 'real-world' educational contexts, the relational information which requires mapping is embedded within other descriptive elements of the structure. Thus the explicit representation of the significant objects and their relationship to each other (i.e. a representation at the structural level) is a non-trivial process with the potential for error. It could be therefore that whilst both types of analogical tasks involve the same core process, that of the mapping of relational similarity, the richer context of problem analogies might mean that it can be more difficult to isolate the appropriate relation. Thus, as we can represent both classical and problem analogies as 4 element structures in Collins and Burstein's analysis, there are other important differences between them which will differentiate effective performance. This performance difference is caused by affecting the ease by which relational similarity is recognised and used. Some of the factors which affect the application of relational similarity have already been investigated and are described below.

2.4.3 Understanding the task requirements - the role of meta-knowledge

As well as challenging the claims made by Piagetian theory, Goswami also drew attention to the role of meta-knowledge. This refers to the knowledge that reasoning using relational similarity, that is by analogy, is an appropriate way to solve the task. Even if children know the relations and are competent in applying the relational similarity constraint, they may still fail to look for relational similarities. Thus, if children are unaware of the goal of the task, that is they do not realise the significance of relational similarity, they will perform poorly. It would not necessarily be true, however, to say that failure was due to a lack of competence in applying the relational similarity constraint. Instead failure may be due to a misunderstanding of the nature of the task.

As has been previously argued, the tendency to reason successively with classical analogies (observed by Piagetian researchers), might have been due to a lack of knowledge about the task requirements, rather than an inability to use the relational similarity constraint. Research using simple problem analogies with young children have provided support for the claim that meta-knowledge has a significant effect on children's analogical reasoning skills.

For example, recent studies which have made the goal of the task explicit have resulted in successful performance from young children. Brown, Kane and Long (1989) provided one group of 7 year olds with explicit instructions concerning analogical mapping. They did this by telling the children a story about the solution to a problem which they had previously been unable to solve. They were then told that they could use the story to help them solve another more difficult problem which they had also previously been unable to solve. 46% of this group transferred the problem solution, compared with only 20% of a group of children who received the same stories but without any mapping instructions. However, although the experimental manipulation resulted in an increase in successful performance, the larger proportion of subjects (54%) still did not transfer the solution from the base to the target problems. Thus it would seem that although the receiving of specific task instructions aided transfer, this was not by itself sufficient to achieve total success.

The role of meta-knowledge was further investigated in the study discussed above. They encouraged 7 year old children to transfer problem solutions by giving them more than one analogy, as well as drawing their attention to the fact that the problem solutions were the same. Thus the children who had received the initial problem a second time were shown a series of pictures which demonstrated its solution. They were then given a third analogical problem, but with no prompts about its similarity to the other two problems. Their performance was contrasted with children who received the same three problems in the same sequence, but with no prompts about any of the problems. This time, the solution was successfully transferred to the third problem by 98% of the children in the experimental group, compared with 39% in the control group. The full experiment therefore revealed the results given in Table 2.1 below.

Table 2.1: Results from Brown, Kane and Long (1989)

	One analogy	Two analogies
Instructions	46%	98%
No instructions	20%	39%

Brown *et al* claimed that their results showed a 'learning to learn effect'. Highlighting similarity between analogies and providing the opportunity to map this similarity onto the target domain means that children will learn that the task goal is to apply the relational similarity constraint. This knowledge about the nature of the task helps them to focus on the relational structure of the base problem and map it across to the target.

This appreciation of the problem goal was shown to be important in a study by Brown and Kane (1988). They designed three pairs of analogical problems, where the relation which linked each pair of problems differed. Thus the two stories in Pair 1 were about a stacking relation, those in Pair 2 were about a pulling relation and those in Pair 3 were about a swinging relation.

3, 4 and 5 year old children in the experimental condition received all three pairs of problems, whereas the children in the control group received only the third pair. For all age groups, performance was greatly improved in the experimental groups. This shows that when children are given the opportunity, they can learn that problem solving using relational similarity is appropriate.

Focusing on structure is important.

Problem analogies have also been used to look at the role of the problem structure in the development of analogical reasoning. Brown *et al* (1986) showed that questioning children about the goal of the base story, a manipulation which was designed to help them focus on the relational structure of the solution, significantly aided transfer to the target problem.

An experiment carried out by Chen and Daehler (1989) investigated whether the ability to abstract a generalised solution schema, in which the relational structure was made explicit, would aid transfer. They gave 6 year old children two analogous stories, followed by an analogous problem. All the subjects were asked how the two base stories were similar. Their responses were scored for the degree of abstraction. Those children who talked about the structure of both stories independently and in terms of the actual objects in the stories scored lower than those who integrated both stories and who had abstracted the relational solution. All the children who had abstracted solution schemas were successful in mapping the solution to the target problem, whereas only 38% of the 'non-abstraction' responders had successfully transferred the solution.

These types of studies are providing us with some evidence which is relevant to the issues raised at the end of the review of the literature on adult analogical reasoning. These were:

1. The importance of the construction of the initial problem representation .

2. The induction of generalised schemas which can be held in long term memory.

The work discussed by Chen and Daehler is directly relevant. It does seem that children can abstract generalised schemas, at least from problems which have been very recently presented, and that these schemas have a significant facilitation effect on analogical mapping.

Brown *et al's* research demonstrates that highlighting the structure of the base problems facilitates performance. However, it does not directly address the issue of how the structural representation is constructed. Problem analogies are inherently ill defined and as such the crucial relations are embedded within a narrative which contains some irrelevant details. It would therefore be difficult to explicitly ask children to represent the relational structure of the problem, as they may not be aware of exactly what we are requiring them to do. Asking them questions about relational structure as we see it does mean that we can highlight its importance, but provides no insight into how children would represent relationships if they had not been cued to do so by the experimenter. We therefore need a task which falls between classical analogies, where the problem structure is explicitly presented to the child as part of the task, and problem analogies, where we often need to remind children about the importance of task relations by beginning to structure the task for them. A suitable analogical task would be one where the child needs to construct the relational task representation himself, but where the requirement to construct it is salient, such that no prompting by the experimenter is needed. This is an important point, and will be returned to at the end of this chapter and at the beginning of Chapter 3.

Before this, we must consider the current knowledge-based theories of the development of analogical reasoning ability, which have been put forward to counter Piaget's structural theory. It will be argued that these accounts also demonstrate a need to investigate the way in which children structure their knowledge, that is, their ability to construct relational task representations.

2.4.4 Knowledge based accounts of the development of analogical reasoning

The criticisms listed above, and the studies which have challenged Piagetian assumptions have led to an alternative view of the development of analogical reasoning. Researchers who have disagreed with the late emergence of analogical competence all share the belief that the development of analogical reasoning is knowledge based. Thus, in certain circumstances, children as young as 10 months old are able to apply the relational similarity constraint (Wagner *et al,* 1981).

Goswami (1992) has argued for an extreme version of the knowledge-based account. She states "the ability to recognise relational similarity may not develop at all. Children may be able to recognise relational similarity at any point in their development" (page 13). Nonetheless, current research (Gentner and Toupin, 1986; Brown, 1989) has still found that analogical reasoning improves with age. There are several subtly different theories which claim to account for this persistent age-related change, whilst still accepting that young children are competent in applying the relational similarity constraint, given optimum conditions. These theories focus on differing factors to account for age-related change in analogical performance.

1. Conceptual system development

There are several theorists which focus on the nature of the underlying conceptual system as a constraining factor in children's performance. Vosniadou (1989) claims that "what develops is not so much the ability to engage in analogical reasoning *per se* but, rather, the conceptual system upon which analogical reasoning must operate" (p.428) Thus mapping of an explanatory structure from a source to a target domain will occur when the structure is present in their source domain knowledge and is also consistent with their developing knowledge of the target domain. If the knowledge base is still immature and does not contain sufficient constraining features, then inappropriate transfer may occur, often where surface features are inappropriately mapped across domains. Adults also engage in this type of reasoning when their domain knowledge is not structured with the necessary relational information, as shown by the research conducted by Chi, Feltovich and Glaser (1981) on novice and expert physicists.

In a similar vein, Gentner (1989) puts forward evidence (Gentner and Toupin, 1986, Gentner, 1988) for a developmental shift in analogical reasoning. She also views the ability to reason about relational similarity as the hallmark of this ability. However, her theory differs from Vosniadou's in that she states that young children show a propensity to reason using surface similarities, with the shift to relational reasoning occurring with age. Nonetheless, it must be recognised that this account is not necessarily arguing for a genuine lack of competence in young children. Gentner has also suggested that domain novices rely on surface similarities (object matching) and has said that this reliance in both young children and novices may well come about because of a lack of relevant domain knowledge. Of course there may be analogies in which base and target domains share surface similarities as well as relational similarities. Gentner (1989) has therefore suggested that children will find this type of analogical problem (termed a 'near' analogy) easier than those which only share relational similarity (termed 'far' analogies). She cites evidence from Gentner and Toupin's study (1986) in support of this claim. Here, 6 year olds could only solve analogies which shared surface and relational similarities, whereas 9 year olds could solve analogies which only shared relational similarities.

In order to determine the cause of the shift demonstrated by Gentner and colleagues, Goswami (1991) reviewed the literature and concluded that because the studies did not test the subject's domain knowledge, there was no evidence to support the increasing cognitive competence hypothesis. Recent studies which have ensured an understanding of the relevant domain knowledge have obtained successful performance in children as young as 3 years. Thus it would seem that the relational shift demonstrated by Gentner could be due to developing knowledge rather than a change in analogical competence *per se*. When children and adults are unable to reason using relational similarity because of a lack of appropriate relational knowledge, they fall back on surface similarities.

Vosniadou's account, discussed overleaf, shares many features with Goswami's (1992) work. Both are concerned with the ability to recognise and use relational similarity. Goswami states that "Children may be able to recognise relational similarity at any point in development......Their ability to do so will depend on the difficulty of the relations on which the analogy is based" (p.13). This is very similar to

Vosniadou's argument - there is a need for relations to be worked out and represented in a conceptual system so that their similarity becomes apparent. The difference between these two theories seems to be that Goswami focuses on the recognition or discovery of the relevant relationships, whereas Vosniadou concentrates on the development of the representation of those relationships. However, as Goswami points out, recognition will depend on the degree of difficulty of the specific relation involved. The depth of knowledge contained in the conceptual system will also depend on the degree of difficulty involved, so it seems that the two theories differ only in focus, rather than in any essential feature.

We can draw a common implication from these theories - children need to possess relational domain knowledge **in an appropriate structure** such that relational similarities can be noticed and therefore used.

2. Information processing capacity development.

A rather different way of looking at the development of analogical reasoning concerns constraints imposed by limited information processing capacity development. Halford (1992) claims that children pass through four stages of structure mapping ability, determined by a maturational increase in the amount of information they are able to process. Each of these stages refers to the number of relations which the child can process in any one time. A more complex task will impose a higher processing load due to more relations being involved. Thus the theory hinges on the number of relations involved rather than an increase in the analogical mechanism *per se*.

In particular, Halford claims that children below the age of 5 years can only reason about single relations, and before the age of about 2 years, they cannot reason about relations at all. However, evidence exists from other studies which suggests that very young children are capable of relational mapping. Willatts (1991) gave infants problem solving tasks in which they had to plan a series of actions which would result in a desired goal. In order to reach a toy they had to remove a barrier between themselves and the toy and then pull on a cloth to bring a container towards them which contained the toy. The subjects had previously practised both actions separately, but in order to plan retrieval of the toy, they had to relate them together, and hold them both in memory at the same time. This can therefore be construed as evidence that babies can reason about one type of relation at a time. Of course, these combinations are between objects which exist in the real world, and not between mental representations. This leads us to a discussion concerning the nature of base and target domains for analogical reasoning. However, this is not explicitly referred to in most of the approaches we have already discussed.

In both the classical analogy and the problem solving studies, mapping has been between base and target domains which exist in some sense in the real world. Halford, however, argues that the reason some mapping tasks can be solved by very young children is because "it is not necessary to map from one mental representation to another, as occurs in analogy" (page 201). This of course adds weight to his working memory capacity argument, as presumably more complex **mental** representations will occupy more space in working memory.

His theory of analogical reasoning describes the mapping process as being between the base domain of an abstract reasoning structure (c.f. Holyoak's pragmatic reasoning schema, 1985) and the target domain (usually the problem) As he has previously described the reasoning schema as being retrieved from semantic memory, then this description implies a mapping between two mental representations (i.e. the schema and the problem solution). It is difficult to see where we can make a clear-cut distinction between mental and non-mental representations, however, as we have no way of saying how task representations are shared between the environment and the problem-solver's internal resources. It could be that they rely on real-world representations of the base domain for both classical and problem analogies. In the former type of problem, the C term is also represented in the real world, so the only element which has to be based on a mental representation is the actual relation between terms and the application of this to the C term. For the target domain in problem analogies, the relation between the relevant task features is often already known. What needs to be represented mentally is the mapping of the appropriate relation from the base domain into the target. It seems therefore that Halford's definition of analogical reasoning is doubtful for two interconnected reasons:

a. If we accept the core concept of analogical reasoning as being solely the ability to map relational similarity from base to target domains, then it does not follow necessarily that distinctions made between mental representations and real-world tasks have any significance with respect to a theory of analogical reasoning *per se*.

b. As pointed out by Goswami (1992) in her commentary on Halford's thesis, we have no way of knowing whether or not children map between real world tasks by constructing mental representations of both base and target domains.

What is more appealing about Halford's work is the emphasis on abstract reasoning schemas, which he argues arise from the child's everyday experience. In a manner similar to that previously described, children abstract generalised relational information which can be used as base domains for analogical problems. He thus argues that much of reasoning is carried out through analogical mapping from these types of schemas.

However, he does not really address the issue of how the structuring of base and target domains occurs. He appears to believe that schemas are represented as complete structures which are retrieved intact and mapped across in their entirety. However, it seems doubtful whether this is the case. Children often build up their knowledge of the world gradually and as such, it seems likely that schemas will be gradually constructed, with potential for errors along the way.

If this is the case, then we might expect these schemas to be initially rather fragile, with the potentiality for interference from irrelevant features in the target domains. Again, we arrive at the importance of the way in which the child constructs the initial task representation containing the important relational information. This representation could be an abstracted generalised ordering schema, or a specific representation of a single base domain.

An alternative information-processing account put forward by Sternberg and colleagues focuses on the sub-processes used in analogical reasoning. Whilst we have already discussed this methodology to look at analogical competence in adults, it has also been used to address developmental issues. By breaking down the ability into various component skills, researchers have been able to investigate whether developmental differences in analogical performance are caused by the differences in the various components.

As previously discussed, Sternberg claimed that six different components were involved. These are:

a. Encoding the analogy
b. Inferring the relation between the A and B terms
c. Mapping or discovering the relation between the A and C terms
d. Applying this relation to the B term to generate a solution
e. Justifying the goodness of the match
f. Responding

Sternberg and Rifkin (1979) tested the componential theory by giving 8, 10 and 12 year old children classical analogies to solve. These were based on drawings of people who either varied on the basis of external features (hats, suits, footwear etc.) or who varied due to integral features such as height , weight, sex etc. They found that use of the different components did not differ qualitatively with age, rather that children just got faster. However, they found no evidence of mapping (item c) by the 8 year olds. They concluded that this provided evidence for the Piagetian argument in that it involves recognition of higher or second order relations. However, as previously discussed, the use of mapping is optional. Item d, the application of the relation, corresponds most closely to the children's ability to apply the relational similarity constraint. If mapping is not necessary for reasoning by analogy, then failure to use this component cannot be cited as evidence for failure to reason using relational similarity. Sternberg and Nigro (1980) also tested the hypothesis that children were reasoning by association. They used verbal classical analogies, but again did not ensure that the subjects understood the relations involved in the analogies. Thus the fact that the children problem solved using associative reasoning rather than relational similarity cannot be used as evidence of a lack of competence in analogical reasoning *per se*.

Whilst componential accounts such as Sternberg's have provided us with a useful methodology, again most of the research carried out in this area can be criticised because the children were not pre-tested for the appropriate domain knowledge. Also, as with the same approaches to analogical reasoning in adults, the description has been concerned solely with the solving of classical analogies. Goswami (1992) has pointed out that both the Inference and Application components in Sternberg's description refer to the ability to work out relations, that is to their relational knowledge. The Application component also concerns children's understanding about the similarity of relations. Because of this, it is not possible to use Sternberg's componential analysis to isolate children's understanding of the relational similarity constraint.

Sternberg's theory is therefore limited in its application. Its concentration on classical analogies means that the initial representation of the relational structure of the task was not considered. On the other hand, Halford's emphasis on the structure of generalised schemas highlights the significance of the representation of relational information.

3. Metaconceptual Development

Brown (1989) has put forward a 'coherent knowledge hypothesis'. This argues that not all knowledge is equal, and that one of the reasons that older children are more successful in unassisted analogies is because their knowledge is better structured. Citing the experiments which manipulated the instructions given to the subjects (Brown and Kane, 1988) and those which provided assistance with structuring the knowledge contained in the base analogy (Brown, Kane and Echols, 1986), she concludes that the better performance demonstrated by the older children is due to two reasons:

1. They possess more efficient general strategies for 'learning to learn', in that they do not need instructional help.

2. These strategies mean that the child has a better understanding of the form or structure in which their knowledge needs to be represented, such that mapping is facilitated.

It is difficult to directly test metaconceptual understanding. However, it seems from the studies referred to above that the way in which children build up integrated representations of the knowledge which they acquire will have an effect on their ability to map analogical relations across domains.

2.5 THE IMPORTANCE OF THE CONSTRUCTION OF RELATIONAL TASK REPRESENTATIONS

We now have considerable evidence that Piaget's claim that children do not reason analogically until the formal operational stage (i.e. around the age of 11 years) is wrong. It seems that when children possess the necessary domain information and it is represented such that the relational structure is highlighted, they can recognise and use relational similarity between domains. All the current developmental approaches, whilst highlighting the crucial importance of the possession of relational knowledge, have therefore also emphasised the importance of how it is structured. The possible exception to this is Sternberg's theory, because he has restricted his account to classical analogies.

Work by Halford has provided us with a way of thinking about how constraints internal to the individual might affect their ability to map relational structures. This thesis will investigate constraints inherent in a task, i.e. those external to the individual, which affect the way in which relational information is represented. This is consistent with Keane's call for further research in the area of pragmatic constraints, discussed earlier in this chapter. It is suggested that this will throw some light on the reasons why some analogical problems are still difficult for children, even though they are able to reason using the relational similarity constraint.

What we are particularly concerned with is children's ability to construct an initial task representation which correctly represents all the significant relational information. This is because many analogical tasks in the real world are concerned with complex domains, with more than one potential item of information which could be mapped. Because of this, it will not be appropriate to use classical analogies, as the domain information has already been structured as part of the task. On the other hand, problem analogies are too poorly defined, such that it is difficult to look at the actual

construction of the task representation. This is because in order to ask the child about the relational structure of the problem, we would have to ask questions which begin to structure it for them.

In order to deal with these constraints, we have decided to select a task in which the structure is explicit, but is not given as part of the task. This way, we can examine the influence of task constraints on the integration of separate pieces of information into one systematic relational representation, such that this relation can be used for analogical mapping. The next chapter deals with the selection of the task for the research which is reported in the following chapters.

CHAPTER 3 : SERIES PROBLEMS AS AN ANALOGICAL REASONING TASK

3.1 CHAPTER OVERVIEW

Chapter 2 identified a need to investigate the role of structural task representations in the development of analogical reasoning. It was suggested that the representation of a domain or problem at the structural level (that is, a representation which depicts salient items and the relationships between them in an integrated manner) would have a facilitatory effect on analogical mapping. Furthermore, it was hypothesised that the early competence in analogical reasoning demonstrated by Goswami and others can often be masked by an inability to represent base and target domains at a common structural level. Such an inability will prevent analogical mapping and so result in poor performance, even though competence in recognising and using relational similarity is firmly in place. In other words, it was suggested that the representation of problem domains with respect to entities and their relations to each other in an integrated system is a necessary precursor to analogical reasoning. Thus the rest of this thesis will explore some of the factors which affect the construction of these types of structural task representations. This will enable us to understand some of the difficulties children encounter when mapping objects and relationships across base and target domains.

It is necessary, therefore, to identify a simple task which would enable a rigorous and systematic investigation of the factors which constrain the internal description of task objects and their relationships to each other in a single, integrated representation. This chapter addresses the choice of task for the subsequent group of experiments, which are then reported in Chapters 4 to 9.

3.2 CHOICE OF TASK

The main aims of the research are as follows:

1. To investigate some of the factors which affect how young children combine relations and objects into a single integrated systematic representation.

2. To investigate how this integrated representation can used to map between problem isomorphs in analogical reasoning tasks.

In order to satisfy these aims, a list of task requirements was constructed.

The task should be:

- simple, such that it is concerned **solely** with objects and their relationships with each other i.e. with no irrelevant details which might inhibit the representation of structural features

- commonly solved by the construction of an integrated structural representation i.e. one which represents domain objects and their relationships

- within the ability range of young children

- able to be used as a base domain in analogical reasoning tasks

- applicable to a variety of domains (both concrete and abstract)

It was decided that the group of tasks known as series problems met the prerequisites listed above. In order to demonstrate the reasons for this, it is necessary to first describe the task and then to give a brief review of the existing literature. Each requirement will then be discussed with respect to the task.

3.3 SERIES PROBLEMS - A DEFINITION

Series problems are based on direct comparisons between two items contained within one premise. Two or more premises are presented, and the task is to make inferences about those items which are not directly related in any one premise:

John is taller than Peter	(premise 1)
Peter is taller than Robert	(premise 2)
Who is taller, John or Robert?	(inferential question)

There is some considerable literature (e.g. Hunter, 1957, Breslow, 1981) concerning this type of task, both in the adult problem solving literature and in that of cognitive development. Much of the research which has addressed the development of children's ability to perform this task has been concerned with whether successful performance requires formal logical competence. This tradition of research began with Piaget's stage theory of cognitive development (1921) He viewed series problems as important because they enabled him to investigate the logical concept of transitivity. A transitive relationship is one in which the relationship between 'John' and 'Peter' and between 'Peter' and 'Robert' (see above) is such that 'John' is necessarily similarly related to 'Robert'. In a further example, the relation 'more intelligent than' is transitive, because it supports valid deductions of the form:

A is more intelligent than B (premise 1)

B is more intelligent than C (premise 2)

Therefore, A is more intelligent than C (conclusion)

N.B. By convention, task items or elements are referred to as A, B, C etc.

Non transitive relations, on the other hand, do not support a similar relation between A and C. For example, the relationship 'likes' is non transitive:

A likes B (premise 1)

B likes C (premise 2)

No valid conclusion

3.4 PIAGETIAN RESEARCH CONCERNING TRANSITIVE REASONING

For Piaget and colleagues (1921, 1970) transitive reasoning* was important as it was viewed as a central element of the concrete operational child's thought. They maintained that a child's concept of relations, causality, time etc. progresses through a series of developmental stages, each one having a qualitatively different structure. This means that the child's conceptions in a certain area of relations (e.g. transitivity) were reorganised as the child attained a different developmental stage.

Inhelder and Piaget (1964) looked at the development of transitive inference from the point of view of the child's developing understanding of different kinds of

---

* Because the literature has been fundamentally concerned with the understanding of transitivity, the problems described above are commonly referred to as transitive inference problems. This thesis continues to refer to them as series problems as it's aim is not concerned with the understanding of formal logical competence i.e. transitivity

relationships. According to them, the child begins with only a categorical understanding of different kinds of relations. For example, taller and shorter are taken to be mutually exclusive, such that an object cannot at the same time be taller than some objects and shorter than others. Once children reach the stage of relativistic understanding of relations (in the concrete operational stage) they can understand that a term can be both tall in relation to A and short in relation to C and so they can deduce the relationship between A and C.

According to Piaget (1972), concrete operational thought is characterised by 8 'groupings'. Transitivity is a property of half of these, those four which are concerned with relations (the other four are concerned with logical classes). A grouping consists of a set of elements (a 'field') and an 'operation' which may be performed between pairs of elements of the field. Grouping V describes the addition of asymmetric relations. These are relations where, when A is related to B this does not imply that B is related to A in a similar relationship. For example, Jane is bigger than Mary is asymmetric as this statement does not mean that Mary is bigger than Jane. However, Jane is married to entails that Bob is married to Jane and so this relation is symmetric. Thus the 'field' of Grouping V is composed of a series of asymmetric relations between two items. Each pair of items which are in relation to each other is known as an 'element'. Thus Jane is bigger than Mary is one element, as the two items 'Jane' and 'Mary' are in relation to each other. When these relations are also transitive, a relationship which exists in two elements in the field which also have one item in common means that the two elements can be joined together. This joining is such that a deduction can be made about the items which are not common to both elements. The joining of elements is termed addition, and is the 'operation' which is performed on pairs of elements in Grouping V. For example, using the transitive relation 'bigger than':

Jane is bigger than Mary	(element 1)
Mary is bigger than Janet	(element 2)
Thus, Jane is bigger than Janet	(as 'bigger than' and 'Mary' are common to both elements)

This addition can be represented symbolically as follows: $(A>B) + (B>C) = (A>C)$

Also, either one or both of the elements listed above can be inverted e.g. Jane is bigger than Mary becomes Mary is smaller than Jane. Transitive inference between only direct relations is known as isotropic (as in the example above, where 'bigger than' is the relation used in each 'element'), whereas that between both direct and inverse relations is heterotropic. Within the domain of relational thought, the ability to co-ordinate inverse relations is also considered to be an essential part of the concrete operational stage. This action may take two forms:

1. The translation of a relation (Jane is bigger than Mary) into its inverse (Mary is smaller than Jane) and back again.

2. The co-ordination of inverse relations which are not inverses of each other around the same term. For example:

Jane is bigger than Mary	(element 1)
Jane is smaller than Sue	(element 2)

Although the relations in the two elements are inverses of each other, the actual elements are not as Jane is the only common item. Thus as well as understanding that

Jane is bigger than Mary entails that Mary is smaller than Jane, it is also necessary to understand that Jane can be bigger than Mary and at the same time be smaller than Sue. Item 2 is especially important for a complete conception of the addition of transitive relations. This is because an understanding of the relative nature of relations is essential for the co-ordination of premise relations around their common term.

Criticisms of Piagetian assumptions

Later research has concentrated on methodological issues concerning transitivity. Various factors have been discussed which could lead either to false positive assessment (subjects are judged to be reasoning transitively when they are not) or to false negative assessment (subjects are judged to be failing to reason transitively when in fact they are)

Reasons for false positive assessment have centred around subjects using alternative methods, such as attaching absolute relational markers to the items (Smedslund, 1963). Thus, in the case of Jane is bigger than Mary, Jane would be designated as big. When asked a question about Jane and Janet, the subject would therefore conclude that Jane had to be bigger than Janet without considering the second premise (Janet is bigger than Mary). These concerns have been easily overcome, however. Smedslund used more than three items, e.g. Jane, Mary, Janet, Sue, for each problem. This ensured that the mid-items, Mary and Janet, were both the larger and smaller items in different premises, e.g. for the item Mary:

Jane is bigger than Mary

Mary is bigger than Janet.

Absolute relational markers can therefore not be associated with 'Mary' and so any inferential questions involving this mid-item, e.g. Who is bigger, Mary or Sue?, requires that premises are combined in order for the correct answer to be inferred.

Several researchers have suggested that one reason for false negative assessment is the difficulty young children experience in memorising task information (e.g. Bryant and Trabasso, 1971). In order to control for this Trabasso (1975) trained subjects to criterion on premise relations and then questioned them on all possible pairs of items (both premise and inferential comparisons). Direct perceptual comparisons were controlled for by using different coloured sticks as premise items and presenting them through a viewing box such that their relative length could not be seen. They were referred to only by their colour, so that no perceptual cues were available.

Subjects as young as 4 years were successful in answering the inferential questions correctly, provided that they received sufficient training. Age differences were due only to the length of time it took for the subjects to learn the premises to a criterion standard, and to the greater difficulty which the youngest subjects (aged 4 to 5 years) had in training with verbal rather than visual feedback. Trabasso therefore argued that previous failures in transitivity tasks by children younger than 7 or 8 years was not due to a lack of conceptual understanding of transitivity, as Piaget claimed. Instead, it was due to the high memory demands which these tasks placed on the subjects, both in initially remembering the premises and in making use of feedback.

Trabasso, Riley and Wilson (1975) continued the research concerning young children's successful performance by examining the relative ease with which subjects made various item comparisons in the 'coloured sticks' task described above. The main finding from these results was termed the distance effect. This concerns the number of deductions required to make the item comparisons. Thus questions concerning the comparisons made in the original premises require no deductions, comparisons made between the first and third items or the second and fourth items require 1 deduction and a comparison between the first and fourth items require 2 deductions. Consider the task below:

> The red stick is longer than the blue stick
>
> The blue stick is longer than the yellow stick
>
> The yellow stick is longer than the green stick
>
> The green stick is longer than the orange stick

Questions comparing the red stick with the blue stick, the blue stick with the yellow stick, the yellow stick with the green stick or the green stick with the orange stick require no transitive deductions as they are concerned with information given in the premises. Questions comparing the red stick with the yellow stick, the blue stick with the green stick or the yellow stick with the orange stick require one transitive deduction as only two premises need to be considered. Comparisons between the red stick and the green stick or the blue stick and the orange stick involve co-ordinating information contained in three premises, resulting in two deductions being made.
The results showed that the greater the number of deductions required to make the comparison, then the faster it was solved. If subjects were encoding the premises separately, and only combining them inferentially when asked to make specific comparisons, one might expect that an increase in the number of deductions required would result in slowing down of the response rate.

Trabasso *at al* argued that the distance effect provides evidence for the fact that subjects integrate the premises into a single ordered representation. They then used this to answer the questions posed by directly making comparisons between individual items in the ordered representation. This eliminated the need for inference *per se*. Thus pairs of sticks which are further apart in the integrated representation will be more easily picked out than those which are closer together and therefore comparison speeds in the former will be shorter.

If we accept Trabasso's argument, then it follows that subjects are not reasoning by transitive inference in the test phase at all, but are making direct comparisons from a fully integrated representation. However, Trabasso maintained that subjects do use transitive inference in order to form the array in the training phase. That is, 'the red stick is longer than the blue stick' is combined with 'the blue stick is longer than the yellow stick' to form

RED BLUE YELLOW (where the left side of the ordering denotes 'LONG')

Nonetheless, it remains unproved whether this use of transitive reasoning demonstrates any explicit conceptual understanding of formal logical procedures, as described in Piaget's stage theory.

3.5 THE CURRENT POSITION CONCERNING TRANSITIVE INFERENCE

The argument therefore, appears to be whether or not young children (typically below the age of about 9 years - the concrete operational stage) are capable of reasoning using transitive inference. Piaget's definition is based on theories of formal logic. If we accept this definition, then the literature to date has not clearly demonstrated a true conceptual understanding of transitivity in young children (below the age of about 9 years). It seems probable that the children Trabasso *et al* tested were solving the problems 'nontransitively', in as much that premises were combined in an additive fashion during the training phase to form a single integrated representation. The additive nature of the process, where each new item is joined onto those which have already been ordered, means that it is perfectly possible to produce a correctly ordered single representation without an explicit understanding of what a transitive relation entails. By using the correctly ordered internal item array, direct comparisons between items could be made during the test phase, and so the use of inferencing was not required. As Breslow (1981) has pointed out, it is not surprising that children will construct and use such a representation in the tasks, as "direct comparisons are clearly more convenient and efficient than inferences" (pg. 347).

The debate (which is still unresolved) seems to have now shifted from the age at which children can reason by transitive inference to considering what the 'transitive task' is actually testing. For example, consider the task below

The red stick is longer than the blue stick	(premise 1)
The blue stick is longer than the yellow stick	(premise 2)
The yellow stick is longer than the green stick	(premise 3)
The green stick is longer than the orange stick	(premise 4)

In order for the children's performance to unequivocally demonstrate a full conceptual understanding of transitivity in the Piagetian sense, we would expect to see the premises above encoded separately in memory during the training phase of the experiment. The comparisons required of the subjects in the test phase would mean that the relevant individual premises were retrieved and the appropriate inferences made.

In the current example, a comparison between the red stick and the green stick would necessitate retrieving premises 1-3 and using the relations given in these to infer the required relationship. In contrast to this, the construction of a single ordered representation means that as the subjects were learning the premises, they would gradually integrate each premise into an integrated ordering as it was presented, in the following manner:

presentation of premise 1

RED BLUE (initial representation, left side denotes 'long')

presentation of premise 2

RED BLUE YELLOW (updated representation)

presentation of premise 3

RED BLUE YELLOW GREEN (updated representation)

presentation of premise 4

RED BLUE YELLOW GREEN ORANGE (updated representation)

There is now a consensus that children can and do solve these tasks by constructing an single ordered representation during the training phase (Trabasso 1975, Trabasso and Riley 1975, Breslow 1981). The point at issue is whether or not this still constitutes an adequate test of the conceptual understanding of transitive relations, as described in Piagetian structural theory. It could be that children are demonstrating appropriate conceptual knowledge of transitivity as they are constructing the integrated representation. However, it is possible to build such a representation in the absence of this explicit knowledge. Each novel item can simply be added in accordance with the relationship described in the new premise without appreciating any necessary inferences which could be made directly from the premise information, due to the transitive nature of the relationship.

For the purposes of this thesis, the issue of a 'transitivity test' is not relevant. The important point is that children can and do solve series problems by the construction of an single integrated array i.e. a structural representation. This behaviour is similar to that of adult novices, although there is also evidence of the later emergence of task dependent strategies (Wood, 1969). He demonstrated that when adult subjects are familiar with transitive reasoning problems, they begin to focus on only the relevant premises which they need in order to make specific comparisons.

3.6 THE NATURE OF THE ORDERED REPRESENTATION

In the examples of Trabasso's research discussed above, the array has been ordered horizontally, such that the left side of the ordering denotes 'long'. However, it must be emphasised that the original research papers contained no details as to the dimension of the ordered array, simply referring to it as a 'single linear array'. The decision to represent the array above in this manner was made solely for illustrative reasons. In fact there is very little discussion in the literature as to it's actual nature. However, work by DeSoto, London and Handel in 1965 is of relevance here. They suggested, like many other researchers, that the crucial step in the solving of series problems is the combination of premise information into a unitary representation. They claimed that this unitary representation consists of a visual image of a vertical or horizontal array. More importantly, they claimed that relationships such as 'better-worse' have a fixed tie to the vertical axis, whereas other relationships (e.g. lighter-darker) are not spatially tied. This theory was expanded on by Handel, DeSoto and London, 1968, who claimed that relationships which have a linguistic linkage to a vertical dimension (for example 'better-worse' employs terms such as 'upper crust' or 'top dog') are tied to a vertical ordering array whereas relationships such as 'farther-nearer' do not have consistent spatial assignments.

Halford, 1992, when discussing series problems as a type of analogical mapping, is unclear as to the dimension of the generalised ordering schema. He states that this is usually 'left to right' or 'top to bottom', but seems to make no distinction between the two. However, his diagrammatic representation of the mapping process, reproduced in Fig. 3.1 below, is labelled using vertical relationships ('above'), but the schema is represented graphically as being in the horizontal dimension.

Fig. 3.1: Series problem mapped into ordering schema (Halford, 1992)

	above		above	
base schema	**TOP** ________	**MIDDLE** ________	**BOTTOM**	
target schema	**JOHN** ________	**PETER** ____________	**TOM**	
	taller than		taller than	

From Halford (1992)

It would therefore seem that there is no real consensus in the literature concerning the dimension of the internal array. Although work by DeSoto and colleagues suggests a preference for vertical orderings, Halford has not taken up this point, as he states that either dimension could be used.

Why do series problems fit with the task requirements listed on Page 58?

1. Objects and their relationships.

These are very clearly defined problems, with no objects which are irrelevant to the task structure which might mislead the subjects. It was felt important to focus in on the representation of relational systematicity, initially without introducing other features which might make this task more difficult and thus confound the data. For this reason, more real world, educationally relevant problems were felt to be inappropriate at this stage of research, as they are often couched in narrative forms which inevitably introduce contextual detail which is unrelated to the relational structure of the task. It seems likely that they might therefore result in problems when children are focusing on the structuring of new knowledge.

2. Solvable by the construction of a structural task representation.

As previously discussed, there is ample evidence in the literature to support the view that children and adult novices solve series problems by the construction of an internal ordered array.

3. Within the ability range of young children

Several researchers have used this type of task to investigate young children's problem solving. Obviously written premise information could be a problem, depending on reading ability. In order to avoid this problem, researchers have introduced either concrete objects or drawings of objects. Pears and Bryant (1990) achieved successful performance with 4- 5 year old children when making comparisons about the spatial position of coloured bricks in a tower (see below for a full description) In a study using comparisons such as height or happiness, Riley (1976) found that 7-8 year old children were successful when using drawings of children's faces.

4. Usability as a base domain in analogical reasoning tasks

Halford (1992) has explained children's success in transitive inference tasks by stating that they map the relationship inherent in the problem onto an ordering schema with which they are familiar. Everyday interaction with their social environment provides experience in ordering people, toys and objects according to height or size etc. Halford suggests that this experience is generalised into an ordering schema, which the child can easily internalise due to the extent of its presence in day to day living. This schema, together with its mapping onto a series problem, is represented diagrammatically in Fig. 3.1 on page 70.

In other words, Halford believes that series problems are solved by analogy to very commonplace orderings which the child learns at an early age. The frequency of occurrence of these orderings means that a generalised ordering pattern is internalised by the child. This can be used as a base domain for the purposes of analogical mapping. This description is very similar to Cheng and Holyoak's (1985) work concerning the use of pragmatic reasoning schemas in analogical reasoning. They define these as abstract knowledge structures induced from everyday experience. Thus the arranging of premise elements into an ordered set amounts to mapping them into a pragmatic reasoning schema, in that it requires using a previously internalised general schema to reason in a new domain which has the same relational structure i.e. is structurally isomorphic.

5. Ease of adaptability to a variety of domains

One of the reasons for investigating the building of structural representations is to consider their importance to the analogical mapping process. Therefore it is desirable that the task selected for study will be able to be used in a variety of domains, both concrete and abstract, as the ease of mapping of the base structure may be differentially affected by the level of abstraction (Gentner, 1989). There are many types of relations, ranging from spatial ('above') to abstract ('happier than') which can be used as a basis for series problems. Spatial relations such as 'above', because they share the surface features of the base schema represented in Fig. 3.1 above, are 'concrete' analogies, whereas relations such as 'happier than' are 'abstract' analogies to the base domain because they share only the deep, relational structure.

For the reasons listed above, it was decided to use series problems to investigate the construction of structural task representations. The following chapter describes the first two studies, which were exploratory in nature. This is followed by Chapters 5 to 9, describing a further eleven experiments in which aspects of the construction of relational representations are systematically investigated.

CHAPTER 4 : THE ROLE OF STRUCTURAL TASK REPRESENTATIONS

This chapter describes two preliminary studies which begin to investigate the role of structural task representations in the development of analogical reasoning, using series problems as a domain. The first of these was carried out in order to ascertain the appropriate age groups and task presentation which it was felt would identify the richest area for the development of further studies. The results obtained in this experiment identified some unforeseen difficulties. The second study was therefore a replication of a previously reported experiment by Pears and Bryant, and was carried out in order to identify a possible source of these difficulties.

4.1 EXPERIMENT 1 - THE DEVELOPMENT OF THE ABILITY TO REASON ABOUT SERIES PROBLEMS IN A SPATIAL DOMAIN

Background

For the reasons listed in Chapter 3, it was decided to use series problems to investigate the construction of structural task representations. One of the overall aims of this research is to ascertain whether some of the apparent age differences in analogical reasoning tasks can be explained by age differences in the ability to represent a task at the appropriate structural level. In view of this the upper limit of the age range chosen for study was fixed at 9 years of age. This is because from the age of 10 to 11 years (the concrete operational stage) analogical competence is no longer questioned by any of the competing theoretical accounts. 5 years of age (the commencement of compulsory education) was selected as the lower age limit for study. This was because the tasks required a level of concentration which would be problematic for preschoolers. An informal pilot study carried out with 3 to 4 year old children showed that they had very little interest in carrying out series problem tasks. The choice of the young age group posed potential problems concerning a suitable type of series problem. Obviously, any task involving reading would not be possible with 4 to 5 year olds. Trabasso's paradigm using coloured sticks of differing lengths was considered, but this involved an extensive training phase in which the individual premises had to be memorised. Again, an informal pilot study at a local nursery school showed that 4 to 5 year olds were not able to learn the premise information to criterion on a task similar to that used by Trabasso *et al,* though whether this was due to a lack of motivation or memory capacity was unclear.

One of the studies which obtained success with 4 to 5 year old children solving series problems was Pears and Bryant (1992). They considered the problem of young children's poor memory span and suggested that there were three ways to overcome this. The first of these was to give children plenty of prior experience with the premises, so that they have the opportunity to learn them thoroughly. There are, however, several objections to this.

The first of these, borne out by the pilot study referred to above, is that young children find memorising the task very difficult (Kallio, 1982). Some researchers have presented the premises serially:

A is bigger than B, B is bigger than C, C is bigger than D, D is bigger than E.

This usually results in eventual success, but this could be due to simply remembering the order of the premises rather than an understanding of the relationships which hold between individual items (Adams, 1978). In the studies which rely on memorising the premises, several of the subjects do not reach criterion in the training phase and so are not tested. It is assumed that these children simply have too short a memory span. Halford and Kelly (1984) have pointed out, however, that their failure could also be due to an inability to understand the information contained in the premises. Thus the exclusion of these children could mean that an artificially high success rate is obtained.

The second solution was an attempt to deal with the objections to the solution discussed above. Studies have been carried out which attempt to lessen or remove the memory load by providing memory aids which symbolise the premise relationships (usually concerned with size) by using spatial position. For example, Halford (1984) used a pegboard with pairs of coloured pegs. The children were required to make transitive judgements about the length of sticks which were the same colour as the pegs. The pegboard remained present throughout the experiment, with the left-hand peg in each pair always being used to represent the longer of the two sticks in each premise. This study showed demonstrated a very low level of success. As Pears and Bryant have commented, this could stem from a difficulty in using one type of relationship to represent another.

Their solution to this was to use spatial relationships, giving premise information using real objects. The spatial relation used was 'higher than' and the premise information was given by means of paired towers of two different coloured bricks. The child is asked to construct a tower of five bricks, using the information given in the premise towers to work out the correct position of each colour. Before the children built the tower, they were asked inferential questions about the relative positions of pairs of bricks. The spatial relation chosen meant that information from two premise towers had to be combined in order to reach the correct answer. This is very different from previous studies, which commonly use length as the task relation. This means that the actual length of individual items cannot be available to the subjects during the test phase, as they would be able to answer any questions by making a direct comparison between the sticks. Thus the use of a spatial relation removes much of the memory load required, as premise information can be visible throughout the study. It also does not require the introduction of any symbolic memory aid, as information can be given using the actual task relation.

Description of Pears and Bryant's study

In view of the above, it was decided to use a similar methodology to that of Pears and Bryant for the current experiment. Their study is therefore explained in detail below, and Figure 4.1 overleaf also gives a diagrammatic representation of the experimental phase. This phase was preceded by an introduction where the children were made familiar with the task materials and requirements. The experimental phase was then split into four sections.

Note The participants were given three types of problem. These differed only in the number of items (bricks) used in the problem - either two, three or four.

Section 1

The premise information was presented to the subject in the form of small premise towers i.e. towers consisting of two different coloured bricks. Thus towers of four items would have three premise towers, towers of five would have four premise towers and towers of six would have five premise towers. These were presented in random order, so that the children could not solve the problem by reading off the order of the colours as they were presented to them.

Section 2

The children were then given duplicates of the different coloured bricks used in the premise towers i.e. four bricks for the 'tower of four' etc.

Section 3

The children were then asked inferential questions concerning the order of colours in a tower which they would build if they preserved the relationships displayed in the premise towers. The questions were always about non-original pairings, for example, using the orderings represented in Fig 4.1 for the 'tower of four':

'Which will be higher, red or blue ?'

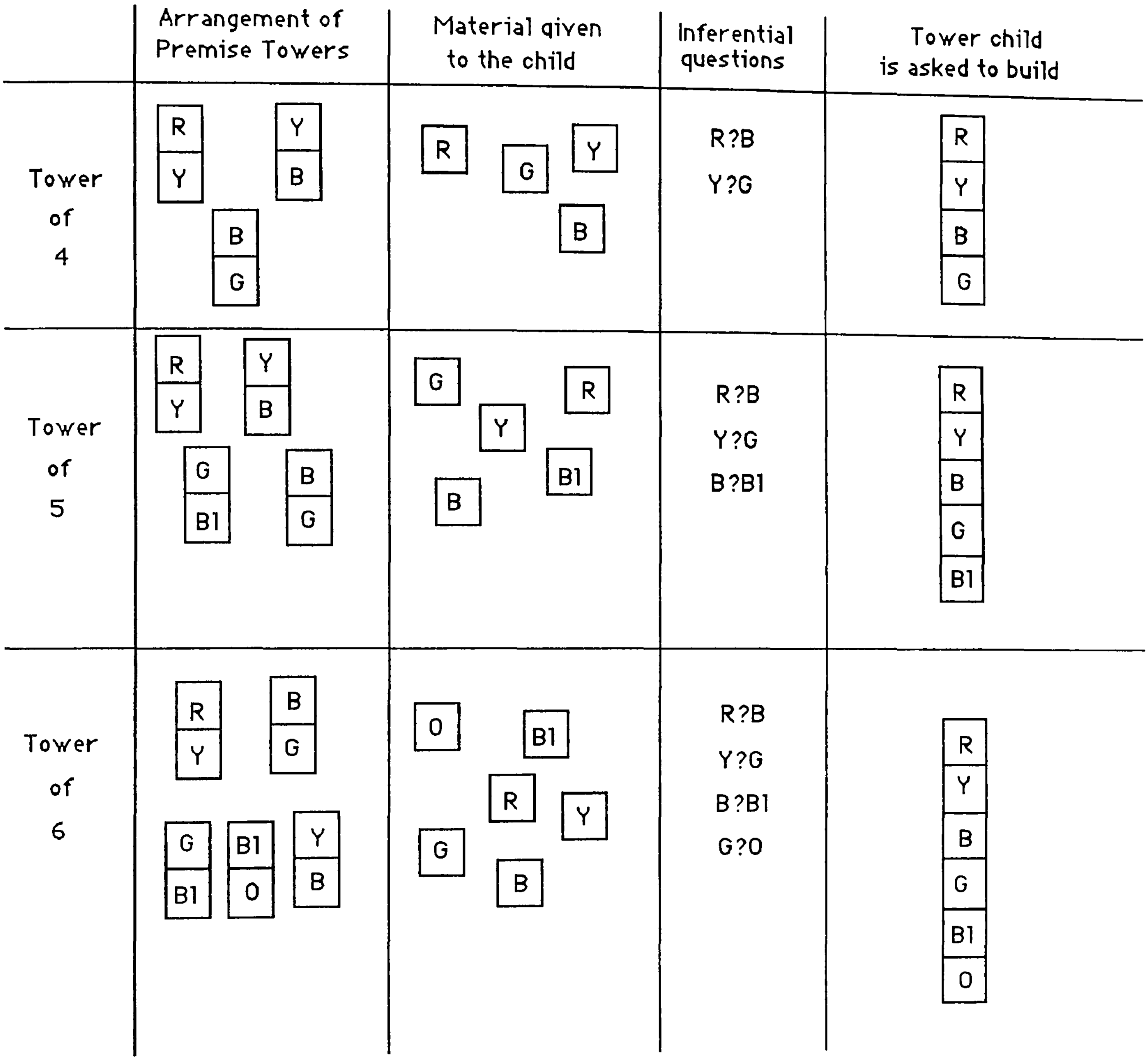
'Which will be higher, yellow or green ?'

Similar questions concerning the relationship between the blue brick and the black brick were added for the tower of five and also between the green brick and the orange brick for the tower of six. Those questions concerning the relationship between the yellow and green bricks (towers of five and six) and also between the blue and the black bricks (tower of six only) were deemed to be 'critical questions'. This is because both of these bricks were as shown as both higher and lower than other bricks. Thus for the tower of five example, the yellow brick is shown in one premise tower as being lower than the red brick, and in another tower as being higher than the blue brick. Also, the green brick is shown in one premise tower as being lower than the blue brick, and in another tower as being higher than the black brick. Therefore, in order to correctly answer a question about the relationship between the yellow and green bricks, the children had to combine information from different premise towers. In contrast, in the tower of five example, correct answers could be given to questions concerning either the red or black bricks by using only one premise tower, as these bricks are 'end-anchors' and are only present in one premise tower. Thus the red brick is always shown in the relation 'higher than', because it is only shown in the premise tower where it is higher than the yellow brick, and the black brick is always shown in the relation 'lower than', because it is only shown in the premise tower where it is lower than the green brick.

Section 4

When the children had answered the inferential questions described above, they were asked to build a tower using the bricks they had been given, and again preserving the relationships displayed in the original premise towers. This was not considered by Pears and Bryant to be a particularly important part of the experiment, as they were using the study to investigate the ability young children to make transitive inferences, rather than to correctly construct an appropriately ordered large tower. They pointed out that "It is in principle possible for the child to use each premise serially to construct the larger tower without ever co-ordinating the information across pairs" (p. 503). Nonetheless, they reasoned that those children who could make transitive inferences about spatial position have at least one appropriate strategy to help them construct the large tower and should therefore be reasonably successful.

Fig. 4.1: Diagram to show Pears and Bryants methodology.



The four year olds in this study performed at significantly above chance on all questions and when constructing the larger towers. Thus there is evidence of early competence in constructing an integrated structural task representation, providing that the relationships relevant to the structure are made explicit and salient by the use of concrete spatial descriptions.

4.1.1 Introduction

In view of the findings reported by Pears and Bryant, it was therefore decided to use a similar paradigm in order to investigate the development of structural reasoning. This first experiment was designed to ascertain whether the precocious performance demonstrated by Pears and Bryant's subjects could be elicited in similar age children when reasoning with another simple spatial relationship, i.e., a horizontal rather than a vertical array* . If an explicit and salient structural relationship is sufficient for success in these types of series problems, then we can expect comparable performance using a similar relationship. It could also be that the use of a simple spatial relationship might facilitate the use of a generalised ordering strategy, as hypothesised by Halford. This would enable reasoning to occur in other more abstract task domains, where the spatial nature of the task structure is not as salient. For example, reasoning about a relation such as 'higher than' which Halford has suggested forms the basis ordering schema for series problems, might encourage subjects to retrieve this schema from memory in order to reason about similar problems based around the relation 'clever than'.

Gentner's (1989) work on analogies, discussed in Chapter 2, also has relevance here. She has differentiated between 'near' and 'far' analogies and has provided evidence that near analogies, i.e. those in which base and target domains share many common surface features, are easier than those which share few or no surface features ('far' analogies). This has been shown both with adult subjects, who typically solve more analogical problems if many surface features are shared, and also with children aged around 6 years, who were successful with problems involving surface analogies, whilst being able to solve those involving deep analogies. An explicitly spatial relationship is, in Gentner's terminology a 'near' analogy, as it shares many surface features in common with the base domain (the ordering schema internalised from

* This will also enable us to investigate the alternative array dimension . However, it must be remembered that the work discussed earlier was referring to abstract evaluative relationships, rather than to concrete spatial ones

everyday life). One might expect that the successful use of an integrated array, internalised from everyday life, to reason about an explicitly spatial domain (a 'near' analogy), would encourage the use of the same strategy to reason about an abstract domain (a 'far' analogy). This is because the 'near' analogy will highlight the importance of the relations between the task items (due to the large amount of surface similarity shared between the base schema and the spatial series problem), thus enabling the same relational structure to be made more salient in the 'far' analogy, which involves an abstract relational series problem.

Aims

This study had two aims:

1. To provide further evidence that the ability to construct an ordered array is a pre-requisite for the ability to solve series problems.

2. To test the following hypotheses:-

a. 4-5 year old children will perform successfully in series problem tasks which are concerned with a simple spatial relationship, such as that used in the Pears and Bryant study, but will be unable to utilise this method of reasoning when given an 'abstract' series problem (a 'far' isomorph of their generalised schema). This is motivated by the work by Gentner discussed above. Based on this, it is hypothesised that 4 to 5 year olds will not yet be able to utilise deeper relational similarities between their generalised schema and the task, without also having surface similarities to act as a cue for analogical reasoning. Furthermore, these children will also be unable to use similarities between the relational structures of the spatial and abstract domains as a cue for analogical reasoning, even when they have just successfully solved the spatial reasoning task. This is because their abstracted schema (Halford, 1992) contains an explicitly spatial element (see Fig. 3.1 on page 70) which they will be unable to isolate from the relational structure, as they are still also relying on surface similarities such as the actual relation used. Thus they will be unable to use the schema to reason about a non-spatial domain, even when it is presented immediately after a spatial task.

Note The relationship 'behind' was chosen, so as to investigate the use of a horizontal array. Work by Walkerdine and Sinha (1975) has shown that children of this age can order objects horizontally, as long as the objects are those which display canonical backs and fronts. For this reason, drawings of different family members, viewed 'side-on' were chosen as the objects. The 'distant' problem isomorph chosen was that of 'happier than'. It was felt that this would carry very few potential preconceptions concerning the relative states of happiness of the different characters used in the study, based on knowledge which the children might import from their everyday interaction. A relationship such as 'cleverer than', on the other hand, might carry with it preconceived ideas that adults are always clever than children, and this would probably influence the way in which the characters were ordered.

b. There will be a developmental progression concerning the use of a generalised ordering schema for the solving of **abstract** series problems by analogy. Piaget's work has demonstrated that, by the age of about 11 years, children are able to successfully solve all analogical problems. Thus we would expect that at this age, the degree of surface similarity between the generalised ordering schema and the series problem would have no effect. Point a. above discussed the hypothesis (motivated by Gentner) that 4 to 5 year olds will only be able to solve spatial series problems, as they will be unable to recognise relational similarities without there also being surface similarities to act as a cue. Again drawing on Gentner's work concerning the developmental shift from surface to deep similarities, we might also expect to see an initial partial use of the generalised ordering schema to reason in the abstract domain. It is hypothesised that this use will firstly occur after the children have successfully solved the spatial series problems. In other words, the proximity of presentation of the two domains will make the structural similarity between them very salient. This salient similarity between the two domains will then be used as a cue to using the generalised ordering schema to reason about the abstract domain. This will, of course, result in an ordering effect, with enhanced performance in the 'far' isomorph when it is preceded by the spatial task. In order to investigate the existence of partial success (dependent on the order of task presentation) in the abstract domain, 7 year old children will be tested, as they are mid-way in age.

c. Older children, around the age of 9 years of age, will not need any cues which highlight relational similarity in order to facilitate problem solving with the abstract relationship. They will spontaneously recognise the structural similarity between the 'far' analogy and their internalised ordering schema. This is because earlier work by Piaget and others has shown that children of this age are capable of analogical reasoning, irrespective of the degree of surface similarity. We would thus expect to see use of the generalised ordering schema to be firmly in place without the need of any cues to prompt its use. This will of course result in there being no ordering effects due to the domain of the series problem.

4.1.2 Method

Design

A three factorial mixed design was used. There were two between subject variables:

- age (5 years, 7 years and 9 years).
- order of task presentation (spatial followed by abstract. i.e. the 'distant' isomorph, and vice versa).

This resulted in six experimental groups, with ten participants in each group.

There was a within subject variable of task type (spatial or abstract reasoning). Each participant received four problems of each task type.*

Participants

Sixty children took part in the experiment. They all attended state nursery (Group 1) or primary schools (Groups 2 and 3) and were of mixed ability. They comprised three age groups:

Group 1 - mean age 4 years 11 months, range 4.7 to 5.2 years

Group 2 - mean age 6 years 9 months, range 6.4 to 7.3 years

Group 3 - mean age 8 years 11 months, range 8.6 to 9.4 years

The children were randomly assigned to 'order of task' groups within their particular age group.

Materials

The stimulus materials consisted of fifteen individual cards (approximately 8.5ins x 4ins). These comprised three copies of five coloured drawings of different people (grandfather, mother, father, girl and boy). The figures were all drawn in silhouette, facing to the left. Copies of the task materials are given in Appendix A.

Procedure

Subjects were tested individually in their school environment . The experimenter first explained the task, using an array of three different characters as a worked example. The children then worked through the four examples without help from the experimenter.

---

* An informal pilot study showed that more than this number resulted in a lack of interest from the 4 to 5 year olds.

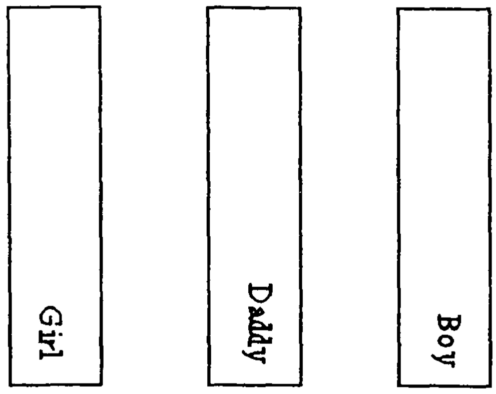
Throughout this thesis, the WISC-R digit span subtest was used as a measure of working memory capacity. The test was administered to the participants before they began the experimental trials. The results from this test enabled us to verify that the working memory capacity of the subject groups was similar to that of their age group norm. Later in the thesis, where comparable age groups were used in one experiment, the test was used to ensure that all the experimental groups had similar scores for working memory capacity. With the exception of Experiments 1, 2 and 10, all the studies were carried out with 7 and 9 year old children as participants. The WISC-R test only supplies age norms for children aged 5 years 6 months and above. Following a successful informal pilot study using the digit span subtest with children aged 4 years 6 months to 5 years 6 months, it was decided to use the test for all age groups, in order to ensure consistency.

Introduction to the task.

The following instructions were read to the children:

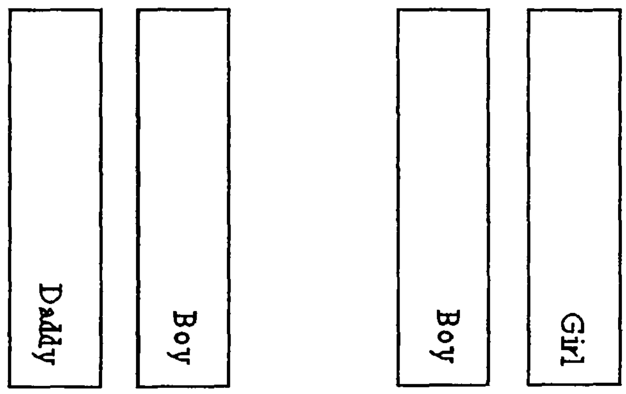
> "I want you to play a new sort of game with me. It's about doing puzzles which put people in the right order. Here are some drawings of different people ...Look, there's a girl, a daddy and a boy".

One copy of each of the appropriate drawings, as depicted in Appendix A, was shown to the children (see below)



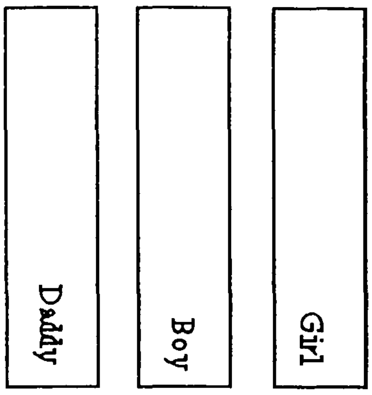
"I'm going to give you some more drawings of the same people so that you can use them as clues to put your drawings in the right order. Here are the clues for this puzzle".

The drawings below were shown to the children.



"The daddy is first over here, then the little boy has followed him. In the other clue, the little boy is first, and the little girl is following him. So the first clue says that the daddy is in front of the little boy, and the second says that the little boy is in front of the little girl. So we can join together the pictures you have to make the right line, because we've got the clues to tell us what to do. Look, like this, daddy *(1)* in front of the little boy *(2)*, and the little boy has to be in front of the little girl *(3)*"

Whilst the last sentence above was being read, the experimenter placed the original cards in the correct order, at the time indicated by the numbers in parentheses (see below)



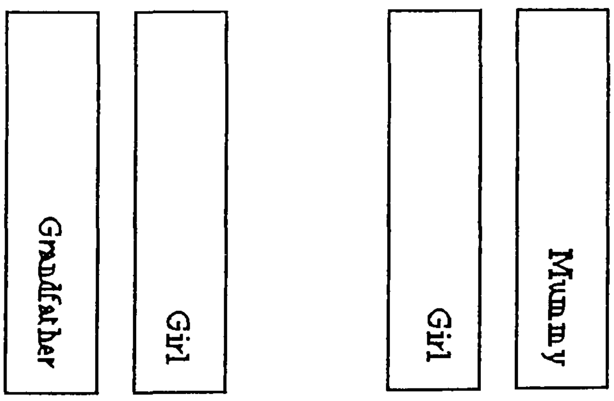
(1) (2) (3)

"Now you do this one by yourself".

A different three cards were given to the subject i.e. grandfather, girl, mummy.

"Here are your clues".

They were then shown their two paired character clues as below.



"Now use these clues to do the puzzle and put your cards in the right order."

Note One 4 year old was unable to do this and so was excluded from the experiment. All the other children were successful.

Experimental task.

Four pairs of cards, i.e. four premises, were used for each problem. This means that the children had to correctly order five different characters for each task. The participants were presented with the premises in non-serial order e.g. boy and mummy, grandfather and girl, daddy and boy, mummy and grandfather. Appendix A gives an example.

The following instructions were read to the children.

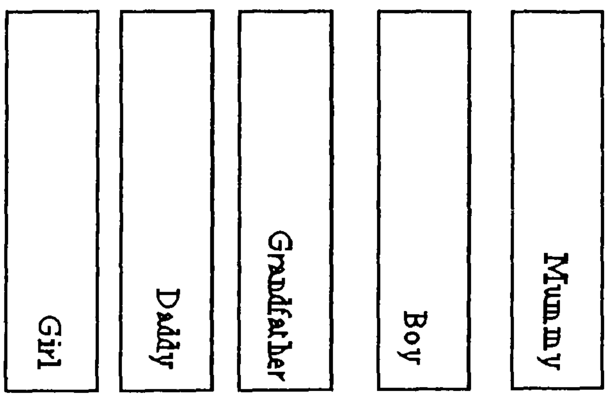
Spatial relation first

> "You're doing really well. Now I've got four more puzzles for you to do. I'll give you drawings of five different people. These people will all be standing in a bus queue and your job is to put them in the right order. I'll give you some clues, just like I did before, so that you can work out what the right order is. Before you put your drawings in order, I'll ask you some questions about where you're going to put some of the different people."

Abstract relation first

> "You're doing really well. Now I've got four more puzzles for you to do. I'll give you drawings of five different people. All of these people are happy, but none of them are exactly as happy anyone else. Another way of saying this is that each person is a bit less or more happy than anyone else. Think about your friends. One of them might very, very happy, because it's their birthday perhaps. Another friend might nearly as happy as the next one because they're going to play with their friend

and to the party. The next friend is quite happy because they're going to the party. And then another friend is not quite as happy because it will be their birthday soon and the last friend might be the least happy because it isn't their birthday for ages. Can you think of something that might make you really, really happy? And then something that might make your mummy nearly as happy? And something that might make you daddy nearly as happy? O.K. This is what the five people are like. They're all happy, but we can put them in order, starting with the person who's the most happy, and finishing with the one who's the least happy. Look at this line here."



The ordering above was shown to the subjects.

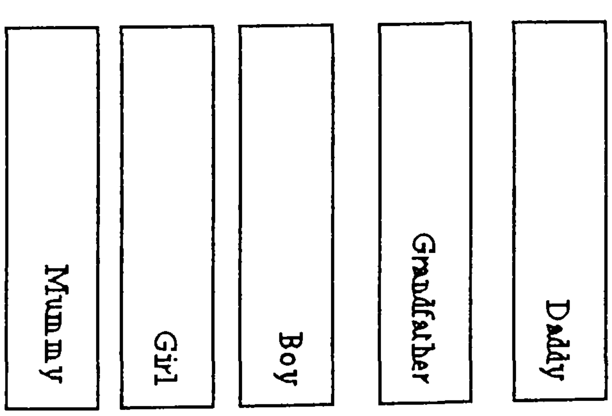
"Here the girl is the happiest, then the daddy, then the grandpa, then the boy and then the mummy. Your puzzles will be about the same five people, but the order they need to be in will change. Your job is to put the drawings in the right order. I'll give you some clues, just like I did before, so that you can work out what the right order is. Before you put your drawings in order, I'll ask you about two of the people and which of the two is happier than the other one."

One copy of each of the five character cards was then put in a pile in front of the subject, followed by the appropriate four pairs of character cards used as premise information. Before the children were allowed to place their cards in the correct order.

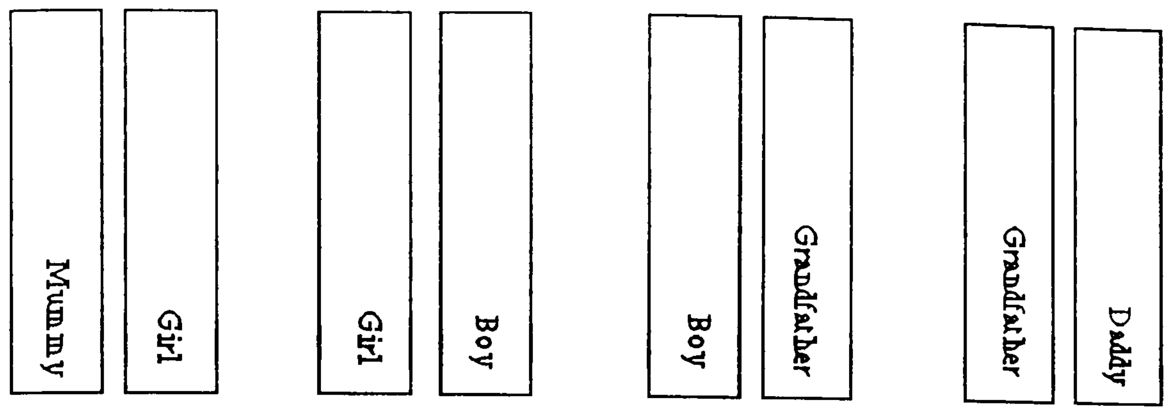
they were asked a series of inferential questions. An example of this for the spatial task would be '

"Who is furthest towards the front of the queue, the daddy or the grandpa?" The same question for the abstract relation would be "Who is the happiest, the daddy or the grandpa?" At the same time the experimenter selected the appropriate pair of cards (the daddy and the grandpa) from the pile in front of the child, so that the child could respond to the question by pointing, if they preferred. As in Pears and Bryant, all the questions were about non-original pairs, for example, the girl and the grandpa, the daddy and the boy and the grandpa and the mummy. Following Pears and Bryant, the critical question was deemed to be the comparison between items B and D (in this example the comparison between the daddy and the boy) This is because both of these characters were as shown as both higher and lower than other characters in different premises. The children were then asked to put the five character cards in the correct order on the table in front of them.

Each child was given four problems for each task type (spatial or abstract). The relative positions of each character and the order of premise presentation (though always random), was varied within subjects. For example, consider the ordering

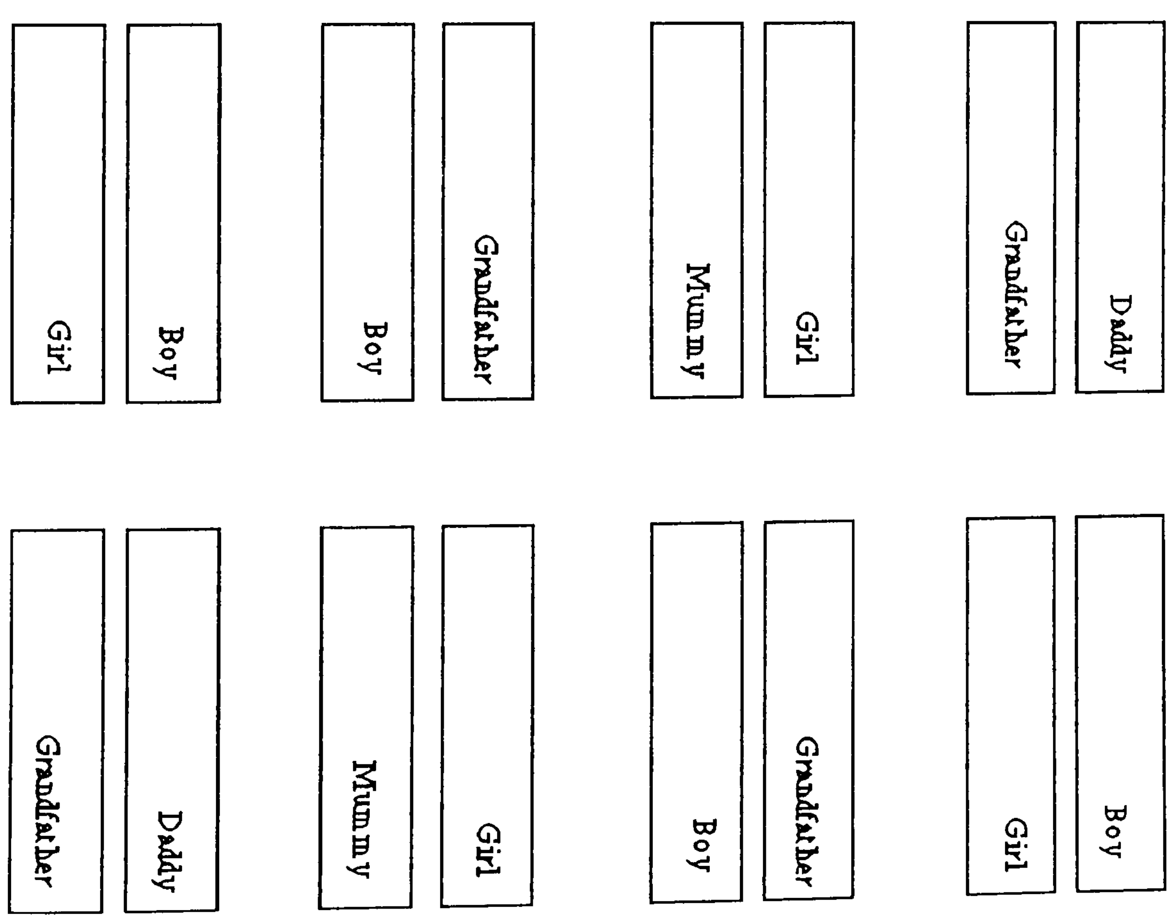


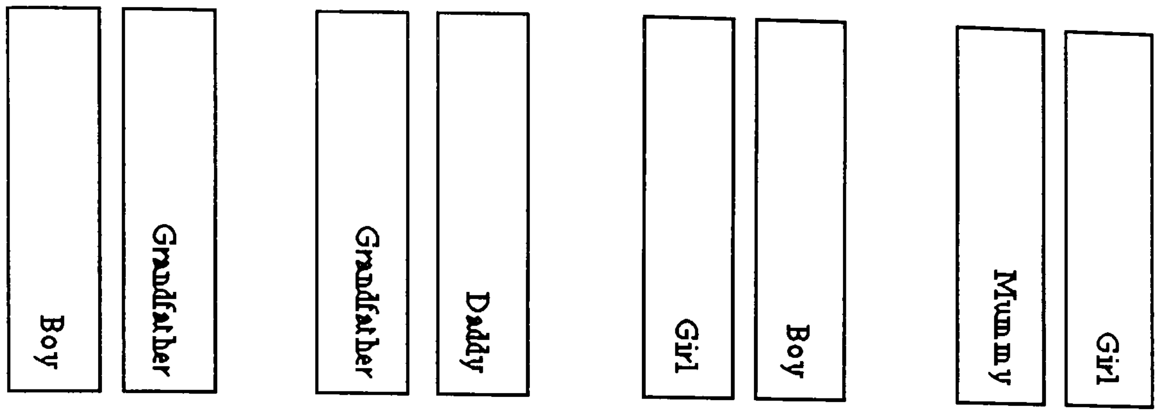
This requires four premises of character pairs as follows:



The order of premise presentation given above is serial, because no actual reordering of premises (and therefore of items) is required. Random orderings are those where some reordering is required. For each 5 item ordering there are 23 possible random premise orderings. Fig. 4.2 shows some examples of random premise orderings of the array given above.

Fig. 4.2: Examples of random premise orderings of the array given above.





The order of premise presentation, though always random, was also varied within groups, such that all 23 possible orderings were used at least once within each condition.

<u>Note</u> As the experiment progressed, it became apparent that the 5 year old children were experiencing considerable difficulty in successfully completing the task. Primarily as a means of improving self-esteem, it was decided to re-order the premises such that they were presented serially. This was done for any trial in which a 5 year old answered all the inferential questions incorrectly and was also unable to build the array correctly.

4.1.3 Results

WISC scores

The results from the WISC digit recall test showed that, although the subjects were performing at a level slightly above that of the norm for their age group, the scoring intervals between the three experimental groups (approximately a score of 2) was in accordance with the scoring intervals between the norm age group scores. The mean scores for each experimental group were as follows:

Group 1 (5 year olds)	6.4

Group 2 (7 year olds)	8.7
Group 3 (9 year olds)	10.5

The relevant average performances on the digit span test are shown below:

Ages between 6 years 8 months and 7 years	8
Ages between 8 years 8 months and 9 years	10

(WISC-R Manual, 1974).

Note

One of the reasons for carrying out this experiment was to enable a direct comparison between with the results reported by Pears and Bryant (1990). In order to do this, the data from the present study will be analysed in two ways. The first of these follows Pears and Bryant's analysis, and is therefore concerned with total number of scores per condition. The second method will consider the mean scores for each condition, and analyses of variance will be carried out, where appropriate.

Answers to inferential questions

Analysis following Pears and Bryant's method

Table 4.1 shows the total number of questions which were answered correctly in each of the six experimental groups. As there were ten participants in each group, and each received four trials in each condition, the maximum number of correct answers in each group is forty.

Table 4.1: Total number of correct answers (critical and non-critical) - scores out of 40

			Task	
Age	**Order**	**Question**	**Spatial**	**Abstract**
5 years	Spat. first	A?C	5	10
		B?D*	0	4
		C?E	6	5
	Abst. first	A?C	5	3
		B?D*	4	8
		C?E	4	9
7 years	Spat. first	A?C	23	19
		B?D*	20	20
		C?E	18	19
	Abst. first	A?C	19	18
		B?D*	21	19
		C?E	17	16
9 years	Spat. first	A?C	31^	31^
		B?D*	24	24
		C?E	31^	26^
	Abst. first	A?C	35^	34^
		B?D*	31~	23
		C?E	28^	32^

* = critical question (see page 91 for a definition)

~ and ^ are explained below.

Since there was only one correct and one incorrect answer for each question, correct answers could have been given by chance 50% of the time (i.e. 20 out of 40 for each summary data point given in Table 4.1). The table shows that only the oldest group (9 year olds) appears to be performing above chance. The 7 year old scores were around chance, and the 5 year olds appear to be performing at well below chance rate.

A binomial test showed that with a probability level of less than 0.05, successful performance at significantly above chance levels occurs when 25 out of the 40 scores are correct. Thus only the 9 year olds performed significantly above chance. They achieved this on all the non-critical questions across all conditions (marked ^ in Table 4.1), and on the critical questions only with the spatial task, and only when it was presented second (marked ~ in Table 4.1).

Analysis of variance.

Due to the floor effects observed for the youngest age group (see Table 4.1) it was not possible to include this group in the following analysis.

In order to investigate the remaining data , a four-way ANOVA was carried out [2 (order - spatial first and abstract first) x 2 (age - 7 years and 9 years) x2 (question - critical and non-critical) x 2 (task - spatial and abstract)]. The first two factors were between subjects, and the last two were repeated measures. It can be seen from Table 4.1 above that each problem involved one critical question and two non-critical questions. Because of this, each data point was converted into a percentage of the maximum score for each condition before being entered into the analysis. Table 4.2 below shows the mean **percentage** of questions which were answered correctly in each of the sixteen experimental groups included in the design given immediately above.

Table 4.2: Mean percentage of correct answers

Standard deviations are shown in parentheses

		Question			
		Critical		Non-critical	
Age	Task order	Task		Task	
		Spatial	Abstract	Spatial	Abstract
7 years	Sp. first	50 (28.86)	50 (33.33)	55.2 (10.54)	47.8 (15.33)
	Abs. first	50 (28.86)	45 (22.96)	44 (13.56)	41.4 (24.47)
9 years	Sp. first	60 (24.15)	55 (19.69)	76.6 (19.95)	71.5 (11.62)
	Abs. first	77.5 (18.44)	57.5 (23.71)	79 (16.67)	82.8 (17.8)

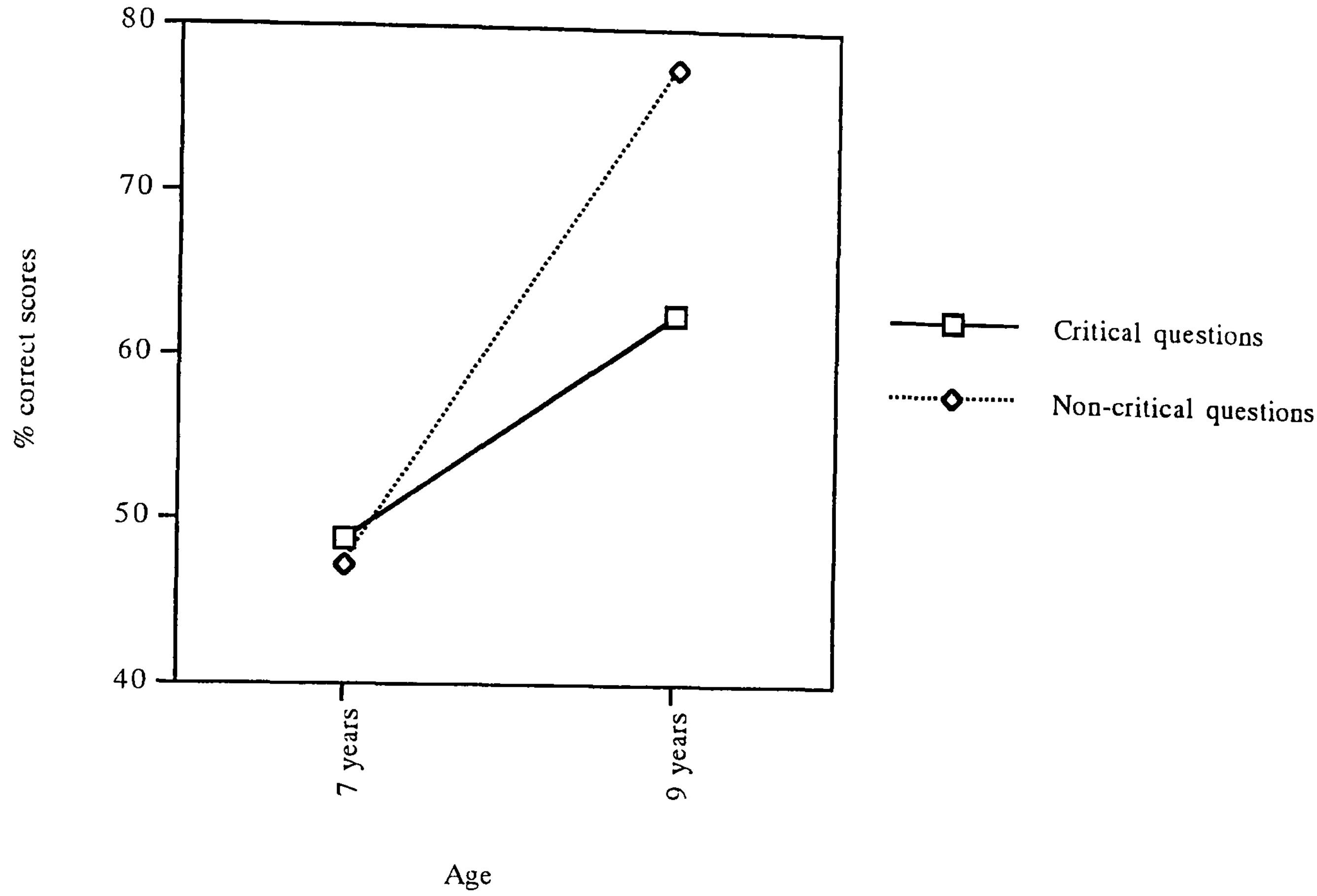
A definition of critical and non-critical questions is given on page 91.

The analysis revealed the following significant effects:

- a main effect of age ($F_{[1, 36]} = 37.896$, $p<0.01$), such that the 9 year olds, with an average percentage successful score of 70%, were more successful overall than the 7 year olds (percentage successful score of 48%).

- a two-way interaction between age and type of question ($F_{[1, 36]} = 4.503$. $p<0.05$) Fig. 4.3 below shows a graphical representation of this interaction.

Fig. 4.3: Age x question type interaction (percentage of correct scores)



The interaction was analysed further, and revealed the following significant simple main effects:

- an effect of age when answering the critical questions, such that the 9 year olds were more successful than the 7 year olds ($F_{[1, 72]} = 6.707$, $p<0.05$).The 9 year olds were successful 62.5 % of the time, whereas the 7 year olds were successful 48.85% of the time.

- an effect of age when answering the non-critical questions, such that the 9 year olds were more successful than the 7 year olds ($F_{[1,72]}=32.730$, $p<0.01$). The 9 year olds were successful 77.5% of the time,whereas the 7 year olds were successful 47.1% of the time.

- an effect of the type of question for the 9 year old children, such that they were more successful with the non-critical questions ($F_{[1, 36]}=7.307$, $p<0.05$). They were successful with the non-critical questions for 77.5 % of the time, and for the critical questions for 62.5% of the time.

Number of children able to order the array

As previously discussed, the question of whether the children could use the information from the premise pairs for the purposes of ordering the concrete array (in this case the tower of bricks) was not of crucial importance in the Pears and Bryant study, as they were primarily interested in the ability to make transitive inferences *per se*. However, it may well be that this will throw some light on the child's ability to determine the structural dependencies which are inherent in the tasks.

Analysis following Pears and Bryant's method

Table 4.3 shows the total number of trials in each condition when the array was ordered successfully using the copies of the premise characters which were provided. As there were 10 subjects in each condition, and 4 trials per subject were counted, the maximum number of successful trials per condition is 40.

Table 4.3: Total number of trials (out of 40) when correct ordering of the array took place

		Task	
Age	**Order**	**Spatial**	**Abstract**
5 years	**Spatial first**	3	4
	Abstract first	4	2
7 years	**Spatial first**	20	31
	Abstract first	31	18
9 years	**Spatial first**	32	34
	Abstract first	35	26

If the arrays were ordered randomly, the chances of them being ordered correctly are 1 in 120 (0.33 out of 40). The number of orderings successfully completed by the 7 and 9 year old children were well above these levels. Even the 5 year olds were managing to do this successfully some of the time. However, it could be that some orders of premise presentation enable easier ordering than others.

Analysis of variance.

Table 4.4 shows the mean number of correctly ordered arrays, using the copies of the premise characters which were provided. Again the 5 year olds have not been included in the analysis, due to the low number of successful scores.

Table 4.4: Mean number of correctly ordered arrays (max=4)

Standard deviations are shown in parentheses

		Task	
Age	**Order**	**Spatial**	**Abstract**
7 years	**Spatial first**	2.0 (0.82)	3.1 (0.57)
	Abstract first	3.1 (0.88)	1.8 (0.63)
9 years	**Spatial first**	3.2 (0.63)	3.4 (0.70)
	Abstract first	3.5 (0.71)	2.5 (0.94)

In order to investigate the above data, a 3 way ANOVA was performed. This was a mixed design - 2 (order - spatial first and abstract first) x 2 (age - 7 years and 9 years) x 2 (task - spatial and abstract). The first two factors were between subjects and the final factor was a repeated measure. The following significant results were obtained:

- a main effect of age ($F_{[1, 36]}=10.286$, $p<0.05$). Overall, the 9 year old children were more successful in completing a correct ordering (overall means were 7 year olds - 2.5, 9 year olds - 3.17).

- a 3 way interaction ($F_{[1, 36]}=5.268$, $p<0.05$). This is shown in the following two graphs.

Fig. 4.4: 7 year old children-interaction between task and task ordering.

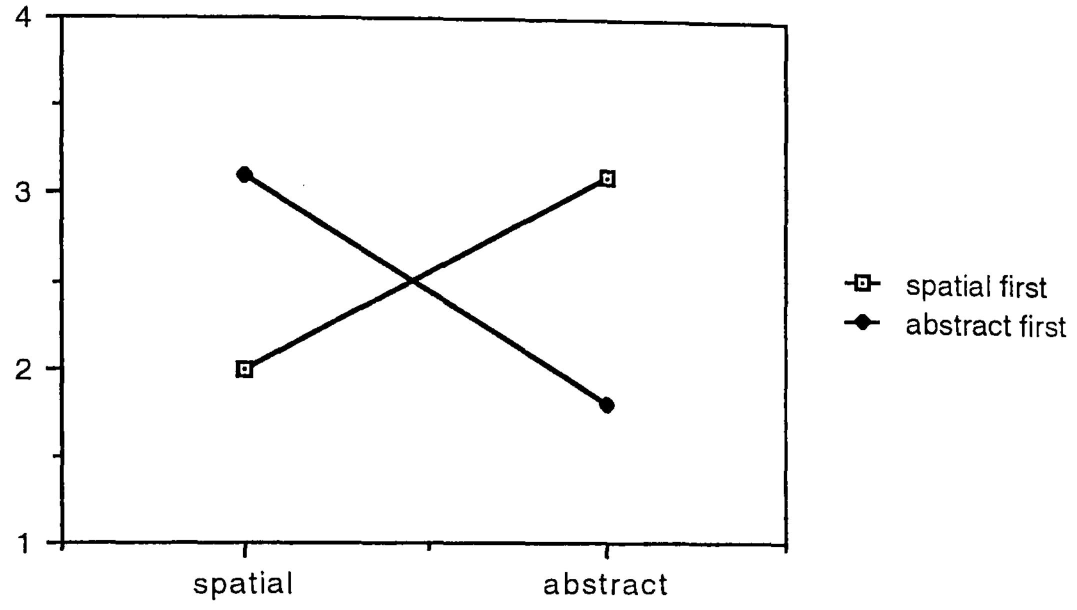
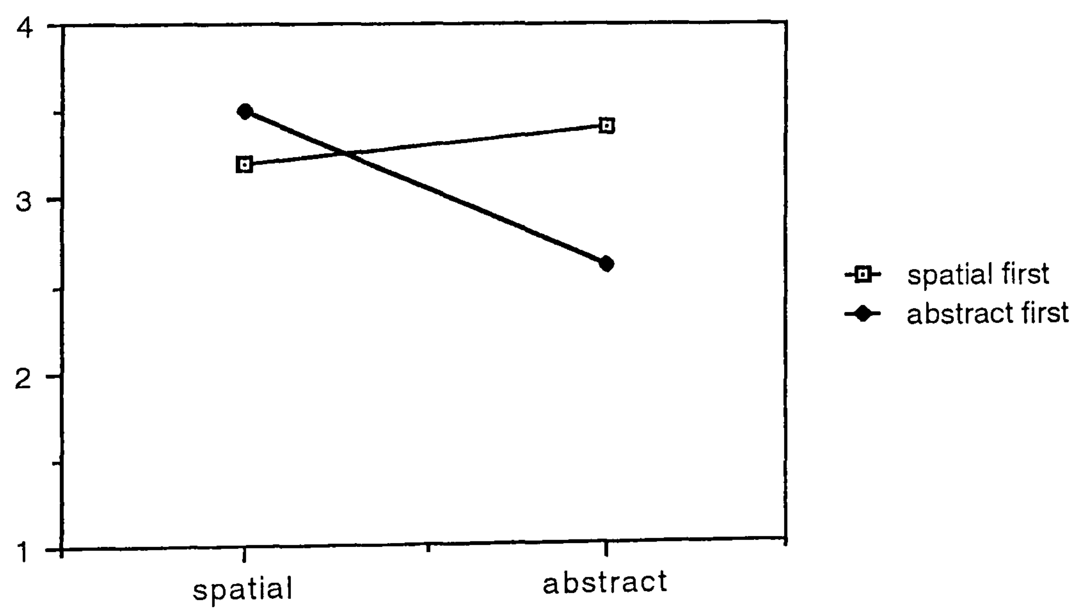


Fig. 4.5: 9 year old children-interaction between task and task ordering.



This interaction was analysed further, and revealed the following significant simple simple main effects:

- an effect of task order when the 7 year olds were building the array with the spatial task, such that the subjects who completed this task second were more successful than those subjects who completed it first ($F_{[1,36]}=7.751$, $p<0.01$). The mean score for those subjects who completed the spatial task second was 3.1 whereas the mean score for those who completed first was 2.0.

- an effect of task order when the 7 year olds were building the array with the abstract task, such that the subjects who completed this task second were more successful than those subjects who completed it first ($F_{[1,36]}=10.826$, $p<0.01$). The mean score for those subjects who completed the spatial task second was 3.1 whereas the mean score for those who completed first was 1.8.

- an effect of task order when the 9 year olds were building the array with the abstract task, such that the subjects who completed this task second were more successful than those subjects who completed it first ($F_{[1,36]}=5.189$, $p<0.05$). The mean score for those subjects who completed the spatial task second was 3.4 whereas the mean score for those who completed first was 2.5.

- an effect of age when the array was being built in the spatial task and when it was presented first, such that the 9 year olds who completed this task were more successful than the 7 year olds ($F_{[1, 36]}=9.224$, $p<0.01$). The mean score for the 9 year olds was 3.2 whereas the mean score for the 7 year olds was 2.0.

- an effect of task type when the 7 year olds were building the array and the spatial task was first, such the abstract task was completed more successfully than the spatial task ($F_{[1, 36]}=17.707$, $p<0.01$). The mean score for the abstract task was 3.1 whereas the mean score for the spatial task was 2.0.

- an effect of task type when the 7 year olds were building the array and the abstract task was first, such the spatial task was completed more successfully than the abstract task ($F_{[1, 36]}=24.732$, $p<0.01$). The mean score for the spatial task was 3.1 whereas the mean score for the abstract task was 1.8.

- an effect of task type when the 9 year olds were building the array and the abstract task was first, such the spatial task was completed more successfully than the abstract task ($F_{[1, 36]}=14.634$, $p<0.01$). The mean score for the spatial task was 3.5 whereas the mean score for the abstract task was 2.5.

Serial re-ordering of premises

It must be remembered that the premises were only re-ordered serially for the 5 year olds. It was done for any individual trial in which the child was unsuccessful in all the inferential questions and also in building the queue. The results of this re-ordering are reported here in order to enable a comparison with similar results reported in Experiment 2.

Inferential questions

Table 4.5 shows the number of questions answered correctly in each trial.

Table 4.5: Total number of questions answered correctly (serial re-ordering for the 5 year old children)

The numbers in parentheses give the total number of children in each trial for whom the premises were re-ordered serially.

Order	Question	Task: Spatial	Task: Abstract
Spatial first	A?C	26(34)~	28(30)~
	B?D*	25(34)~	25(30)~
	C?E	27(34)~	27(30)~
Abstract first	A?C	21(33)	22(31)~
	B?D*	23(33)~	19(31)
	C?E	26(33)~	22(31)~

* = critical question

It can be seen that all the questions were answered correctly at a frequency above that of chance (50% success rate).

At the 0.05 probability level, a binomial test showed that a score of 19 or more out of 30, 20 or more out of 31 and 22 or more out of 33 or 34 was significantly better than chance. Thus, for 10 out of the 12 questions (marked ~ in Table 4.5), the children were performing significantly above chance.

Building of array

Table 4.6: Total number of successful queue completions (serial re-ordering for the 5 year old children)

The numbers in parentheses give the total number of children in each trial for whom the premises were re-ordered serially.

	Task	
Order	**Spatial**	**Abstract**
Spatial first	27 (34)	26 (30)
Abstract first	25 (33)	23 (31)

The equivalent percentage figures are given in Table 4.7 below.

Table 4.7: Percentage of successful queue completions- serial re-ordering

These are expressed as a percentage of the total number of serial re-orderings.

	Task	
Order	**Spatial**	**Abstract**
Spatial first	80	83
Abstract first	73	74

A 2-way analysis of variance was carried out on the above data and revealed no significant differences ($F_{[1, 18]}=0.044$, $p>0.05$). However, it has to be remembered that the individual successful scores which were analysed are percentages of the total number of trials for each child where the premises were re-ordered serially. Thus a successful percentage score of 100% could signify that a child was unsuccessful with random ordering for all his trials (a total of 4) and then successful with them all when re-ordered serially. On the other hand, a less successful percentage score of 50% could signify that a child was successful with random ordering for half his trials (a total of 2) and then successful with only one of those two when re-ordered serially. Thus the child who presumably has a better grasp of the task (because they are able to reason using random ordering for at least some of the time) has a lower percentage successful score for this analysis, as we are considering only serial re-ordering. Indeed, the main reason that this analysis was carried out was to enable a comparison to be made with Experiment 2.

4.1.4 Discussion

Random premise ordering

Answers to inferential questions

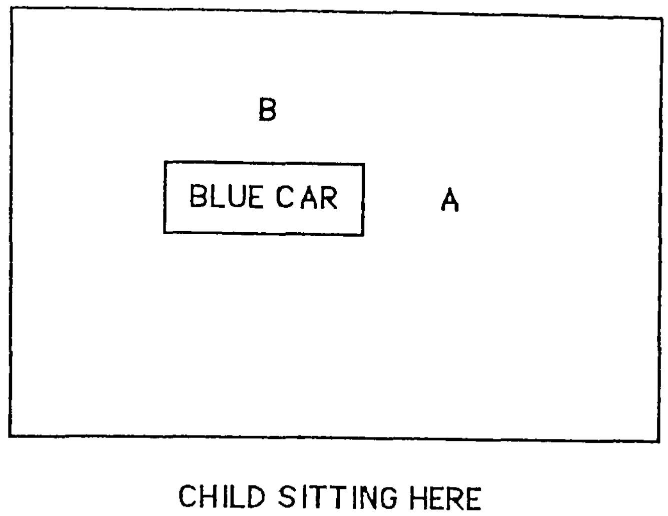
It can be seen from the results that there is no support for the experimental hypotheses. The precocious performance obtained by Pears and Bryant in their study was not replicated with either the 5 or the 7 year old children, who did not perform significantly above chance on either the critical or the non-critical questions. A binomial test performed on the data showed that only the 9 year olds were performing significantly above chance when answering some of the questions, whilst the 5 year olds were significantly below chance. This can be easily explained by this group's great reluctance to guess the answers. The cause of this reluctance was not totally clear, however. It could be that the type of question they were being asked was unfamiliar. The children were attending the local nursery school. To date, most of their experience in answering questions has been concerned with those problems which can be solved by observation. For example, they might be asked to look at a picture of children playing and asked 'Who's holding the red ball?' It could therefore be that answering questions which require active problem solving is very difficult. Also, some of the children remained rather timid throughout testing, which could explain their inhibitions.

The improved performance by the oldest age group was also shown by the analysis of variance. They were more successful than the 7 year olds overall, and also when considering both the critical and non-critical questions separately. If we look solely at the 9 year olds performance, they showed a significant improvement when answering the non-critical questions in comparison with the critical questions.

There was however no interaction of age with either the task (i.e. whether spatial or abstract reasoning was required) or task ordering. Thus the main effect of age can be explained by its effects on the type of question. It had no effect on successful performance in either the spatial or the abstract task, regardless of the order in which these tasks were presented.

These results are interesting for several reasons. Firstly, the poor performance shown by the 5 and 7 year old children needs investigating further, especially when comparing it with the Pears and Bryant study. The discrepancy could be for several reasons, one of which concerns the actual relational term used here (behind), which might have given the children more problems than the 'on top of' relationship used by Pears and Bryant. In a three dimensional space, there is a potential ambiguity in that 'behind' might mean either 'to the right of' or 'on top of'. For example, in the situation below, a request to the child to place a red car behind the blue car might result in a placement at either position A or B. If the children were answering the inferential questions by constructing an internal integrated array, then this potential ambiguity may have affected performance.

'Birds eye' view



It is considered unlikely this type of confusion is happening in the current study, as the subjects were working only in a horizontal dimension, with figures which had obvious fronts and backs, so they only had a placement choice of right or left. Also, nothing was observed during the experiment which suggested any confusion. Nonetheless, Pears and Bryant showed that similar aged children were apparently more at ease when working with a tower of bricks and using a vertical spatial relation ('higher than') than the children in this study who used a horizontal relation ('behind'), together with drawings of people (nonetheless, they were still not successful in completing the critical inferencing questions). Other likely reasons for the lack of replication of Pears and Bryant's results are methodological (though the detailed Method section was very closely followed) or cultural, as the Pears and Bryant study was carried out in France. This could have resulted in differing educational methods, or length of time in formal education. In order to address these possible reasons, it was decided to replicate the study reported by Pears and Bryant (1990). The report of this begins on page 121.

As previously discussed, it seems likely that children will solve series problems by constructing an integrated internal array. This is because it imposes a lighter working memory load than that of encoding individual premises (Breslow, 1981). The children in the current study did not need to commit the premise information to memory, however, as the premise towers remained in view throughout the experiment. Nonetheless, the 5 and 7 year olds were unable to correctly answer any of the inferential questions at a rate significantly above that of chance. It seems therefore, that these children were still experiencing some difficulty in actually integrating premises internally (that is, constructing a structural task representation). They were not allowed to actually construct the full tower using real bricks until after they had answered the inferential questions, thus some internal manipulation of premise information was still required.

Notwithstanding the differences between the current study and that of Pears and Bryant, as discussed above, it seems from this experiment that the evidence concerning children's ability to build an integrated **internal** spatial array is not as robust as expected. The results from the two studies suggest that some of the difficulties which 5 year olds experience in building an integrated internal array might depend on the actual spatial relationship used. However, the older (7 and 9 year old) children in the current study were also not performing at ceiling level when answering the inferential questions, suggesting that they also hadn't fully mastered the task of internally integrating premises into an ordered array.

Using the evidence from the construction of real arrays

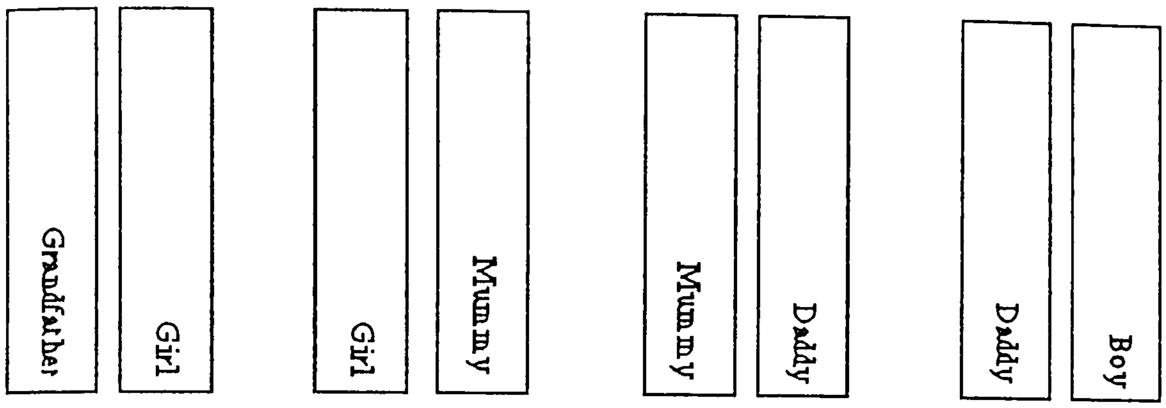
For the reasons given above, the analysis of variance of the 'concrete array' data was carried out. If we accept that children, at least initially, answer these types of problems by constructing an internal integrated ordering, then it seems reasonable to look at their performance in constructing a similar external problem representation. This should highlight some of the difficulties involved, and it will enable us to consider whether the source of the child's difficulty lies in the inability to identify and use the appropriate structural relationships. Riley (1976) working in the domain of series problems, has shown that data patterns obtained when children have been working with external orderings correlate with their performance when required to perform all operations mentally. She used drawings of six children's faces, with the first name written underneath. She told her subjects that they were going to learn about six imaginary children, who differed on one dimension (for example, weight). During the training phase, the children learned the comparative weights of the adjacent pairs of children. They were then tested about the comparative weights of non-adjacent pairs. Half of the subjects had access to duplicate drawings of the six faces during the training phase. They were told that they could use these to help them remember the premise relations. All of these children which were tested in the weight condition organised the pictures into a linear order. They spent much less time in training, and made fewer errors than those children who could not use duplicate drawings. However, both groups were fastest with those comparisons involving end-anchors. Of the internal pairs, both groups of subjects showed an inverse relation between decision time and number of intervening items between the required comparison (see Chapter 3 for a further explanation of these). Bearing this in mind, it is likely that the problems which arise during external integration of items will reflect similar issues concerning internal integration.

Constructing the external array

It can be seen from the results that even the 5 year old children were able to build the array correctly for some of the time. However, the numbers of successful attempts were quite low for these children, and probably represent tasks when a relatively easy order of presentation of premises was given. This suggestion is explained more fully below.

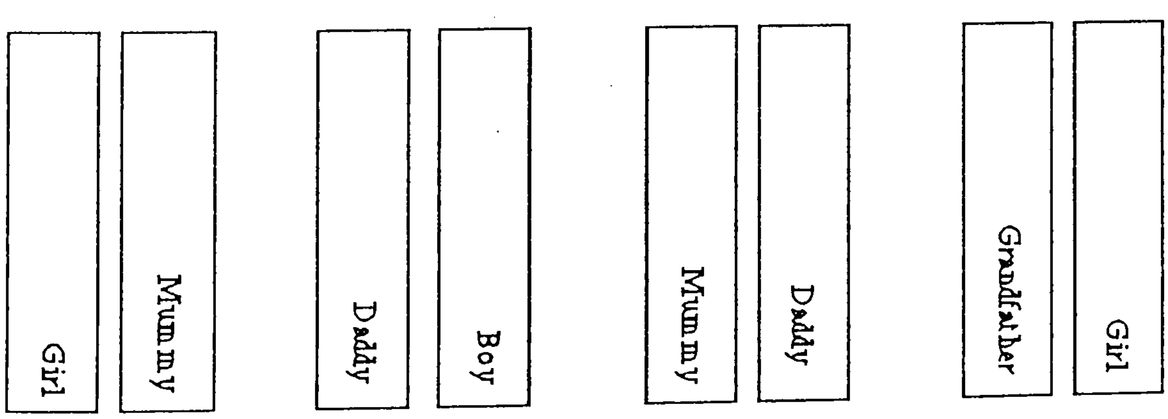
There were four character pairs used as premise information. An example of this is given in Fig. 4.6 below.

Fig. 4.6: Example of paired-character premise information



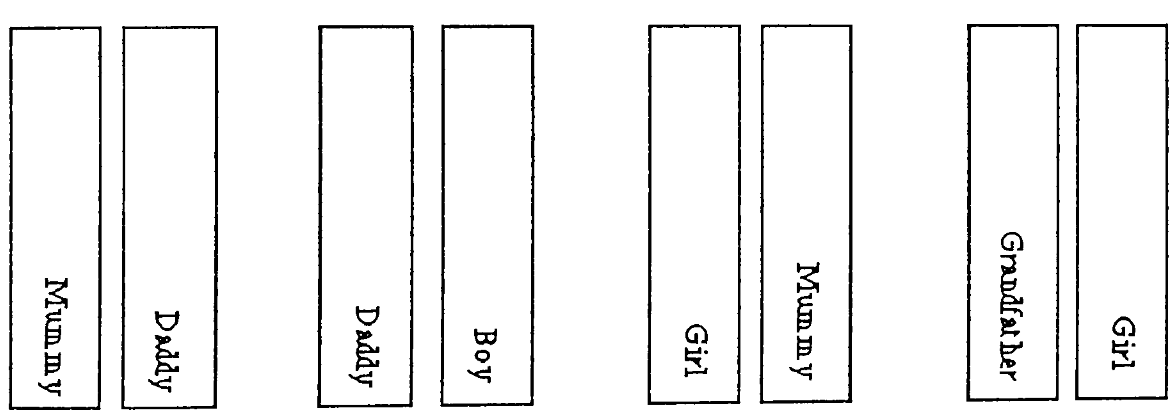
The ordering shown above is serial, because no manipulation of the premise positions is required to correctly order the array. None of the subjects in this experiment were presented with the premises ordered serially. Any order other than that represented above has been termed 'random ordering'. This means that at least one of the premise positions needs to be altered before the array can be ordered. Any 4-premise problem has 24 possible orderings, 1 of which is serial and the other 23 are random. Some examples of random orderings for the equivalent serial ordering above are given in Figs. 4.7, 4.8 and 4.10 below.

Fig. 4.7: Example of a random premise ordering



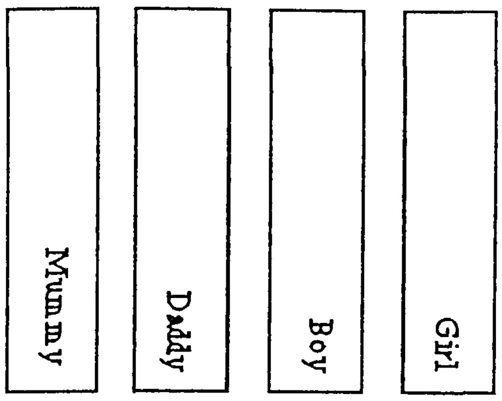
Each condition in this study had 10 subjects, who completed 4 trials in each condition. Each subject always used a different random ordering for each trial they completed. However, these orderings were not varied **systematically** within subjects, or between groups. Nonetheless, records were kept of the different types of random premise ordering that each child received. It appears from these that the requirement to place an item in front of, rather than behind, a partially ordered array inhibited successful performance, and therefore that the number of these requirements in any particular trial determined that trials level of difficulty. It seemed to the experimenter that the subjects considered the premises strictly in the order that they were presented (however, no formal records were kept concerning this). This would explain why the need to place an item at the front of a partially ordered array sometimes arose. Fig. 4.8 shows an example of a suggested complex random premise presentation of the ordering given in Fig. 4.6 above, and also illustrates the suggestion made above.

Fig. 4.8: Example of a suggested complex random premise presentation.



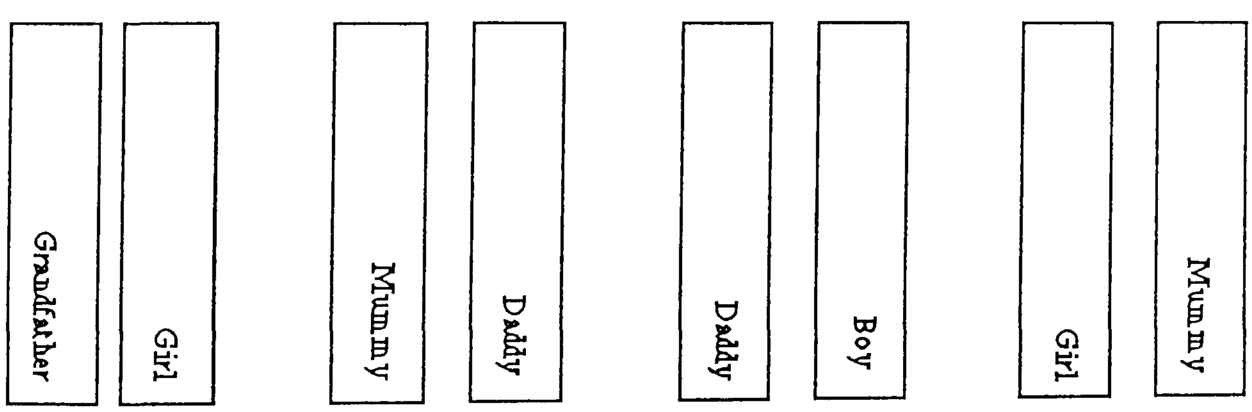
If, as is suggested, the child considers each premise in the order it is presented, he or she will begin to order the items by first placing the 'Mummy' and 'Daddy' cards, followed by the 'Boy' card. The next novel item encountered by scanning the array from left to right is the 'Girl' card. The partially built array will therefore be as in Fig. 4.9 below.

Fig. 4.9: Partially built array



In order for the premise information to be followed correctly, the 'Girl' card needs to be placed at the **front** of the current array. It appears that in the current experiment, this posed considerable difficulty. In contrast, Fig. 4.10 shows a simple random premise presentation for the same array.

Fig. 4.10: Example of a suggested simple random premise presentation.



In this example, provided that the premises are considered in the same order that they are presented, the array can be successfully constructed without the need to place any items at the front of a partially ordered array.

Because of the lack of systematic variation of types of premise orderings in the current experiment, further consideration of the data concerning the above is not possible.

Also, it is conceivable that the children had difficulty in realising that the mid-items were represented in two different premise-pairs, in that they were two copies of the same drawing. In order for the premise information to be successfully integrated, the child needs to amalgamate these two representations into one character. In the example above, the children are looking at two copies of the girl, mummy and daddy. They need to be aware that they only need one of each of these characters to construct a linear ordering. If the children were trying to construct an array still using two drawings of the mid-items then their performance will be greatly inhibited, as only one drawing of each character was made available to them. This will be investigated in Chapter 5, together with a systematic consideration of the effects of premise ordering.

Due to the large number of nil scores in the 5 year olds data only the 7 and 9 year old's results were used in the analysis of variance. The results of this showed that, overall, the 9 year olds were more successful than the 7 year olds, and also showed that the different age groups were performing differently from each other, according to task and order of presentation. This difference is explained more fully below.

The 7 year olds showed a simple practice effect. This is because their performance always demonstrated a significant improvement when reasoning with the task type (either abstract or spatial relationship) which was presented second. Thus there was no effect of the actual task types, and the children appeared merely to be learning how to manipulate the items so as to construct the correct array.

The 9 year old children needed no practice when the spatial relationship was presented first. Their performance was virtually at ceiling with this spatial relation in this condition and therefore showed no significant improvement with the second relationship (abstract).

However, when the abstract relation was presented first, the 9 year old children were not performing at ceiling (even though they were given a simple introduction). The children in this condition showed a significant improvement when they completed the second set of trials (the spatial relationship task). Also, the 9 year olds performance on the abstract task was significantly better when it was presented second than when it was presented first. However, performance when the spatial task was presented second was not significantly better than when it was presented first (presumably because performance was already at ceiling).

These results suggest that 9 year old children are able to map task structure from a generalised ordering schema in order to solve series problems. However, there appears to be a distinction between mappings made on the basis of surface similarities and those made on the basis of deep similarities (Gentner, 1989). For the purposes of this experiment, surface features are defined as the actual relationship used. Thus the spatial task and the generalised schema share surface features, as they both use the relation 'behind'. Deep features are defined as the relational structure. Thus the abstract and spatial tasks share deep features as, whilst different actual relationships are used, they share the same structure.

Those trials which did not share surface features with the schema were not successfully solved by analogical mapping, unless deep similarities in the relational structure of the task and schema were made salient by the prior presentation of a similar task. This task, as well as being structurally isomorphic to both the task in question and the generalised schema, also shared the surface features of the schema. The salient surface similarity between the spatial relationship task and the generalised schema, and the salient deep similarity between the two task types (spatial and abstract) enabled analogical mapping to occur between the schema and the abstract relationship task. Without the benefit of this 'bridge', which provided a link between the two task types and the generalised schema, the 9 year olds were unable to map relational structures from their generalised schema to the abstract series problem. This result provides some support for Gentner's argument that children are at first reliant on surface features when reasoning analogically. It seems that this reliance stems from a need to make structural similarities salient. Once surface similarity has been recognised and used, the close proximity of another problem which shares a deep structure cues mapping based on the deep relational similarity between series problem domains.

However, the results of most interest appear to be those shown by the 7 year old subjects. These children did not immediately recognise how to relate items together into an integrated representation, irrespective of the relation in which they were working. Thus, although there were salient surface similarities between the spatial task and their hypothesised generalised schema, analogical mapping did not occur until the children had experienced several trials. This practice effect occurred irrespective of the type of task. It seems therefore that there are constraints other that of surface similarity which affect successful analogical mapping in order to solve series problems. Due to repeated exposure to very similar problems, the 7 year old subjects in this study were able to eventually overcome these constraints. However, children are not often in this position and need to be able to map between relational similarities in novel tasks. Focusing on the actual construction of an external ordered array should give some insights as to the constraints children face in constructing and using structural relationships. If we can identify some of these, we could present problems in a way which will support a systematic and structural problem representation. It is hoped that this might facilitate analogical reasoning by providing a bridge across which appropriate mappings will more readily be made. It may be that this will result in more immediate mapping, without the need to focus on initial surface similarities.

4.1.5 Summary

Experiment 1 had two aims, which are summarised below.

1. To provide further evidence that the ability to construct an ordered array is a pre-requisite for the ability to solve series problems.

2. To test the hypothesis that 5 year old children would only be able reason successfully using the spatial task, whilst 9 year old children would be successful irrespective of task type. It was also suggested that there would be an interim situation (probably around 7 years) where the children would be able to transfer successful analogical mapping from the spatial domain to the abstract one.

It has been shown that the ability to integrate individual premises into a integrated array is very difficult for the 5 year olds, and initially difficult for the 7 year olds, though they do get better with practice. These children also had difficulties in answering inferential questions. Thus this study adds support to arguments made by Trabasso and others that series problems are solved by the construction of an integrated array. In respect of hypothesis 2, successful performance was not shown at all by the 5 year old children. Nonetheless, the oldest group (9 year olds) did provide some evidence that spatial series problems are, at least initially, easier than abstract ones. However, it seems from the 7 year olds performance that there are constraints affecting the integration of premises into an ordered array which occur irrespective of the task domain (spatial or abstract).

For this reason, it was decided to conduct further studies to concentrate on how children construct an external ordered array when given a common series problem. The following set of experiments has been designed to investigate the existence of two possible constraints.

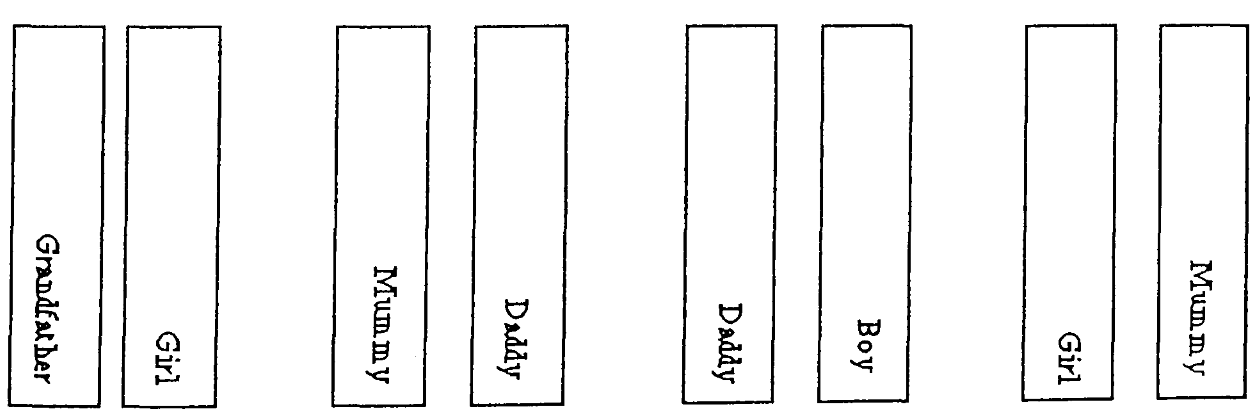
1. The effects of the order of premise presentation.

Although the order of presentation in the current study was random, there appeared to be certain orderings which the children were more successful with. It was noticed that the children seemed to consider the premises strictly in order of presentation. This means that, for certain premise orderings, items will have to be placed in front of a partially ordered array. It is proposed that those orderings which require this will give particular problems.

2. An inability to deal with redundant mid-items.

It is conceivable that the children did not realise that in order to successfully integrate the premise information, only one drawing of each character was required, even though all the mid-array characters were represented twice in the premise information. This is illustrated in Fig. 4.11 below.

Fig. 4.11: Example of random premise information.



In the example above, the 'Daddy, the 'Girl' and the 'Mummy' are represented twice in separate premises, as they are all 'mid-items' characters. In order to integrate the separate premises into one array, the children need to be aware that two similar drawings are actually different representations of the same character. A lack of this knowledge would have prevented them from successfully ordering the array, as this is dependent on integration of premise information.

It was decided, based on work by DeLoache in 1991, to use photographs as stimulus materials, rather than drawings. This is because DeLoache showed that photographs were the easiest graphic representation for young children to understand. This is also consistent with work reported by O'Connor, Beilin and Kose (1981) who argued that 6 year olds believe in the fidelity of photographs, as they are more likely to accept illogical outcomes as true when they were represented in photographs rather than in drawings. These two studies suggest that children might find it easier to accept that two identical photographs in separate premises were of one character and therefore realise the necessity to use only one photograph of a character when building the integrated array.

Further chapters in this thesis will therefore describe the group of studies which addressed the two factors listed above.

Before we begin to consider issues concerning the construction of an ordered external array, it is necessary to address the poor performance shown by the 5 year old children in this study, when compared with the results obtained by Pears and Bryant (1990).

As previously discussed, the difference in results could be due to the actual relational term used. The current study used the relation 'behind', thus ordering in a horizontal dimension, whilst the Pears and Bryant study used the 'on top of' relation, thus working in a vertical dimension.

In order to investigate this, it was decided to replicate Pears and Bryant's experiment. If their results are also replicated, we will have some evidence that the relational term 'behind' poses more problems for 5 year old children when solving series problems than the relational term 'on top of'.

4.2 EXPERIMENT 2 - REASONING ABOUT SERIES PROBLEMS IN A VERTICAL SPATIAL DOMAIN (A REPLICATION OF PEARS AND BRYANT, 1990)

4.2.1Method

Design

A one-factorial repeated measures design was used. Each subject was presented with twelve trials. Four of these trials involved working with a tower of four bricks, four with a tower of five bricks and the final four trials used a tower of six bricks. Thus the independent variable was the number of bricks in the tower. The measures of performance were the number of questions answered correctly and the number of correctly ordered towers. See page 80 for a diagrammatic representation of the task.

Participants

The fifteen children who participated in the study were all attending a state nursery school in a mixed catchment area. Their mean age was 4 years 9 months (range 4 years 6 months to 4 years 11 months). Five children did not succeed in building a tower in the practice trial and were therefore excluded from the experiment.

Materials

The task materials were different coloured bricks measuring approximately 1cm cubed. Thess fitted together in a similar manner to 'LEGO' bricks. Photographs of the task materials are given in Appendix B.

Procedure (this has been taken from Pears & Bryant 1990)

Introduction to task

The children were introduced to the task materials, and were told that they would be playing a game with them. The children were already familiar with the task materials, as they are used regularly for number work. Each child was then shown two 'premise towers' each consisting of two bricks of different colours. For example, there might be a tower with a red brick above a green brick, together with a tower with a green brick above a black one. The child was then given three bricks, in this case black, red and green and asked to build one three-brick tower so that they kept the relationships already shown to them in the two smaller towers. The experimenter then pointed to the two-item towers, explained the relations to the child, and told them to use these to help them to build their tower.

Experimental task

See page 80 for a diagrammatic representation of the task.

The children were presented with twelve trials in all, divided into three blocks of four, in the following order:

Block 1 - they were given three two brick towers to work with. This meant that the tower which they eventually had to build and about which they were asked inferential questions consisted of four bricks.

Block 2 - as above but with four two brick premises and a five brick eventual tower.

Block 3 - as above but with five two brick towers and a six brick eventual tower.

The relative position of the different colours was changed on every trial and varied between trials.

Each trial (12 for each subject) followed the following procedure:

The premise information (in the form of two-brick towers) was presented to the subject. It was presented in random order, so that the children could not solve the problem by reading off the order of the colours as they were presented to them.

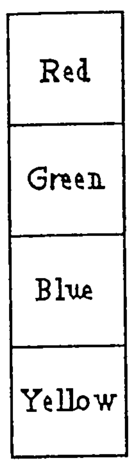
The children were then given duplicates of the different coloured bricks used in the premise towers i.e. four bricks for the 'tower of four' etc.

They were then asked inferential questions concerning the order of colours in the tower which they would build if they preserved the relationships displayed in the premise towers. The questions were always about non-original pairings-see Fig. 4.12 below.

Fig. 4.12: Examples of completed towers and associated inferential questions

For the tower of four

Example completed tower



Inferential questions

Which will be the highest-the red brick or the blue brick?

Which will be the highest-the yellow brick or the green brick?

For the tower of five

Example completed tower

Red
Green
Blue
Yellow
Black

Inferential questions

Which will be the highest-the red brick or the blue brick?

Which will be the highest-the yellow brick or the green brick?

Which will be the highest-the blue brick or the black brick?

For the tower of six

Example completed tower

Red
Green
Blue
Yellow
Black
Orange

Inferential questions

Which will be the highest-the red brick or the blue brick?

Which will be the highest-the yellow brick or the green brick?

Which will be the highest-the blue brick or the black brick?

Which will be the highest-the green brick or the orange brick?

The questions 'Which will be the highest-the yellow brick or the green brick?' (towers of five and six) and also 'Which will be the highest-the blue brick or the black brick?' (tower of six only) were deemed to be critical questions as they involved two items which were not end points and which had been shown as both higher and lower items in the premise towers.

When the children had answered the questions, they were asked to build a tower using the bricks they had been given, and again preserving the relationships displayed in the original premise towers.

Children who answered all the inferential questions incorrectly and were also unable to build the tower correctly in any one trial were then presented with the premises ordered serially for those trials.

WISC-R digit recall scores were also taken from each subject.

4.2.2 Results

The results from the WISC-R showed that the mean score for the subjects used in this study was 6.3. This is similar to the group of comparably aged children used in Experiment 1.

Note As in Experiment 1 the analysis of data has replicated that carried out by Pears and Bryant (1990). This was to enable a direct comparison between their results and those of the current study.

Random presentation

Inferential questions

Table 4.8 shows the total numbers of subjects (out of a possible 15) who answered the inferential questions correctly in each condition.

Table 4.8: Total numbers of subjects who answered the inferential questions correctly (max. =15)

Trials ->	1	2	3	4
4 TOWER				
A?C	8	10	8	12~
B?D	10	8	11~	13~
5 TOWER				
A?C	8	8	11^	10
B?D*	9	9	9	8
D?E	10	11^	8	8
6 TOWER				
A?C	9	8	10	9
B?D*	9	10	9	7
C?E*	8	9	8	8
D?F	12`	6	7	7

* = critical question (see page 20 for a definition)
^, ~ and `. See below for explanation.

As in Experiment 1, the children could have been correct by chance 50% of the time. Table 4.8 shows that more than half the children produced correct answers on 33 out of the 36 total questions asked.

At the 0.05 probability level, a binomial test showed that a score of 11 or more out of 15 was significantly better than chance. This occurred in none of the critical questions. Furthermore, a significantly better than chance performance was only obtained for the non-critical questions in 3 out of the 8 instances for the 'tower of four' (marked ~ in Table 4.8), 2 out of 8 for the 'tower of five' (marked ^ in Table 4.8) and in 1 out of 8 for the 'tower of 6' (marked ` in Table 4.8).

Comparison with Pears and Bryant (1990)

Pears and Bryant report that all the noncritical inferential questions were being correctly answered significantly more times than would have been expected by chance. The number of correct answers obtained for the critical questions was also significantly above chance in 8 out of the 12 possible questions. The table below shows a direct comparison of significantly above chance answers, between this study and Pears and Bryant (1990).

Table 4.9: Comparison of significantly above chance answers (current study with Pears and Bryant)

Trials ->	1		2		3		4	
	This exp.	P.&B.	This exp.	P.&B.	This exp.	P.&B.	This exp.	P.&B.
4 TOWER								
A?C		sig		sig		sig	sig	sig
B?D		sig		sig	sig	sig	sig	sig
5 TOWER								
A?C		sig		sig	sig	sig		sig
B?D*						sig		sig
D?E		sig	sig	sig		sig		sig
6 TOWER								
A?C		sig		sig		sig		sig
B?D*		sig		sig		sig		sig
C?E*				sig				sig
D?F	sig	sig		sig		sig		sig

* = critical question (see page 20 for a definition)

Comparison with Experiment 1

It is appropriate here to reconsider the results obtained in Experiment 1, where the children were asked to build a queue consisting of 5 people. Only the data from the spatial task has been used for comparison purposes, and only when it was presented first. This was to enable as close a match as possible between task types. In the current study, using only the 'tower of 5' data, no critical questions and only 2 out of 8 noncritical questions received significantly more correct answers than the chance rate. However, scores which represented a 50% success rate were always obtained for both types of questions. This is to be contrasted with the corresponding conditions in Experiment 1 (the queue task). None of the questions (critical or non-critical) were answered at a rate significantly above that expected by chance (with a probability level of 0.05).

Building the towers

Pears and Bryant argued that the successful building of the actual towers was of less importance than the results obtained from inferential questioning. Nonetheless, they reasoned that those children who could make transitive inferences about spatial position have at least one appropriate strategy to help them construct the large tower and should therefore be reasonably successful. For the purposes of this thesis, however, the construction of the tower is crucial, as it gives us an insight into the subjects ability to build a structural task representation by integrating separate premises. Table 4.10 shows the number of children who were able to successfully build the tower in each problem type.

Table 4.10: Number of successful completions of tower (scores out of 15)

Problem	Trial 1	Trial 2	Trial 3	Trial 4	Totals
4 TOWER	10	7	8	12	37
5 TOWER	10	7	8	10	35
6 TOWER	3	5	5	4	17

Because of the importance of this data for the research question addressed in this thesis, it was decided to deviate from Pears and Bryant's analysis methods and carry out an analysis of variance. A one-way analysis of variance was carried out on the total number of correct answers for each of the three problem types. This showed a significant effect of the number of bricks in each tower ($F_{[2, 28]}=20.303$, $p<0.01$). The means for this analysis are given in Table 4.11 below.

Table 4.11: Mean number of successful tower completions (collapsed over trials) Max = 4

Standard deviations are shown in parentheses

Problem	Successful tower completions
4 TOWER	2.47 (0.99)
5 TOWER	2.33 (0.72)
6 TOWER	1.13 (0.92)

Tukey comparisons showed a significant difference between 'tower of 4' and 'tower of 6' ($q=8.18$, $p<0.01$) and also between 'tower of 5' and 'tower of 6' ($q=7.36$, $p<0.01$).

Comparison with Pears & Bryant

This study reported percentages of correctly built towers. These are reproduced, together with the data from this study, converted into percentages, in Table 4.12.

Table 4.12: Percentages of correctly built towers

Problem	This study	Pears & Bryant
4 TOWER	62	83
5 TOWER	58	70
6 TOWER	28	64

Comparison with Experiment 1

Again, it is appropriate to compare this experiment with the results obtained in Experiment 1. Thus, the 'tower of 5' data has been used. Considering the data from Experiment 1, only 5 year old children working with the spatial task has been used for comparison purposes, and only when it was presented first. This was to enable as close a match as possible between task types. 58% of the children were able to build this tower successfully, compared with 8% of the same age children who were able to build the queue successfully.

Serial re-ordering of premises

It must be remembered that the premises were only ordered serially in those individual trials in which the child was unsuccessful in all the inferential questions and also in building the tower (or the queue in Experiment 1).

Inferential questions

Table 4.13 shows the number of questions answered correctly in each trial.

Table 4.13: Number of questions answered correctly (serial presentation)

The numbers in parentheses give the total number of children in each trial for whom the premises were reordered serially. The last column of the table gives the total number of serial re-orderings **when questions were answered correctly** expressed as a percentage of the total number of serial re-orderings i.e. the first number in the totals column expressed as a percentage of the number in parentheses.

Trials	1	2	3	4	Totals	% totals
4 TOWER						
A?C	1(3)	0(3)	2(4)	2(2)	5(12)	42%
B?D	2(3)	3(3)	1(4)	1(2)	7(12)	58%
5 TOWER						
A?C	2(3)	1(4)	2(4)	1(3)	6(14)	43%
B?D*	0(3)	2(4)	3(4)	2(3)	7(14)	50%
D?E	2(3)	0(4)	3(4)	3(3)	8(14)	57%
6 TOWER						
A?C	3(3)	2(5)	2(5)	2(6)	9(19)	47%
B?D*	1(3)	1(5)	3(5)	1(6)	6(19)	32%
C?E*	2(3)	0(5)	1(5)	0(6)	3(19)	16%
D?F	2(3)	1(5)	0(5)	3(6)	6(19)	32%

Only 17 out of a total of 36 data entries are at chance or above. Due to the sample sizes the data was collapsed across trials (see 'totals' column above) and a binomial test was carried out. This showed that none of the questions (critical or non-critical) in any of the three problem types were answered correctly significantly more frequently than would have been expected by chance (0.05 level). Significantly above chance performance (0.05 level of probability) would have occurred if 9 or questions had been answered correctly for the 'tower of four' (where a total of 12 were reordered serially), 10 or questions had been answered correctly for the 'tower of five' (where a total of 14 were reordered serially) and 13 or questions had been answered correctly for the 'tower of six' (where a total of 19 were reordered serially).

Comparison with Pears & Bryant

In the equivalent trials in the Pears and Bryant study, all the questions were answered correctly at a rate above chance. Only one score, that of the critical question in the first trial for the 'tower of 4' fell below the 0.05 level of significance.

N.B. Because Pears and Bryant used serial presentation with a different subject group, the sample sizes were larger (34) and they were also broken down across trials. On carrying out an analysis of variance - 2 (random or serial ordering) x 2 (tower size-5 or 6) x 2 (critical or noncritical question), Pears and Bryant found no significant effect of ordering. A similar analysis was not carried out to compare the effects of ordering in this study, due to the small sample sizes.

Comparison with Experiment 1

As with this study, serial presentation was given only to those children in Experiment 1 who were totally unsuccessful (in any one particular trial). However, due to the greater lack of success in Experiment 1, the sample sizes were much bigger. Table 4.14 reproduces the number of questions answered correctly by the 5 year old children in each of the conditions in Experiment 1.

Table 4.14: Number of questions answered correctly in Experiment 1 (serial presentation)

The numbers in brackets give the total number of children in each trial for whom the premises were reordered serially.

Order	Question	Spatial	Abstract
Spatial first	A?C	26(34)~	28(30)~
	B?D*	25(34)~	25(30)~
	C?E	27(34)~	27(30)~
Abstract first	A?C	21(33)	22(31)~
	B?D*	23(33)~	19(31)
	C?E	26(33)~	22(31)~

* = critical question

It can be seen that all the questions were answered correctly at a frequency above that of chance (50% success rate).

At the 0.05 probability level, a binomial test showed that a score of 19 or more out of 30, 20 or more out of 31 and 22 or more out of 33 or 34 was significantly better than chance. Thus, for 10 out of the 12 questions (marked ~ in Table 4.14), the children were performing significantly above chance. This can be compared with the data from the current study, where a success rate significantly above that of chance was not detected for any of the towers (though this data had been collapsed across trials).

Building of Tower

Again, the sample sizes for this are low. However, 39% (9 out of 23) of the children who attempted to build the 'tower of 4' after serial presentation were successful, 28% (7 out of 25) were successful for the 'tower of 5', and 63% (27 out of 43) for the 'tower of 6'.

Because the data has been collapsed across trials to obtain these figures, and different subjects attempted this task in different trials, it is not possible to perform an analysis of variance in order to ascertain whether or not the effect is significant.

Comparison with Pears & Bryant

Table 4.15 shows the percentage of successful scores for both studies.

Table 4.15: Percentage of successful scores for the current study and Experiment 1

	This study	Pears and Bryant
4 TOWER	39	87
5 TOWER	28	75
6 TOWER	63	71

Comparison with Experiment 1

Table 4.16: No. of successful queue completions- serial re-ordering (Experiment 1)

The numbers in parentheses give the number of successful queue completions which occurred after serial re-ordering expressed as a **percentage** of the total serial re-orderings.

Order	Spatial	Abstract
Spatial first	27 (80)	26 (83)
Abstract first	25 (73)	23 (74)

The average successful percentage rate of queue completions from the above data is 78%. The equivalent percentage from the current study for the 'tower of 5' task is 28%.

4.2.3 Discussion

The results obtained show some improvement in the performance reported in Experiment 1 (the queue task), although comparable results to those observed by Pears and Bryant were not obtained.

The discussion is therefore split into two sections. The first section discusses the results of the current study and compares them with those obtained using the queue task (Experiment 1), whilst the second compares the current results with those reported in Pears and Bryant (1990).

Comparison with the queue task (Experiment 1)

Further details of this study can be obtained by reading the full report. However, the main differences in the method are as follows:-

- the mean age of the comparable group of children in Experiment 1 was two months older than in the current study.

- the subjects in Experiment 1 completed 8 trials altogether. All of these were 5 item tasks, four involved a spatial relation and four involved a more abstract relation. In the current study the children completed twelve trials in all, four each of four, five and six item tasks. All the trials involved the same spatial relation.

- the spatial relation used Experiment 1 was 'in front of' / 'behind', whereas in the current study it was 'above' / 'below'.

- real objects i.e. bricks, were used in this study. Representations of objects, i.e. drawings of people, were used in Experiment 1.

It is not considered likely that the difference in age groups between experiments is having an effect on the results obtained. On the whole, the slightly younger age group show an improvement in performance over the older children. Also, considering the data from 5 year old age group in Experiment 1, there was no significant effect of the relationship used, or of the order in which this was presented (i.e. no practice effects) Indeed, this age group children seemed almost to disregard the explanation of the actual relationship. It therefore seems unlikely that either age or the spatial/abstract relationship change is having an effect, so we need to look elsewhere for an explanation of the observed differences between studies. In view of the points listed above, it is probable that this will be found in any or all of the following:

1. Differing number of trials per subject (8 or 12), resulting in a possible practice effect.

2. Whether or not the number of items in trials was increased for individual subjects (5 items throughout or four, then five, then six). This may have resulted in a 'easy to hard' phenomenon, or a 'learning to learn' effect.

3. The dimensionality of the spatial relation (horizontal or vertical).

4. The type of stimulus used (representations or real objects).

The discussion which follows will attempt to address the differing results obtained in the light of the above.

Inferential questions with random premise presentation.

In Experiment 1, the subjects were performing at a rate significantly below that of chance. In the current study, a 50% success rate i.e. at chance, was obtained for all the questions. It appears that the most likely explanation for this is familiarity with the task materials. The experimenter noted that the children were very happy to guess at an answer when they were working with the bricks. Nursery age children are used to working with colour names, and also using very similar bricks in their counting games, when they actually build them into towers. This has to be contrasted with the actual task which the children undertook in Experiment 1. They are used to dealing with drawings of people, and the concept of bus or supermarket queues is not unfamiliar to them. However, the experimental task context itself was unfamiliar to the children. The experimenter noted that the children were quite reluctant to interact with the drawings, whereas it was difficult to control the play sufficiently in the current study. This unfamiliarity has presumably resulted in much less task involvement, so that the children are unwilling to guess at an answer.

However, although the more familiar task environment resulted in greater involvement, there is still very little evidence of success in the current study when answering the inferential questions. None of the critical questions were answered above chance, and only two out of eight of the noncritical were. It would seem that the children are beginning to occasionally order the items appropriately. Both of the successful occasions were when the children could use an 'end-anchor' to support their problem-solving (i.e. for a non-critical question), but there were other occasions when inferencing using 'end-anchors' was not successful. These limitations could either be due to:

- capacity limitations when processing the premises in order to construct the array.

- metacognitive considerations in that the children are not aware that the construction of a fully ordered array would be an efficient strategy.

- not knowing how to interpret all the premise information so as to correctly construct the array.

Building the array with random premise presentation.

These results address two of the three possible areas of difficulty which have been detailed immediately above. The children were explicitly asked to build an external array using the actual premise items. Thus there should be no problems relating to internal processing limitations, or to metacognitive considerations concerning appropriate strategies. However, only 58% of the children in this study were able to build the tower, so it would seem that the subjects are still experiencing difficulties in correctly interpreting the information which they have been given. Also, there was a significant decrease in performance between **actual construction** of the 'tower of 5' and that of the 'tower of 6'. Thus, even when processing limitation problems have been eliminated by the presence of a real array which can be externally manipulated, performance is still inhibited by a greater number of premises. It therefore appears that the children were making errors when they actually interpreted the premise information, which resulted in incorrect brick orderings. If this was the case, then it is not surprising that the larger towers produced worse performance, as more items give more opportunity for error.

The percentage of successful tower completions in this study is considerably higher than that obtained in Experiment 1 (58% as compared with 18%). There are several possible reasons for this:-

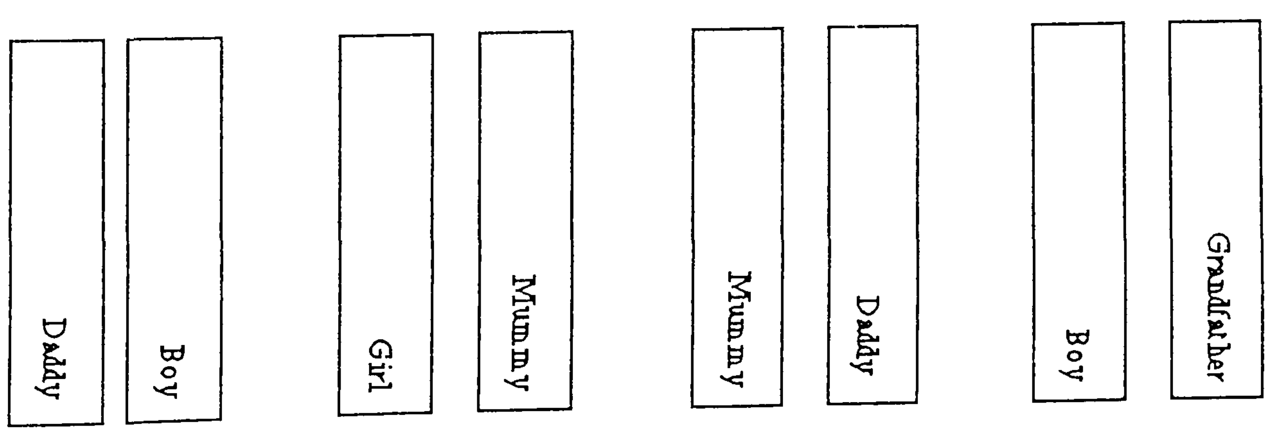
- familiarity with the task environment, as previously discussed. It is felt, however, that this cannot be the sole reason for the observed difference, otherwise we would expect to see performance approaching ceiling in the current study.

- a possible practice or 'learning to learn' effect. It seems unlikely that this is occurring here. The data from Experiment 1 was averaged over eight trials, and there was no evidence of any improvement towards the latter part of a subject's performance. The comparable results from the current study was averaged over trials 5 to 8 (i.e. the 'tower of 5' data). There was no evidence of any practice effect here, as the percentages dropped as the number of trials increase, presumably due to an increased number of items.

- an interaction between the dimensionality of the relationship used and the dimension in which the premises were presented. In both experiments, the premises were actually placed in front of the child in a left to right order (see Fig. 4.13 below).

<u>Fig. 4.13: Actual premise placements (random)</u>

Experiment 1



This study

Black	Red	Blue	Yellow
Red	Blue	Yellow	Green

It could be that the type of placing used in Experiment 1 is interfering more with the construction of the array in Experiment 1. Here, the dimension in which premise information is presented is also the dimension in which the items need to be ordered (horizontal in both cases). This could be causing some interference, and will be investigated in more detail in Chapter 8.

- actual objects were used, rather than drawings. It does not appear probable that this will be having an effect. Children of this age are used to looking at picture books etc., and none of the subjects expressed any surprise or seemed ill at ease when they were initially introduced to the drawings. However, it is intended that future studies in this thesis will investigate differing effects between using drawings or real objects as task items.

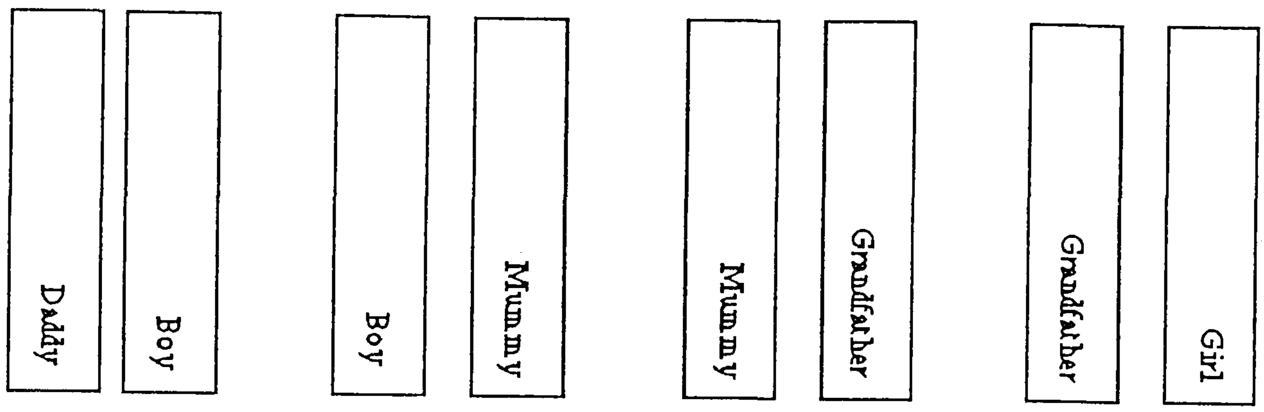
The bricks that were used in this study were similar to 'LEGO' bricks, in that they had to be pushed onto each other. When two bricks had been joined together, it was necessary to exert some force before these bricks could be separated. In contrast to this, the drawings were simply mounted onto pieces of card, so that 'joining' two items meant just placing them next to each other. Thus, provided that the two bricks presented together in a premise were correctly joined when the child began to construct the array, it was more difficult for them to subsequently split the two bricks in error, than it was when dealing with the drawings. It seemed to be the case in Experiment 1 that the children were mistakenly inserting new characters into the middle of a partially ordered queue. In the current experiment, subjects were discouraged from disjoining partially ordered towers by experiencing some physical difficulty. Providing that the pairs of bricks which replicate premise information have been correctly joined, this will result in less splitting of bricks which have been placed into appropriate pairs. This also has some drawbacks, of course, in that bricks which have been incorrectly added to partially ordered arrays are equally difficult to remove!

Inferential questions with serial premise presentation.

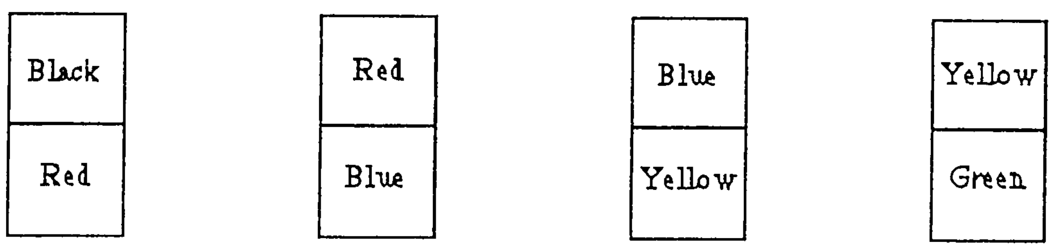
For both this study and Experiment 1, only the subjects who answered all the questions incorrectly and were unable to build the external array received the premises reordered serially. In this study, an improvement in performance was not obtained by the reordering. In Experiment 1, however, performance was improved as 11 out of 12 questions were answered correctly. It seems likely that the reason for this is again the interaction between the relationship used and the dimension in which the premises were presented. With serial presentation, the mode of premise placement used will facilitate performance in Experiment 1. Now, the children only have to eliminate adjacent mid-items in order to be successful. In Experiment 1, this is trivial, due to the already horizontal view of the premises. In the current study, however, further manipulation needs to be carried out, in that the resultant array has still to be ordered vertically. An example of this is given in Fig. 4.14 below.

Fig. 4.14: Actual premise placements (serial).

Experiment 1.



This study



Building the array with serial premise presentation.

The relevant data in the current study shows a large improvement from the 'tower of 5' to the 'tower of 6' (18% to 63%). This would seem to be demonstrating a large practice effect. Serial ordering of the premises does improve performance, but only when the children have become familiar with the actual mechanics of building the tower. This pattern of results was not seen with random presentation, where a large decrease in performance was seen in the 'tower of 6' data. Thus it seems that serial presentation removes some of the cognitive load, but the actual mechanics of the construction also need to be perfected. The relevant results from Experiment 1 showed that 58% were able to build the 5 person queue, when presented with the data ordered serially. This has been averaged out over conditions, as no differential or practice effects were observed. Thus it seems that, for the queue task, the children needed no practice in building the array. In the current study, comparable performance was 18%, but, after practice (even with more items), success rate became equivalent to Experiment 1.

So, in summary:-

With random presentation, children are generally more successful with this task (bricks task) than with Experiment 1 (queue task). However, the children do not show a significant success rate in answering the inferential questions in either experiment. When they constructed an external array, they were more successful with the current study, though their performance was still not at ceiling and was inhibited when the number of items in the array was increased.

The situation is more complex with serial presentation. In the current study, no facilitation of performance was achieved in answering the inferential questions. In Experiment 1, more of the questions were answered correctly at a rate significantly above that of chance, when the premises were reordered serially. If we consider the construction of the external array, the position changes again. Serial presentation of premises in the current study resulted in a large facilitatory practice effect (trials 4 to 8 compared with trials 9 to 12), even though the number of items was also increased with practice (the opposite effect was observed with random presentation, where performance with the 'tower of 6' was significantly worse than the other two). In Experiment 1, serial presentation resulted in an immediate and overall improvement, which was comparable with the eventual performance after practice in the current study.

If we consider the possible explanations for these differences, which were introduced at the beginning of this section, it would seem that we are left with two probable effects :-

1. The interaction of the dimension in which the premise information is presented, and that of the actual relationship used. This is investigated further in Chapter 8.

2. The actual dimension in which the task was situated-vertical relations are easier than horizontal ones. It is difficult to think of a reason why this might be so, however.

Comparison with Pears and Bryant (1990)

The only currently observable difference is the subject group. Pears and Bryant used French children with a mean age of 4 years 6 months. This study used English children, with a mean age of 4 years 9 months.

The results obtained in this study have failed to replicate the performance obtained by Pears and Bryant. In the random presentation of premises, performance was worse in this study when looking both at inferential questions and at external array completion. It also showed no signs of improving through practice (though this could have been cancelled out by the use of more items in the later trials). Serial presentation appeared to have an effect on answering inferential questions in both studies, though Pears and Bryant's subjects were presumably at ceiling, and the children in this study were still not performing significantly above chance on any questions. Building of the external array using serial presentation was showing an improvement due to practice in this study (even with an increased number of items). Pears and Bryant report a slight decrease in performance due to an increase in items. It could well be that, with more practice, the subjects in this experiment would show equivalent performance to that obtained by Pears and Bryant. It seems likely that their performance would then begin to deteriorate due to an increased number of items, though it is unclear when this would happen.

Other than cultural differences, there appears to be no explanation for these results. The better performance was produced by the slightly younger age group. Both groups of subjects were attending the equivalent of state nursery schools, though of course there may well be significant differences in the type of education received.

4.3 GENERAL SUMMARY

In conclusion, the research reported in this chapter has resulted in three main outcomes.

Firstly, it has been shown that series problems are a suitable domain in which to investigate the development of structural representations.

Secondly, we have evidence that 5 to 9 year old children have difficulty in integrating premises into a single array, and that this affects the solving of series problems using analogical reasoning.

Finally, there is some support for the proposition that there are constraints inherent in the task which inhibit the integration of separate relations into a structured representation.

This final point will be investigated further in the following chapters.

CHAPTER 5 : TASK CONSTRAINTS IN THE BUILDING OF STRUCTURAL REPRESENTATIONS (PART 1) CONCERNING INFORMATION ORDERING AND 7 YEAR OLD CHILDREN

The findings from Experiment 1 indicated that the 5 and 7 year old children were unsuccessful in answering inferential questions in series problems, and also, for the most part, in building the external ordered array i.e. the queue of people correctly ordered. This could be completed by considering each pair of items presented as a premise and integrating the information contained in them. Furthermore, the findings were consistent with the evidence already reviewed which shows that:

1. Child and adult novices solve series problems by forming an internal, integrated representation (Trabasso, Riley and Wilson, 1975; Wood, 1969).

2. Children who use concrete item representations to form an external integrated array produce patterns of results which are very similar to results when they solve similar problems without any external representations (Riley, 1976).

Because of the above, it is now accepted that the lack of success in solving series problems demonstrated in Experiment 1 was due to difficulties in constructing a single, integrated representation of the relational information contained in separate premises. This failure to correctly represent the task at the structural level meant that the children were unable to make appropriate mappings between the task and their generalised ordering schema (Cheng and Holyoak, 1985; Halford, 1992). It seemed that the difficulties the children experienced was demonstrated by their comparable difficulty in building the external array. In other words, constraints in the construction of an integrated internal representation were demonstrated in the 'real-world' by corresponding constraints in the construction of a single, integrated external representation.

Internal and external constraints

There is a tradition of research which has investigated the internal constraints on the development of problem solving abilities. Following on from Piaget's structural theory, neo-Piagetians such as Case (1972) and Pascual-Leone and Goodman (1979) have drawn attention to the role of working memory capacity in children's general cognitive development. They maintain that a growth in this capacity accounts for children's increasing ability to successfully process and act upon information, both in everyday situations and in educational environments. In a similar vein, Halford has drawn attention to the function of working memory in the mapping of analogical relations.

Researchers such as Donaldson (1978) have questioned Piaget's structural theory by investigating the effects of external factors which impinge on children's cognitive performance. They have shown that children's perceptions of the task requirements and their beliefs about the experimenter's expectations can mask competence in some Piagetian tasks. It is necessary, therefore, to go beyond characteristics internal to the child in order to fully investigate the underlying reasons for children's performance in any given domain.

There is little reported research concerning the effects of external task constraints* in the domain of relational representation. There is, however, an emerging theoretical perspective in the cognitive domain in which the situated nature of cognition is a central

---

* Researchers such as Crisafi and Brown (1986) and Holyoak, Junn and Billman (1984) have shown how varying the amount of surface similarity in problem analogies affects performance. In these studies, the items which required mapping were varied in order to vary the similarity. In the context of this thesis, external task constraints refer to those factors which are inherent in the way in which the task is structured.

issue. Here, the interaction of the individual with other individuals, with artefacts and with the environment is recognised as crucial. The idea is that cognition is distributed, embodied in the environment and in artefacts and in tasks as much as in the heads of individuals. This has been demonstrated by researchers such as Carraher, Carraher and Schliemann (1985) working with Brazilian street children. Although they were unable to deal with formal arithmetic problems, they were very skilled in dealing with the calculations required to sell candy. The message received from this type of research is that 'cognitive context' has a fundamental part to play in problem solving. It is the interaction of task, environment and the individual which results in performance. It would therefore appear both useful and productive to think about factors which originate from outside the original cogniser in relation to the construction of structural representations. These factors, whilst not discounting the theories of workers such as Halford, will add a further dimension to current knowledge concerning the formation and mapping of systematic relational representations.

The aim of the following set of studies, therefore, is to begin to investigate the task constraints which are operating on the building of structural task representations. Experiment 1 showed that the 7 year old children were able in certain circumstances to correctly construct an external ordered array, although their performance was still error prone. For this reason, a decision was made to use similar aged children in the following experiment.

5.1 EXPERIMENT 3 - ORDERING AND STIMULUS CONSTRAINTS ON THE INTEGRATION OF INFORMATION

Rationale

Experiment 1 identified two initial task constraints which might have been inhibiting successful performance. These were:

1. Elimination of redundant items.

It is conceivable that the children had difficulty in realising that the mid-items were represented in two premises and that in order to successfully construct the array, the two tokens of the same person have to be amalgamated into one character. If the children were trying to construct an array still using two drawings of a mid-item then their performance will be greatly inhibited. Research has shown that photographs are perhaps the easiest graphic representation for young children to understand (DeLoache, 1991; O'Connor, Beilin and Kose, 1981). For this reason it was decided to compare the use of drawings and photographs as types of stimulus, to ascertain whether the use of photographs will facilitate the elimination of redundant mid-items.

2. Order of Premise Presentation.

Although the order of presentation in Experiment 1 was random, there were some types of ordering which the children seemed to find easier than others. This was also borne out by the youngest children in this study, who only appeared able to construct an array with certain types of ordering. It was noticed first of all that the children seemed to consider the premises strictly in order of presentation. It seems that those orderings which required a premise to be placed in front of a partially constructed array (i.e. a left to right violation) gave particular problems. Because of this it was decided to systematically vary the order of premise presentation between conditions. A 5 item array has 24 possible orderings. In view of this, it was decided to reduce the number of items to 4, so that all possible orderings could be covered in one study. The possible premise orderings for such an array are as follows:

1. AB BC CD

2. AB CD BC

3. BC CD AB

4. BC AB CD

5. CD BC AB

6. CD AB BC

1. is a serial presentation, where the relational information does not require restructuring, and can therefore be used as a demonstration.

2. involves only the insertion of the third premise between the first two.

3. and 4. require only one premise to be placed using a left to right violation (L-R) i.e. A in front of BCD or A in front of BC and then join D.

5. and 6. require either two straightforward L-R's or the joining of the second and third premise to form a 3 term array and then an L-R. This would appear to be a complex set of actions which can be carried out in a variety of ways and it will be interesting to see whether a common strategy will emerge.

Taking the above factors into account, the following hypotheses were formulated:

1. There will be an improvement in performance if photographs, rather than drawings are used.

2. The following three types of premise presentation will result in decreasing levels of performance:

Level 'Insert' - Order 2

Level 'L-R' - Orders 3 & 4

Level 'Mixed' - Orders 5 & 6

5.1.1 Method

Design

A two factorial between subject design was used. The two factors were as follows:

1. Method of stimulus presentation. 2 levels:

   Photographs

   Drawings

2. Ordering of premises. 3 levels:

Level I (Insert)	AB CD BC
Level L-R (Left to Right violation)	BC AB CD
	BC CD AB
Level M (Mixed)	CD BC AB
	CD AB BC

Note Serial ordering was used as a worked example.

Thus there were 6 experimental groups.

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task. The 'order of placement' details were also recorded, for both correct and incorrect orderings.

Participants

The participants in this study were 60 mixed ability 7 year olds from state first and primary schools, with predominantly middle class catchment areas (mean age 7 years 1 month, range 6 years 8 months to 7 years 4 months). They were randomly assigned to one of the six experimental groups.

Task description

The relation used was 'behind'. This was based on work done in Experiment 1. 4 item arrays were used, so as to constrain the numbers of possible premise orderings. All the items were side views of women, distinguishable from each other by wearing different coloured jumpers. Thus a completed array would be 4 women in a queue, facing forwards and viewed from the side, with the front of the queue being to the left of the observer's visual field. The orders of colours used were randomly varied, so as to give 6 different orderings of four colours (yellow, green, red, purple). Appendices C and D show examples of completed linear orderings.

Each subject completed six tasks. The two groups who completed level I had six different colour ordering examples of the one type of premise ordering, and the other four groups received three examples of the two orderings covered in their groups. Thus each subject received one example of each of the different colour orderings. The fit of colour order to example was randomly varied between subjects so that each premise ordering was associated with 5 out of the 6 colour orderings.

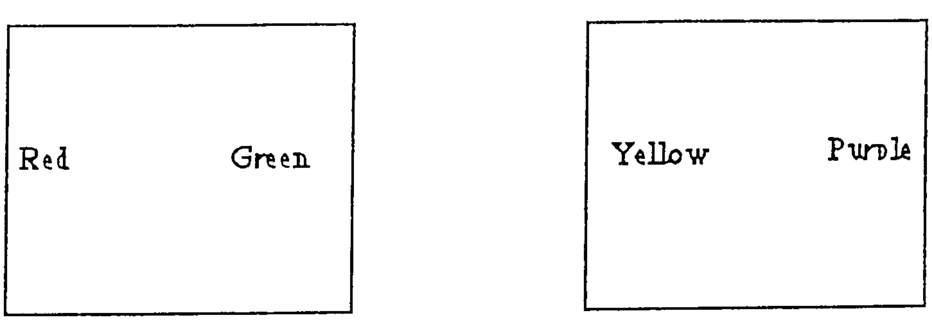
All of the subjects in groups L-R and M were presented with alternate task type orderings and this was counter-balanced between subjects. Thus, **for each group**, half of the subjects received order a b a b a b and the other half received b a b a b a.

Materials

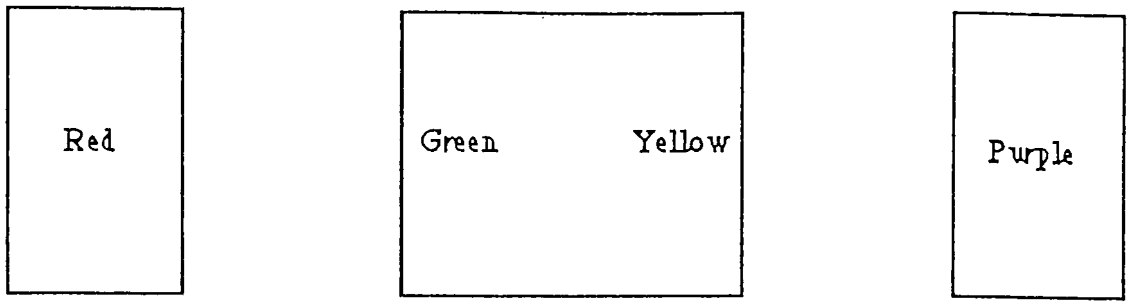
Photographs

Six different colour photographs were taken, each showing the 4 colours in a different order. The women used were all between the ages of 22 and 28 years, dressed in black trousers and the appropriately coloured jumper and with a neutral facial expression. They stood in front of a plain background approximately 20 ins apart and facing forwards, to the photographer's left. Colour prints (6ins x 4ins) were made, and a further 3 copies of each print were obtained. Two of these prints were cut so that the three different pairings of adjacent characters were shown (see Fig. 5.1 below). These three pairings formed the premise information.

Fig. 5.1: Example of 3 character pairings used as premise information



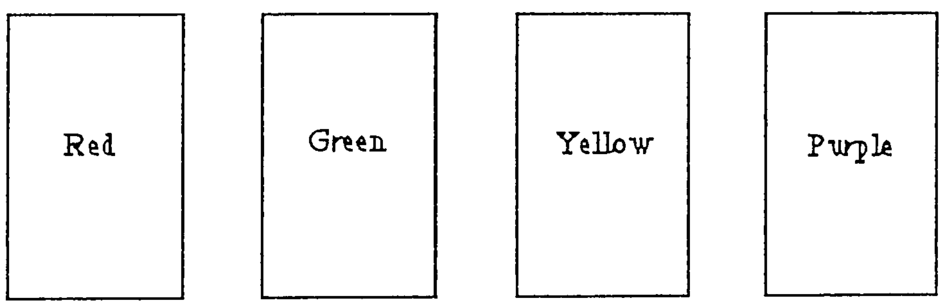
From 1st copy of complete photograph



Discarded From 2nd copy of complete photograph Discarded

The third was cut so that the four separate characters were displayed, each as a 'part photo' (see Fig. 5.2 below). These were given to the subjects to use when constructing the linear ordering.

Fig. 5.2: Example of individual characters used by subjects when ordering the array



Copies of the task materials are given in Appendix D.

Drawings

The same colours, orderings and sizes were used as for the photographs. The figures were drawn by hand, photocopied and then coloured. They were then cut as described above. Copies of the task materials are given in Appendix C.

Procedure

Subjects were tested individually in a quiet room. The experimenter first explained the task, using a complete photograph and serial ordering of premises as a worked example. The children then worked through the six examples, using copies of the individual 4 items to build a concrete array. Final ordering, time taken and order of item placement were recorded by the experimenter.

Wording of instructions

> "I've taken photos/done drawings of four people standing in a queue...Look, like this one... Can you see that they're all wearing different coloured jumpers? You see that the person wearing the <yellow> jumper is first, then the <red> and then the <green> and last of all the <purple> one. I also made some copies....they were exactly the same. I cut one of the copies down the middle, so that the first two people in the queue were in the first half and the last two people were in the second half.......look...like this.....I then got another copy and cut off the first and last people in the queue and threw them away, so that I only had the middle two people left in this part...look...like this..... The other photos/drawings have got the same people in, but they're in a different order in each one. In the puzzles that I want you to do for me, you'll have to work out which order the people are in, because I'm not going to show you the big photo/drawing until the end. You can use these part photos/drawings as clues to help you do the puzzle .....look, I'll show you how........<Use serial ordering> Here are four separate photos/drawings of the four people. We can use these to build up our own queue from the clues. Look at these two people here. The photo/ picture shows us that the person wearing the green jumper is standing behind the person wearing the red jumper...OK? So you use your little photos/drawings to start to build the queue.....now look at this next one.. Who's standing behind the person with the green jumper?....OK, so now you use your photos/pictures again.....Now

use them again when you've worked out what this last photo/picture is showing you......Right, now you've put all of the people in the order which these three part photos/pictures told you about and so you've finished this puzzle. Well done.....There are six more for you to do. They all use the same people but the order of the people will be different and the order of the photos /pictures which I give you might make the puzzles a little bit harder."

WISC-R digit recall scores were also taken from each subject.

5.1.2 Results

Note Each subject completed six experimental trials. However, a large number of the subjects erroneously tried to complete the first trial by recalling the demonstration ordering. For this reason data from the first trial of each subject has been discounted in the following analyses.

WISC scores

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (Insert ordering-photographs)	8.5
Group 2 (Left-to-right violation ordering-photographs)	8.7
Group 3 (Mixed ordering-photographs)	8.5
Group 4 (Insert ordering-drawings)	8.9
Group 5 (Left-to-right violation ordering-drawings)	8.4
Group 6 (Mixed ordering-drawings)	8.8

The average performance on the digit span test for ages between 7 years and 7 years 3 months is a score of 8 (WISC-R Manual, 1974).

Number of correct answers

Table 5.1 shows the mean number of correct answers for each experimental group.

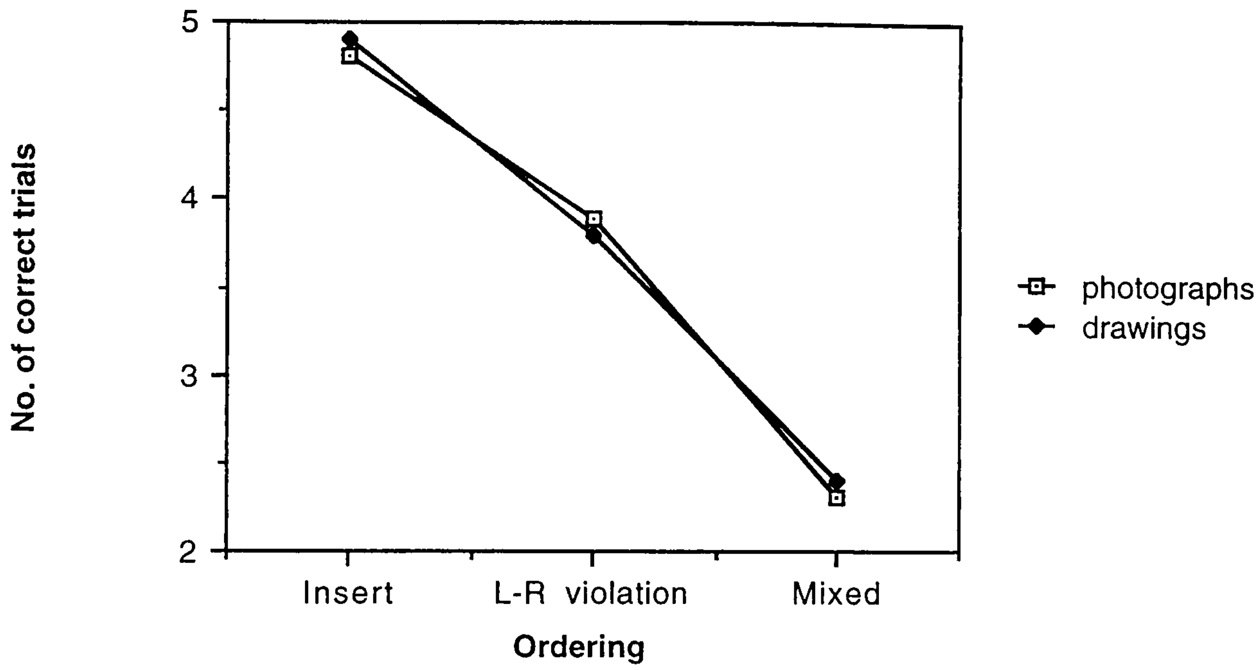
Table 5.1: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

Stimulus	Ordering		
	Insert	L-R violation	Mixed
Photographs	4.8 (0.42)	3.9 (0.88)	2.3 (1.34)
Drawings	4.9 (0.32)	3.8 (1.23)	2.4 (1.17)

An ANOVA [2 (stimulus) x 3 (ordering), both between subjects factors] was carried out on the above data and showed a main effect of ordering ($F_{[2, 54]}$=33.204, p<0.01). Fig. 5.2 shows a graphical representation of the above data.

Fig. 5.2: Effects of stimulus and ordering (number of correct answers)



In view of the lack of a significant effect of the type of stimulus presented, the stimulus data was collapsed. The means for the resultant three groups were 4.85, 3.85 and 2.35 respectively. A one-way analysis of variance was carried out on this data, which showed a significant effect of ordering ($F_{[2, 57]}=34.947$, $p<0.01$). Tukey comparisons were carried out to investigate this. They revealed the following significant effects:

- the subjects were more successful with Insert ordering than they were with L-R violation ($q=4.70$, $p<0.01$).

- the subjects were more successful with L-R violation than they were with Mixed ordering ($q=7.05$, $p<0.01$).

Time taken to complete trials (correct answers only)

Table 5.2 shows the mean time taken (in seconds) for each experimental group.

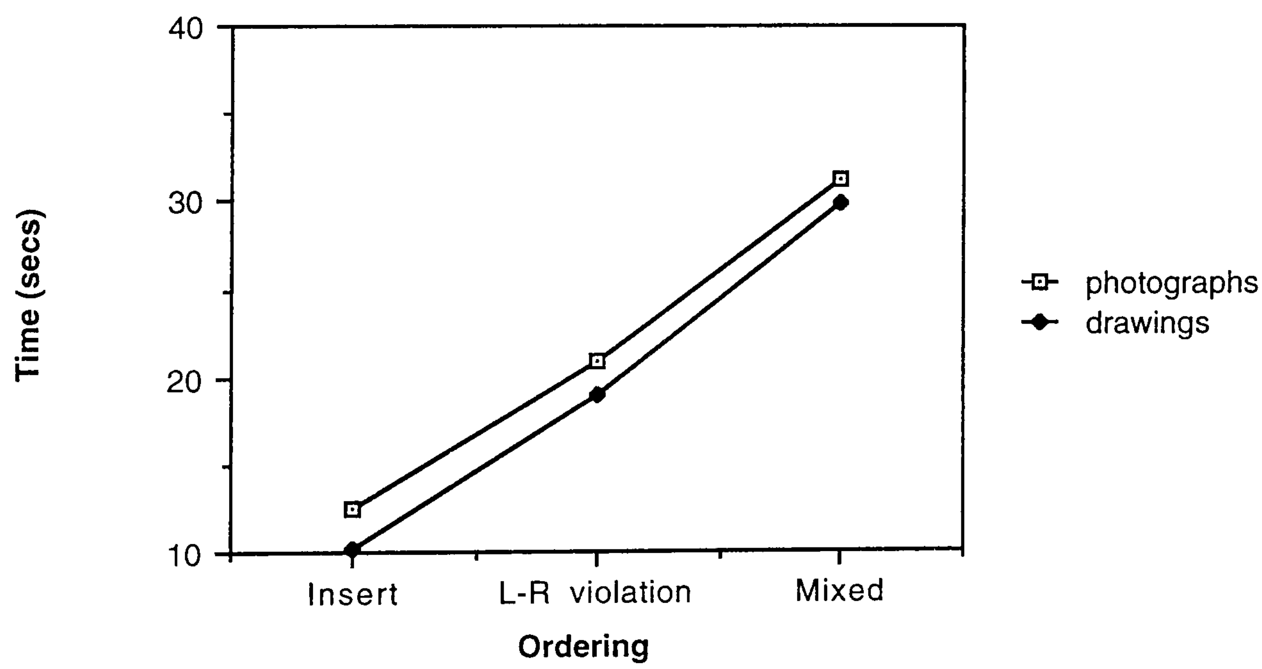
Table 5.2: Mean time taken to complete trials (correct answers only)

Standard deviations are shown in parentheses

Stimulus	Ordering		
	Insert	L-R violation	Mixed
Photographs	12.56 (4.26)	20.80 (3.30)	31.19 (4.61)
Drawings	10.24 (1.77)	18.92 (3.94)	29.89 (3.91)

An ANOVA [2 (stimulus) x 3 (ordering), both between subjects factors] was carried out on the above data and showed a main effect of ordering ($F_{[2, 54]}=131.135$, $p<0.01$). Fig. 5.3 shows a graphical representation of the above data.

Fig. 5.3: Effects of stimulus and ordering (time taken to correctly solve tasks)



In view of the lack of a significant effect of the type of stimulus presented, the stimulus data was collapsed. The means for the resultant three groups were 3.39, 3.66 and 4.21 respectively. A one-way analysis of variance was carried out on this data, which showed a significant effect of ordering ($F_{[2, 57]}=129.370$, $p<0.01$). Tukey comparisons were carried out to investigate this. They revealed the following significant effects:

- the subjects were quicker with Insert ordering than they were with L-R violation ($q=10.04$, $p<0.01$)

- the subjects were quicker with L-R violation than they were with Mixed ordering ($q=12.66$, $p<0.01$).

Order of placement data

Whilst the subjects were completing each task (as before, five were counted in total) the experimenter took details of the order in which they constructed the external array. One record sheet was used for each child, and separate lines were used to record each separate action which the child made. An example of this is given below.

Fig. 5.4: Example record sheet

1. R P

2. R P Y

3. Y R P

4. Y R P G

Note The letters Y, R, P and G refer to the different coloured jumpers worn by the characters. Line 1 signifies that the first action was that of placing the 'red' and 'purple' characters together in one action. Line 2 signifies that the subject left the placing made in line 1 and added the 'yellow' character to the end of the partially built array. Line 3 signifies that the child then removed the 'yellow' character and re-ordered it at the front of the partially built array Line 4 signifies that the subject left the placing made in line 3 and added the 'green' character to the end of the partially built array.

The record sheets were then scored and collated in respect of the following items.

1. It has previously been suggested that the children employ a left to right strategy, with respect to the consideration of premise information, when they construct the array. For example, if they were given BC AB CD they would begin the array by placing items B and C, followed by items A and B (or just by joining on A) and then by C and D, rather scanning the array to select out the items in the correct order i.e. A B C D. The table below gives the mean number of trials in which the left to right construction strategy was employed.

Table 5.3: Mean number of trials where a left to right take-up of information was used (max=5)

Stimulus	**Ordering**		
	Insert	**L-R violation**	**Mixed**
Photographs	4.7	4.1	4.4
Drawings	4.8	4.5	4.6

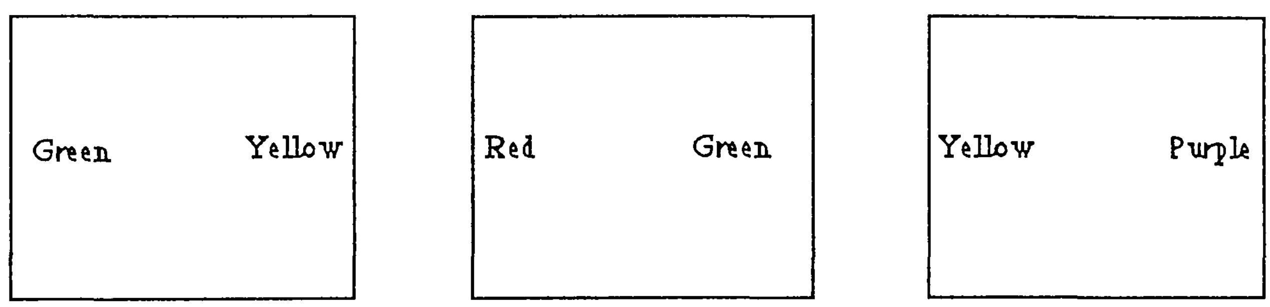
The overall mean number of trials when a left to right take-up of information was used was 4.5 (out of a maximum of 5).

2. The left to right strategy has also been investigated with respect to the position in which the children place the items. Hypothesis 1 suggested that children have difficulty in placing items in **front** of a partially constructed array. In view of this, and also the partial support obtained in the analysis of variance previously described, it was decided to look more closely at the error patterns, and also any spontaneous correction of orderings, to see whether further support could be obtained for this. In order to do this, a distinction has been made between successful trials due to self-correction and those which were successful from the outset. An answer which is correct from the outset is one in which no incorrect 2, 3 or 4 item orderings have been made during the problem solving process, whereas a self-correction involves the spontaneous correction

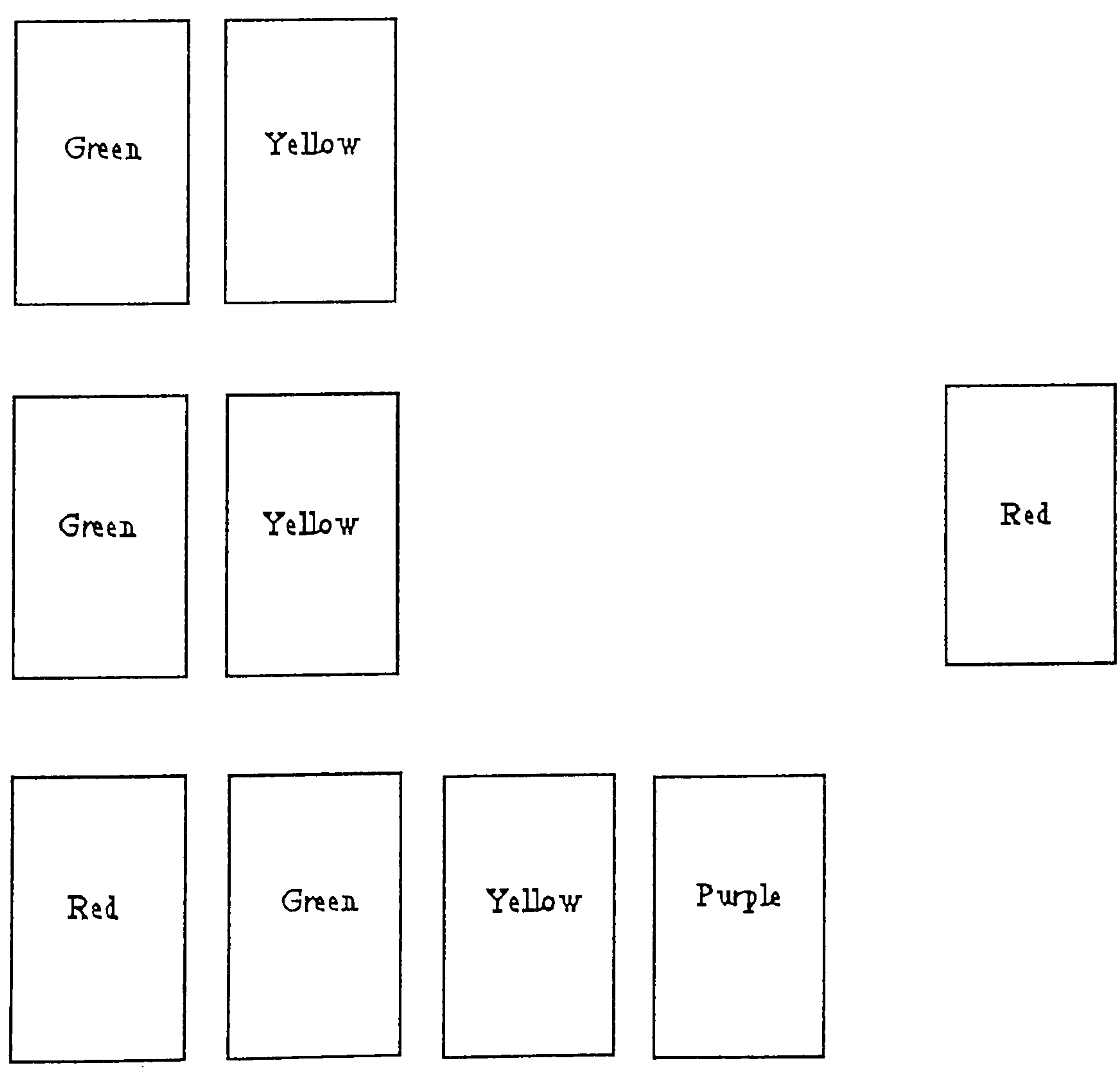
of incorrect 2, 3 or 4 item orderings. An example of this distinction is given in Fig. 5.5 below.

Fig. 5.5: Examples of 'correct from outset' and 'spontaneous self-correction' solutions

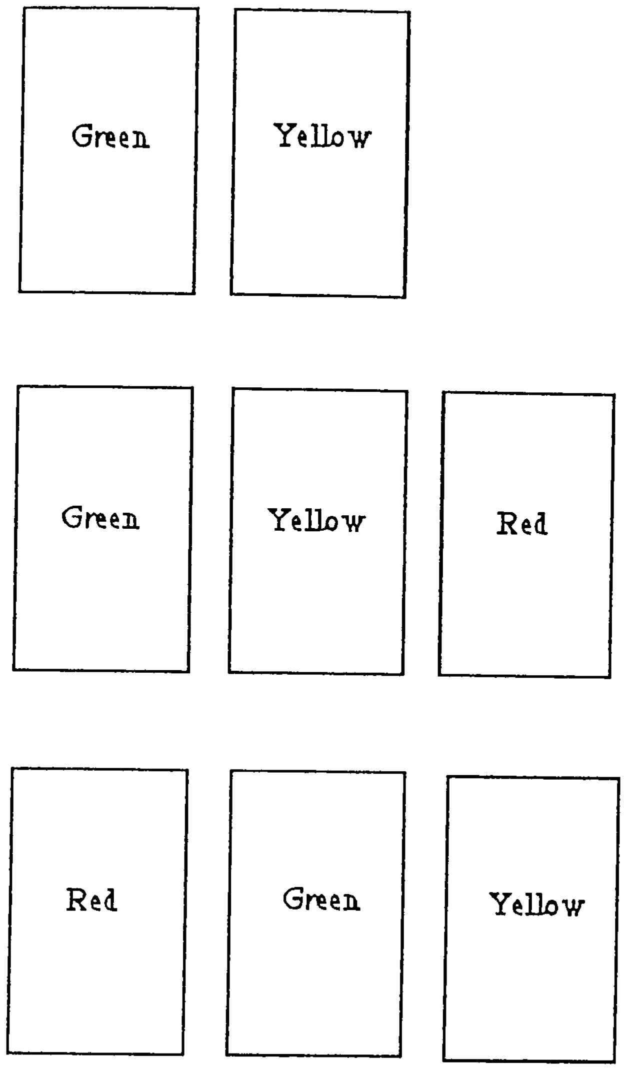
Premise presentation



Correct from outset solution i.e. with an observable break in the array



Self-correction solution i.e. with no observable break in the array



Because of the very small amount of errors and self corrections in the 'Insertion' groups, consideration has been restricted to the L-R violation and Mixed groups.

Table 5.4: Total number of left to right violation errors

The figures in parentheses show the total number of all errors in that group. The two figures in each cell are then converted to a percentage.

	Ordering	
Stimulus	**L-R violation**	**Mixed**
Photographs	8 (11) 73%	19 (27) 70%
Drawings	10 (12) 83%	22 (26) 85%

The overall mean percentage of left to right violation errors is 78%.

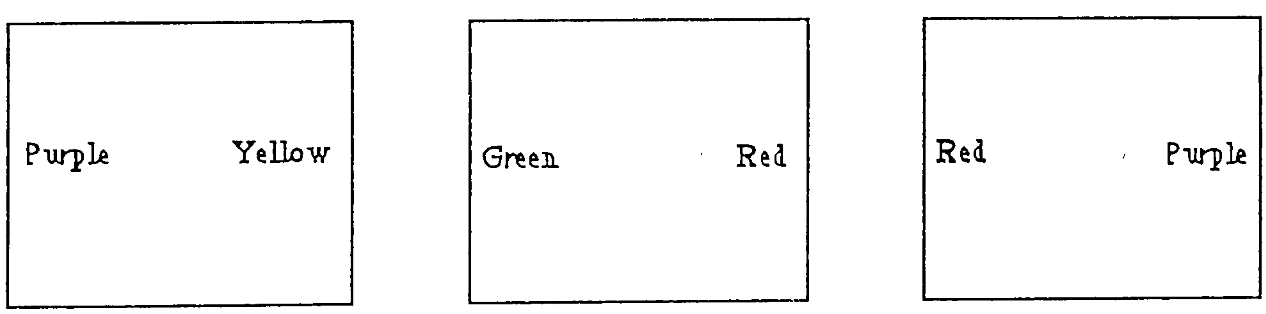
Table 5.5: Total number of successful left to right violation self-corrections

The figures in parentheses show the total number of all errors in that group. The two figures in each cell are then converted to a percentage.

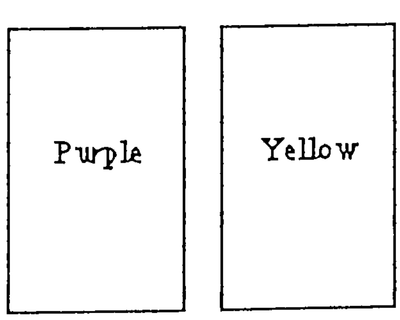
	Ordering			
Stimulus	**L-R violation**		**Mixed**	
Photographs	13 (16)	81%	12 (17)	71%
Drawings	14 (16)	88%	15 (20)	75%

The overall mean percentage of successful left to right violation self-corrections is 79%.

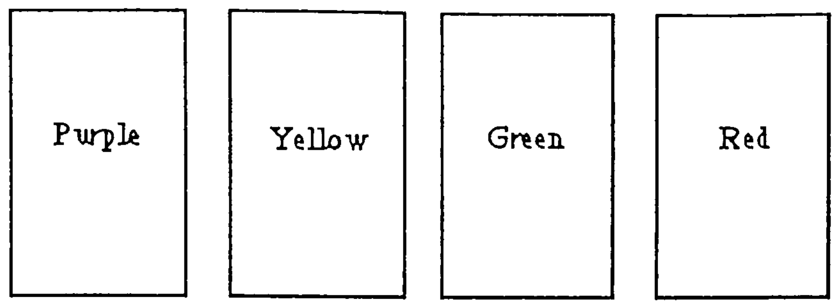
3. The order of placement data was also scrutinised for instances of 'premise pair' splits. Consider the premise ordering below.



As subjects invariably use a left to right take-up of information, a typical ordering would be as follows

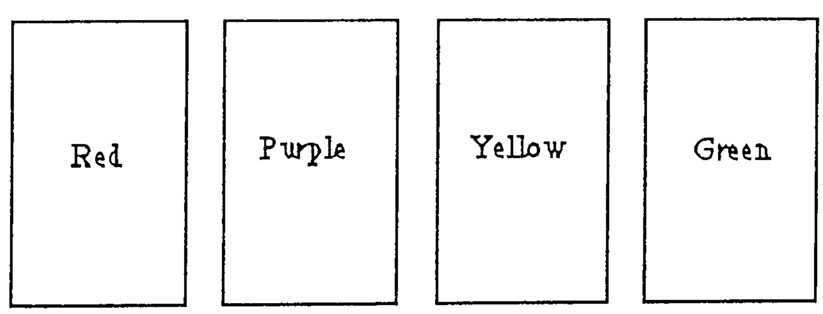


(from 1st premise)



(from 2nd premise)

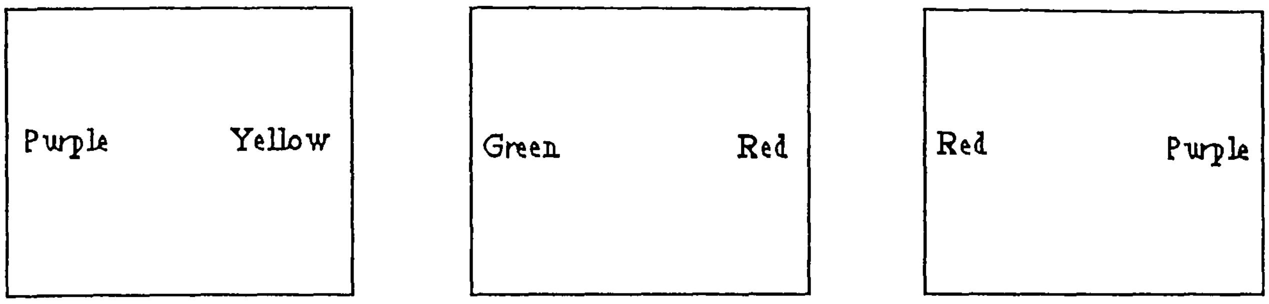
At this stage, typically the subject would indicate that they had finished ordering the queue (this would of course be an incorrect ordering), or they would look at the 3rd premise and so reconsider their answer. It seemed that many of the children re-ordered the array by moving only the last character (red jumper) in order to comply with the 3rd premise. The children would then indicate that they had completed the task, with the queue being ordered as follows (still incorrect).



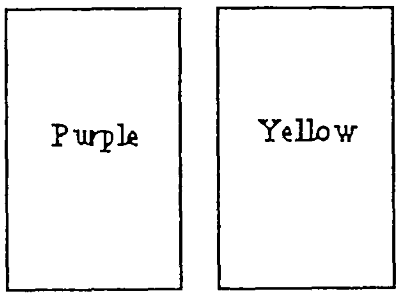
However, in re-ordering as above, the child has failed to take account of the information given in the 2nd premise, in that the red jumper is now no longer behind the green jumper. This type of error has been termed a 'premise pair split' because the subject has erroneously split characters which were given as immediately next to each other in the premise information.

In contrast to this, a child who re-ordered by taking note of the information given in all three of the original premises would carry out the following placements.

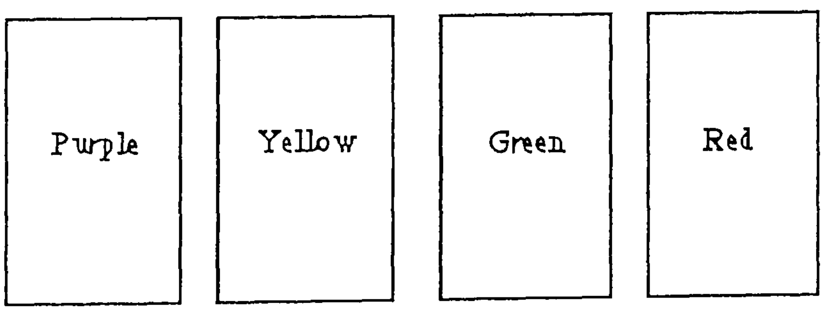
Premise presentation



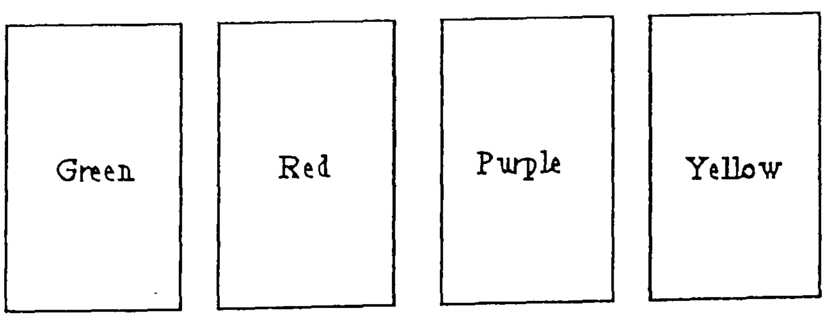
1st placement of items



2nd placement of items



Re-ordering following consideration of third premise - note that red and green jumpers have been re-ordered in one action due to the information given in the second premise.



In view of this, the CD AB BC task type was examined for evidence of splitting of a previously ordered premise pair.

Table 5.6 : Total number of 'previously ordered premise pair' splits - (CD AB BC) tasks

The figure in parentheses shows the total number of all initially correct, self-corrections and errors in that group. The two figures in each cell are then converted to a percentage.

Stimulus	Init. correct		Corrections		Errors	
Photographs	0 (1)	0%	5 (5)	100%	14 (14)	100%
Drawings	0 (2)	0%	8 (9)	89%	19 (19)	100%

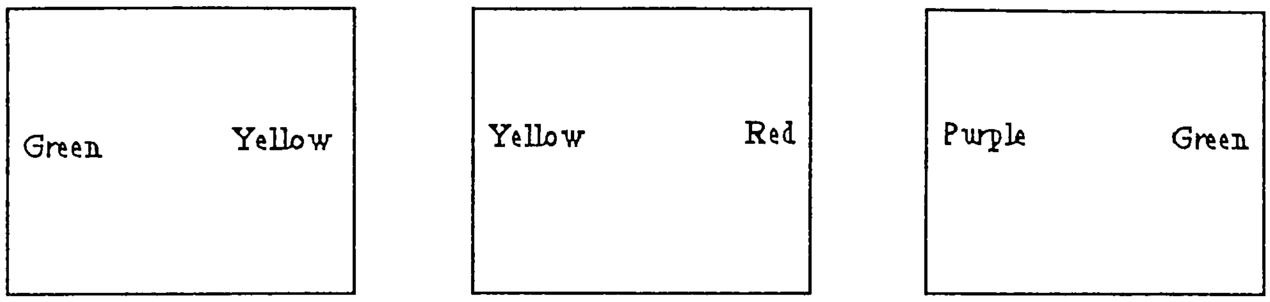
5.1.3 Discussion

Elimination of redundant items

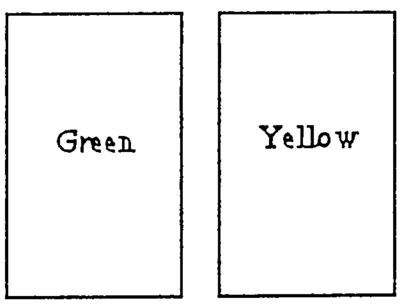
The first experimental manipulation, that of comparing photographs with drawings as stimulus material, was designed as a means of testing whether or not the participants had problems in accepting that the two middle character items were each represented twice in the original premise information. In order for a correct solution to be reached, it was necessary to eliminate the redundant character tokens. However, there was no significant difference between the two conditions. Also, there were no comments from the subjects which indicated that they were confused by there being only one drawing of each character available to them when ordering the array. This is illustrated in Fig. 5.6 below.

Fig. 5.6: Example of a potential situation where subjects might have been confused

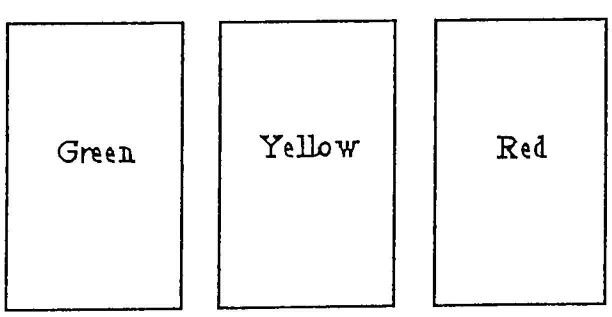
Premise presentation



1st placement



2nd placement (note that only the character (the red jumper) is added, although the second premise has two characters, (the yellow jumper and the red jumper).



We must now consider whether the experimental manipulation was an adequate test of the hypothesis concerning the redundancy of mid-item information. It could be that the current lack of evidence in support of this proposition is due to the experimental procedure, as the subjects were only given one token of each character to work with. Thus the providing of a single token might have been causing confusion when the array was being ordered, if the subjects were not aware that one token for each mid-character had to be eliminated. This could have contributed to a lack of understanding of the task and therefore also to the children's lack of success. Although we have evidence from the current experiment that the order of premise presentation significantly affects performance, inhibition might also be occurring as a result of an erroneous belief that duplicate tokens of the mid-characters (the red and yellow jumpers in the example given above) were needed. An obvious way to test this would be to compare the performance of children who had only one token of each character available, with the performance of those who had more than one. A further experiment was therefore designed in order to carry this out, and is reported in Chapter 7.

Order of premise presentation

It can be seen from the results section that there is empirical support for hypothesis 2. This showed that different premise orderings had a differential effect on performance. Both types of data i.e. number of correct solutions and time taken to reach a correct solution, provided a similar pattern of results. Task difficulty was primarily affected by the number of left to right violation of premise orderings which were required.

It seems that there are at least three levels of 'order of premise presentation' difficulty for a four item array. Thus the representation of structural systemacity in series problems is not an obvious initial process for 7 year old children. Even when the similarity between the task domain and the type of spatial reasoning required is made very explicit, children still need practice in recognising and using item relationships.

The problems with L-R violation are interesting. Obviously, these children are literate, and so this could well be a cultural phenomenon produced by interaction with text. The study used a left to right presentation of premises, which might have encouraged the children to build as they would construct written text. Therefore, it may prove enlightening to present premises vertically. This will be addressed in a further study. Also, the data from the Mixed group is inconclusive. Performance is significantly worse than the L-R group, but it is not clear why. There are several types of ordering patterns which could be used in these problems. In view of the results obtained in the other two conditions, it seems likely that the subjects were experiencing difficulty due to the number of left to right violation of premise orderings which were required. On the other hand, the mixed condition actually contains two subtypes of task - CD AB BC and also CD BC AB. These are presented alternately. The necessity to change strategies in order to deal with the two subtypes may be contributing to the difficulty. It would be interesting therefore, to run a further two level, between subject experiment, which compares CD AB BC with CD BC AB. Also, it could be that the former sub-type is more difficult than the latter, due to the absence of a common item in the first two premises (sub-type CD BC AB has the C item common to the first two premises). This design would allow a direct comparison to be made. Several of the children in the Insert ordering (AB CD BC) commented that there was no need for the last premise and were happy to join AB to CD without the confirmatory premise BC. However, although ordering CD AB BC produced many errors, examination of the placement details revealed a partial reordering, often to BCDA, which showed some consideration of the last premise. This could be due to an added conceptual understanding provided by the alternative task in the Mixed group.

It is also possible that one of the reasons for the 7 year olds tendency to employ inappropriate left to right strategies concerns the nature of the task. This was unfamiliar to the children, and some of them appeared to be having difficulty in grasping what was required of them. This could be due to the to the lack of context contained in the task. It was explained to the subjects that they were solving puzzles about working out the right order for people in a queue, but it could be that this explanation did not provide the children with a good reason for trying to understand what was required of them.

There is evidence (Donaldson, 1978) that children's performance can be facilitated by the use of contextualised 'real-world' tasks which are meaningful and motivating. In view of this, it was decided to carry out a study to compare performance on two series problems. One of these will be the same problem as that encountered by the children in the current study. The other will use same horizontal spatial relationship, but will require the child to work with everyday well-known toys, and will provide a supporting narrative. This study is reported as Experiment 4.

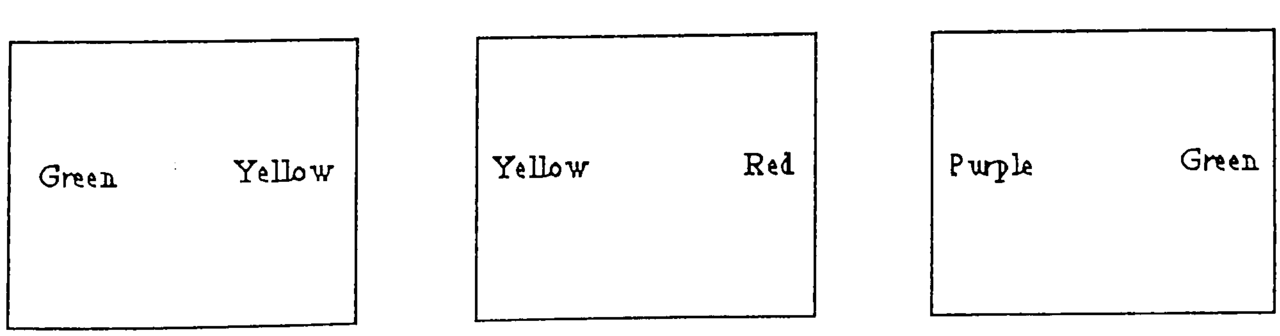
Order of premise placement data

This data has shown that an average of 79% of the subjects errors and self corrections were due to them erroneously placing a character token at the end (i.e. to the right) of a partially built array when they should have been placing them at the front. This of course provides further support for the hypothesis that children's performance in integrating premise information into one array is inhibited when they are required to violate a left to right ordering pattern.

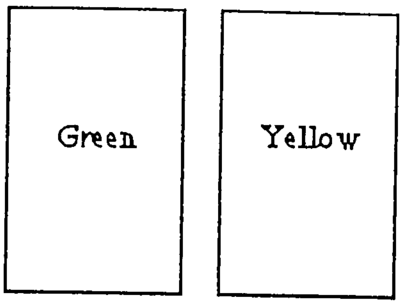
We can also see that for an average of 4.5 out of 5 trials, the subjects always used a left to right takeup of information. This means that they scanned and acted on the premise information in the order in which it was presented to them, that is in a left to right order. This provides us with a reason as to why the children need, for certain types of orderings, to violate their left to right rule. Fig. 5.7 below gives an illustration of this.

Fig. 5.7: Example of left to right violation

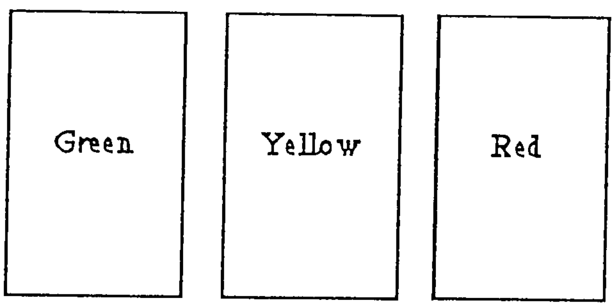
Premise presentation



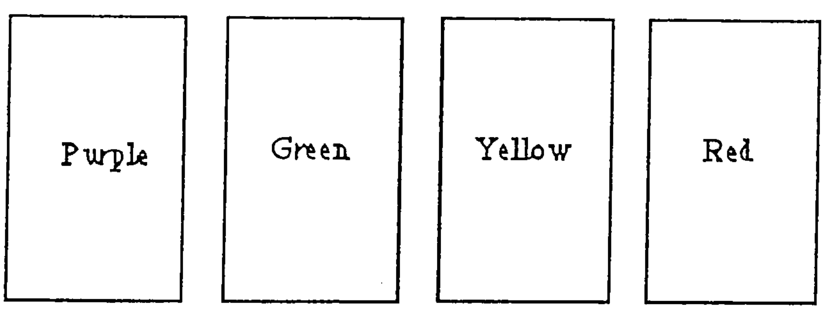
1st placement



2nd placement

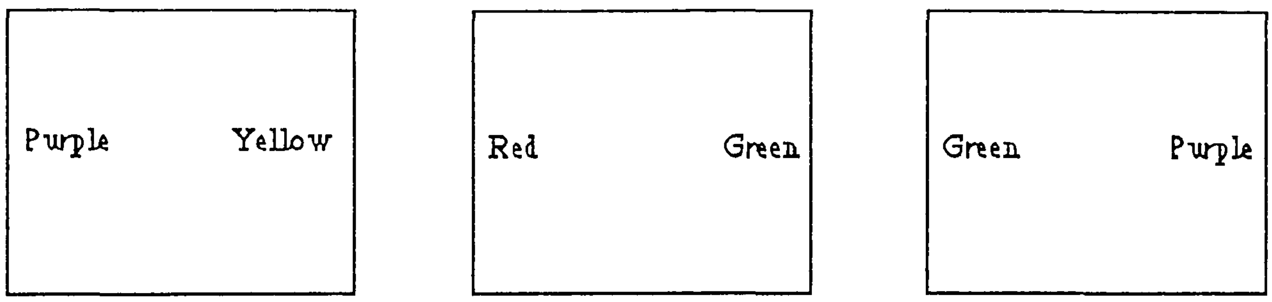


3rd placement

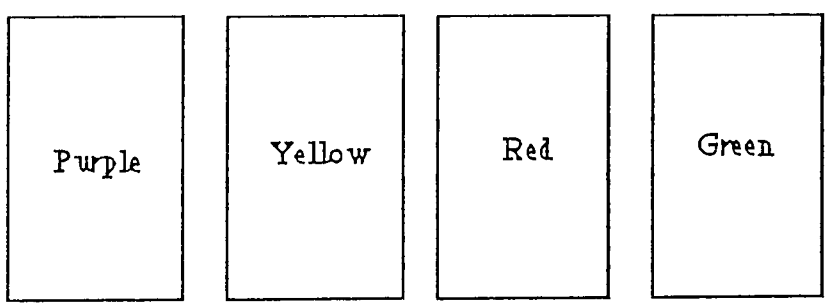


Because the premise have been dealt with strictly in the order in which they were presented, the 3rd placement required that the new character be placed at the front (i.e. to the left) of the partially ordered array. This therefore required violation of the left to right ordering rule. If however the children had searched the premises in order to find the first 'end-anchor', i.e. the first character in the array, and had continued by searching for the 2nd, 3rd and then 4th characters, then no left to right violations would have been necessary.

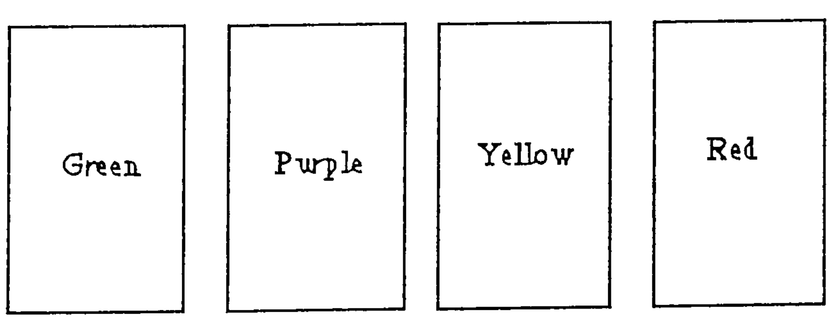
The order of premise placement data also appears to be showing that a further difficulty is in realising that an item pairing given in an original premise should not be separated. Initally, the children appear to only consider one premise at a time. Consider the premise ordering below.



An common initial (incorrect) ordering is:



Subjects would usually reorder this as:



This is because they then view the third premise. This of course has only involved the re-ordering of one item (green jumper) rather than both the characters which were linked in the second premise (the red jumper and the green jumper), which would result in a correct solution. This is of relevance to the situation observed in Experiment 2 (Pears and Bryant replication). The bricks used were ones which were moulded to form a join (similar to 'LEGO' bricks) so it was much easier to move items as pairs rather than individual bricks. It is interesting to speculate on why the 'split pairing' is occurring. It could be that the children do not fully understand all the information inherent in the initial premise, or just that they forget to utilise this information when re-ordering, as they are no longer directing their attention towards it.

Conclusion

To summarise, therefore, it appears that the inhibition in performance in the construction of a structural task representation demonstrated by the 7 year old children in Experiment 1 was probably not due to problems concerning the elimination of redundant mid-items. A further experiment is however necessary to investigate this factor (see Chapter 7).

The current experiment has provided evidence that the 7 year olds experience difficulty in constructing an ordered array and this is dependent on the order of premise presentation. However, it could be that this effect is due, at least in part, to the child working with an unfamiliar task. Experiment 4 will therefore investigate whether this effect is also found when children are working with familiar toys and provided with a supporting context.

We have collected some evidence that the 7 year olds difficulty in this study stems from an inappropriate application of a left to right ordering strategy. This occurs because they act on premise information strictly in the order in which it is presented to them. Experiment 1 demonstrated that 9 year old children were more successful that 7 year olds in constructing an ordered array. A further study (Experiment 5 in Chapter 6) needs to be carried out to ascertain whether this improvement in the older children's performance was due to the surrender of the inappropriate left to right ordering strategy used by the 7 year olds in the current study.

5.2 EXPERIMENT 4 - THE EFFECT OF A FAMILIAR TASK ON THE INTEGRATION OF INFORMATION

There is evidence (Donaldson, 1978) that children's performance can be facilitated by the use of contextualised 'real-world' tasks which are meaningful and motivating. In view of this, it was decided to conduct the following study, which compares performance on series problems. Both use the same horizontal spatial relationship, but one uses equivalent stimulus materials to Experiment 3, whilst the other uses everyday well-known toys, together with a supporting narrative.

5.2.1 Method

Design

A one factorial between subject design was used.

The two levels of 'type of stimulus material' were 'toy train' (Condition 1) and 'drawings of queue' (Condition 2).

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task. The 'order of placement' details were also recorded, for both correct and incorrect orderings.

Participants

The participants in this study were 10 mixed ability 7 year olds (mean age 7 years 1 month, range 6 years 7 months to 7 years 6 months) from state first and primary schools, with predominantly middle class catchment areas. The subjects were from the same schools as those in Experiment 3, and were assigned to Condition 1. The data for Condition 2 (drawings of queue) was taken from Experiment 3 (Condition Left to Right violation and Drawings).

Task description

Condition 1

As in Experiment 3, except that the task items were four different coloured carriages from a toy train. The carriages could be coupled together and decoupled. The premise information was presented by using drawings of two carriages coupled together (3 drawings for each 4 item task). The drawings were approximately 6 ins x 2.5 ins, and the carriages were coloured appropriately. Each carriage had a character seated in it, facing towards the front, so that the carriages had obvious backs and fronts. An example of premise information is at Appendix E.

All the subjects worked with the following task orderings:

BC AB CD

BC CD AB

This corresponds to level L-R (Left to right violation) used in Experiment 3.

Condition 2

As in Experiment 3, except that all the subjects worked with the following task orderings:-

BC AB CD

BC CD AB

This corresponds to level L-R (Left to right violation) used in Experiment 3.

Materials

Condition 1

A toy 'LEGO' train, with an engine and four different coloured carriages, with a character facing towards the front of each carriage, was used as the 'ordered array. The children were handed the carriages and asked to put them in the right order. Premise information was presented using 3 drawings of pairs of carriages, coloured appropriately, for each ordering. Appendix E shows an example of premise information and the appropriately ordered array.

Condition 2

As in Experiment 3, but using the drawings only. Appendix C shows an example of premise information and the appropriately ordered array.

Procedure

Condition 1

As in Experiment 3, except the children were handed the four carriages to work with. The wording of the instructions was as follows:

"I'm going to tell you a story about something that happened in a nursery school, just like the one you've got here. The teacher wanted the young children to learn their colours. She decided to make it more fun by using a toy train, just like this one (show engine and four coupled carriages) Look. first there's a <red> carriage, then a <blue> one, then <green> and last it's a <yellow one> She asked the children to choose any two carriages, and make a drawing of them. They could choose any two, but they had to colour them in the right colours and put them in the right order. So you see that some children choose the first two <red then blue>, some chose the next two <blue then green> and some chose the last two <green then yellow>. Whilst the children were busy working, the teacher went out of the room. Some of the children started to play with the train, and they took the carriages apart, like this <separate the engine and all the carriages>. They then realised that they needed to join all the carriages together in the right order before the teacher came back, otherwise she would know that they hadn't been working. They couldn't remember the right order, so they had to use the drawings which they'd made as clues. Look, you can do it like this <use serial ordering and work through example as in Experiment 3>. I want you to imagine that you're in this class, and you have to use the drawings to work out the right order for the carriages. There are six to do. They all use the same carriages, but the order that they should be in will be different and the order that I give you the drawings in might make it a little bit harder."

Condition 2

As in Experiment 3.

WISC-R digit recall scores were also taken from each subject.

5.2.3 Results

Note. Each subject completed six experimental trials. However, a large number of the subjects in the first study erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier studies the first trial from each subject has been discounted.

WISC scores

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (toy train)	8.2
Group 2 (drawings of queue)	8.4

The average performance on the digit span test for ages between 7 years and 7 years 3 months is a score of 8 (WISC-R Manual, 1974).

Number of correct answers

Table 1 shows the mean number of correct answers for each experimental group.

Table 5.7: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

Type of stimulus material	
Toy train	**Drawings of queue**
3.6 (1.17)	3.8 (1.23)

A one way ANOVA was carried out on the above data and showed no significant effect of the type of stimulus material ($F_{[1, 18]}=0.138$, $p>0.05$).

Time taken to complete trials (correct answers only)

It was felt that the children may take longer to join carriages together (Condition 1) than they would to place drawings in order (Condition 2). In view of this, 10 children were asked to complete three trials each of the Insert ordering (Experiment 3), which has previously been established as producing performance approaching ceiling, but whilst working with the toy train. However, there were no significant differences between the mean scores obtained and those of the children in Experiment 3 who completed the Insert condition using the drawings (mean scores of 11.3 seconds and 10.24 seconds respectively, $F_{[1, 18]}=3.163$, $p>0.05$).

Table 5.8 shows the mean time taken (in seconds) for each experimental group.

Table 5.8: Mean time taken to complete trials (correct answers only)

Standard deviations are shown in parentheses

Type of stimulus material	
Toy train	**Drawings of queue**
20.83 (2.39)	18.92 (3.94)

A one way ANOVA was carried out on the above data and showed no significant effect of the type of stimulus material ($F_{[1, 18]}=1.711$, $p>0.05$).

Order of placement data

Whilst the subjects were completing each task (as before, five were counted in total) the experimenter took details of the order in which they constructed the external array. These were then examined with a view to identifying any differences in the use of left to right strategies, as reported in Experiment 3.

Table 5.9: Mean no. of trials where a left to right take-up of information was used (max=5)

Standard deviations are shown in parentheses

Type of stimulus material	
Toy train	**Drawings of queue**
3.8 (1.23)	4.5 (0.97)

A one way ANOVA was carried out on the above data and showed no significant effect of the type of stimulus material ($F_{[1, 18]}=1.995$, $p>0.05$).

Table 5.10: Total no. of Left to Right violation errors

N.B. the second figure in each cell shows the total number of errors in that group

Type of stimulus material	
Toy train	**Drawings of queue**
12 (15)	10 (12)

Table 5.11: Total no. of successful Left to Right violation self-corrections

N.B. the second figure in each cell shows the total number of self-corrections in that group

Type of stimulus material	
Toy train	**Drawings of queue**
13 (14)	14 (16)

5.2.4 Discussion

The lack of any significant differences between conditions for both the number of correct answers and the time taken to complete the trials means that there is no evidence to support the experimental hypothesis i.e. that working with a situated, concrete task facilitates performance.

It would seem, therefore, that the provision of a 'real task' does not affect performance on series problems for the 7 year old children. This is further borne out if we consider the order of premise placement data. The toy train condition appears to be producing just as many left to right violation errors and self-corrections as the queue task, and there is no significant difference in the left to right take-up of information.

There are three possible reasons for this:

1. Inappropriate choice of 'real task'

It could be that, if a different real task had been selected, facilitation of performance would have occurred. This appears unlikely, however. The children who used the toy train certainly seemed more 'at home' with the task, and there were no puzzled expressions or inappropriate questions, as there were with the children who used the queue task. Also, there were no indications from the order of premise placement data of any lessening of the left to right strategies which have been observed in the earlier experiments. It could be that the presentation of premise information using real objects, rather than drawings, would have an effect, although it is not obvious exactly how this would facilitate performance. On the contrary, asking the children to order using duplicates of the objects used in the premises might cause some confusion, and it is difficult to think of a cover story which would be convincing and motivating.

2. Time taken to couple the carriages

It was considered possible that the children would take longer to join the carriages together than they would to place the drawings in order. However, there was no difference in times between the performance on the train and queue tasks when the children were using the Insert condition (where performance approaches ceiling). Nonetheless, it could still be that the decoupling of incorrectly ordered carriages, when the subjects self correct errors, takes longer than the re-ordering of the queue task. This might explain the lack of difference in times between conditions, whereby any facilitation in performance gained by the real task is being offset by the longer time needed to re-order potential errors.

If this is the case, the train task ought to be producing significantly more correct solutions. As this did not occur, it seems unlikely that the potentially more difficult manipulation of the task materials is having any effect.

3. Constraints on performance are not due to familiarity or motivational factors.

This would appear to be the most likely explanation, if we consider the current data and the results from previous studies. It seems that 7 year old children are experiencing difficulties with these types of tasks due to the inappropriate use of a left to right strategy, which is still used when they work with familiar task materials within a motivational context. The evidence from the current study, together with that from Experiment 3, suggests that this rule is very robust for a large proportion of 7 year olds.

Experiment 1 compared the performance of 5, 7 and 9 year old children in the domain of series problems and showed that only the 9 year old children were able to answer the inferential questions and correctly order an external array for a significant proportion of the tasks. The two studies reported in this chapter have shown that the degree to 7 year old children were required to violate a left to right ordering strategy significantly affected successful performance when constructing the external array. Details were also taken of the order in which the children placed the individual items. Collation of this data showed that the children relied almost exclusively on a left to right ordering strategy in the following ways:

- they employed a left to right take up of information.

- spontaneous self corrections and errors involved erroneously placing an item at the end of a partially ordered array, rather than in front.

The following chapter will discuss similar experiments, but carried out with 9 year old children. We know that there is a difference in overall performance between 7 and 9 year old children (Experiment 1), and that the 7 year old's difficulty stems from an inappropriate L-R ordering strategy (Experiment 3). We will therefore be able to find out whether the age difference in performance observed in Experiment 1 is due to differences in the application of this strategy.

CHAPTER 6 : TASK CONSTRAINTS IN THE BUILDING OF STRUCTURAL REPRESENTATIONS (PART 2) CONCERNING INFORMATION ORDERING AND 9 YEAR OLD CHILDREN

Experiment 1 compared the performance of 5, 7 and 9 year old children in the domain of series problems and showed that only the 9 year old children were able to answer the inferential questions and correctly order an external array for a significant proportion of the tasks. The 7 year olds could order the array but only after practice and were unable to answer the inferential questions (see Chapter 4).

A further study (Experiment 3, discussed in Chapter 5) investigated the performance of the 7 year old children in order to get some purchase on their initial inability to carry out this ordering. It was found that the degree to which they were required to violate a left to right ordering strategy significantly affected successful performance when constructing the external array.

Details were also taken of the order in which the children placed the individual items. Collation of this data showed that the children relied almost exclusively on a left to right ordering strategy in the following ways:

- they employed a left to right take up of information.

- spontaneous self corrections and errors involved erroneously placing an item at the end of a partially ordered array, rather than in front.

The experiments reported in this chapter investigate the existence of an inappropriate left to right ordering strategy in the performance of 9 year old children.

6.1 EXPERIMENT 5 - THE EFFECT OF ORDERING ON THE INTEGRATION OF INFORMATION FOR 9 YEAR OLD CHILDREN

Rational

The current experiment replicates Experiment 3, but using 9 year old children. We know that there is a difference in overall performance between 7 and 9 year old children (Experiment 1), and that the 7 year olds difficulty stems from an inappropriate L-R ordering strategy (Experiment 3). This study will enable us to find out whether the age difference in performance is due to differences in the application of this strategy. If the reason for the older children's improvement in overall performance was due to the realisation that left to right ordering is inappropriate in the L-R Violation and Mixed orderings, then we would expect that these specific conditions will both lead to improved performance for the 9 year old when compared with the 7 year olds.

6.1.1 Method

Design

A one factorial between subject design was used.

The three levels of premise ordering were as follows:

Level I (insert)	AB CD BC
Level L-R (left to right violation)	BC AB CD
	BC CD AB

Level M (mixed) CD BC AB

CD AB BC

NB Serial ordering was used as a worked example.

Experiment 3 also looked at the effect of a further factor (type of stimulus material). However, no significant effects were found, and so this is not being investigated in the current study.

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task. The 'order of placement' details were also recorded, for both correct and incorrect orderings.

Participants

The participants in this study were 30 mixed ability 9 year olds (mean age 9 years 1 month, range 8 years 8 months to 9 years 5 months) from state first and primary schools, with predominantly middle class catchment areas. The subjects were from the same schools as those in Experiment 3. They were randomly assigned to one of the three experimental groups.

Task description

As in Experiment 3

Materials

As in Experiment 3, but using the photographs only. An example of the task materials is given at Appendix D.

Procedure

As in Experiment 3.

6.1.2 Results

N.B. Each subject completed six experimental trials. However, a large number of the subjects in the previous studies erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier ones the first trial from each subject has been discounted.

WISC scores

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (Insert ordering)	11.9
Group 2 (Left-to-right violation ordering)	11.6
Group 3 (Mixed ordering-)	11.8

The average performance on the digit span test for ages between 9 years and 9 years 3 months is a score of 11 (WISC-R Manual, 1974).

Number of correct answers

Table 6.1 shows the mean number of correct answers for each experimental group.

Table 6.1: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

Ordering		
Insert	**L-R violation**	**Mixed**
4.7 (0.48)	4.7 (0.48)	3.8 (0.63)

A one way ANOVA was carried out on the above data and showed a significant effect of ordering ($F_{[2, 27]}=9.346$, $p<0.01$). Tukey comparisons showed that there was a significant increase in number of correct trials when comparing Insert ordering with Mixed ($q=5.30$, $p<0.01$), and also when comparing L-R violation with Mixed ($q=5.30$, $p<0.01$).

Time taken to complete trials (correct answers only)

Table 6.2 shows the mean time taken (in seconds) for each experimental group.

Table 6.2: Mean time taken to complete trials (correct answers only)

Standard deviations are shown in parentheses

Ordering		
Insert	**L-R violation**	**Mixed**
14.58 (2.44)	13.44 (2.39)	18.22 (3.70)

A one way ANOVA was carried out on the above data and showed a significant effect of ordering ($F_{[2, 27]}=7.372$, $p<0.01$). Tukey comparisons showed that there was a significant decrease in the time taken to correctly complete the orderings when comparing Insert ordering with Mixed ($q=3.95$, $p<0.05$), and also when comparing L-R violation with Mixed ($q=5.20$, $p<0.01$).

Order of placement data

Whilst the subjects were completing each task (as before, five were counted in total) the experimenter took details of the order in which they constructed the external array. These were then collated in respect of the following items:

1. Use of a left to right strategy, with respect to premise information consideration.

It was found that this was the predominant strategy in Experiment 3. For example, if the children were presented with BC AB CD , they would begin the array by placing items B and C, followed by items A and B (or just by joining on A) and then by C and D, rather scanning the array to select out the items in the correct order i.e. A B C D. The table below gives the mean number of trials in which this type of construction strategy was employed by subjects in the current study.

Table 6.3: Mean number of trials where a left to right take-up of information was used (max=5)

Ordering		
Insert	**L-R violation**	**Mixed**
4.9	2.8	4.1

2. Use of a left to right strategy, with respect to the position in which the children place the items.

Again, it was found that this was the predominant strategy for the 7 year old children in Experiment 3. In view of this, it was decided to look more closely at the error patterns and spontaneous correction of orderings in the current study. In order to do this, a distinction was made between successful trials due to self-correction and those which were successful from the outset. An answer which is correct from the outset is one in which no incorrect 2, 3 or 4 item orderings have been made during the problem solving process, whereas a self-correction involves the spontaneous correction of incorrect 2, 3 or 4 item orderings. An example of this distinction is given in the Results section of Experiment 3 (see Chapter 5).

Because of the very small amount of errors and self corrections in the Insert groups, consideration has been restricted to the L-R violation and Mixed groups.

Table 6.4: Total number of Left to Right violation errors

The figure in parentheses shows the total number of errors in that group. The two figures in each cell are then converted to a percentage.

Ordering	
L-R violation	**Mixed**
2 (3) 66%	7 (12) 58%

Table 6.5: Total number of successful Left to Right violation self-corrections

The figure in parentheses shows the total number of self-corrections in that group. The two figures in each cell are then converted to a percentage.

Ordering	
L-R violation	**Mixed**
7 (9) 78%	9 (10) 90%

3. Splitting of pre-ordered premise pairs

The data from Experiment 3 showed that a very high proportion of the tasks involved the splitting of a pair of items which had been ordered for the subjects in the original premises, particularly in the CD AB BC task. In view of this, the current data was scrutinised for similar evidence.

Table 6.6: Total number of 'previously ordered premise pair' splits - (CD AB BC) tasks

The figure in parentheses shows the total number of all initially correct, self-corrections and errors in that group. The two figures in each cell are then converted to a percentage.

Init. correct	Corrections	Errors
0 (10) 0%	5 (7) 71%	5 (8) 63%

Comparison with data from the 7 year old children (Experiment 3)

Because the subjects in Experiment 3 and this study were taken from the same schools, and the same materials and procedure were used throughout, it was decided to combine the relevant data from the two studies in a direct statistical comparison. Experiment 3 was a two factor experiment, where one of the factors involved a comparison between using photographs and drawings as stimulus material. No significant differences were found, so this was not investigated with the 9 year old children. Only the data from the 7 year old children who used the photographs has been used in this comparison, so as to equate with the stimulus material used in the current study.

Number of correct answers

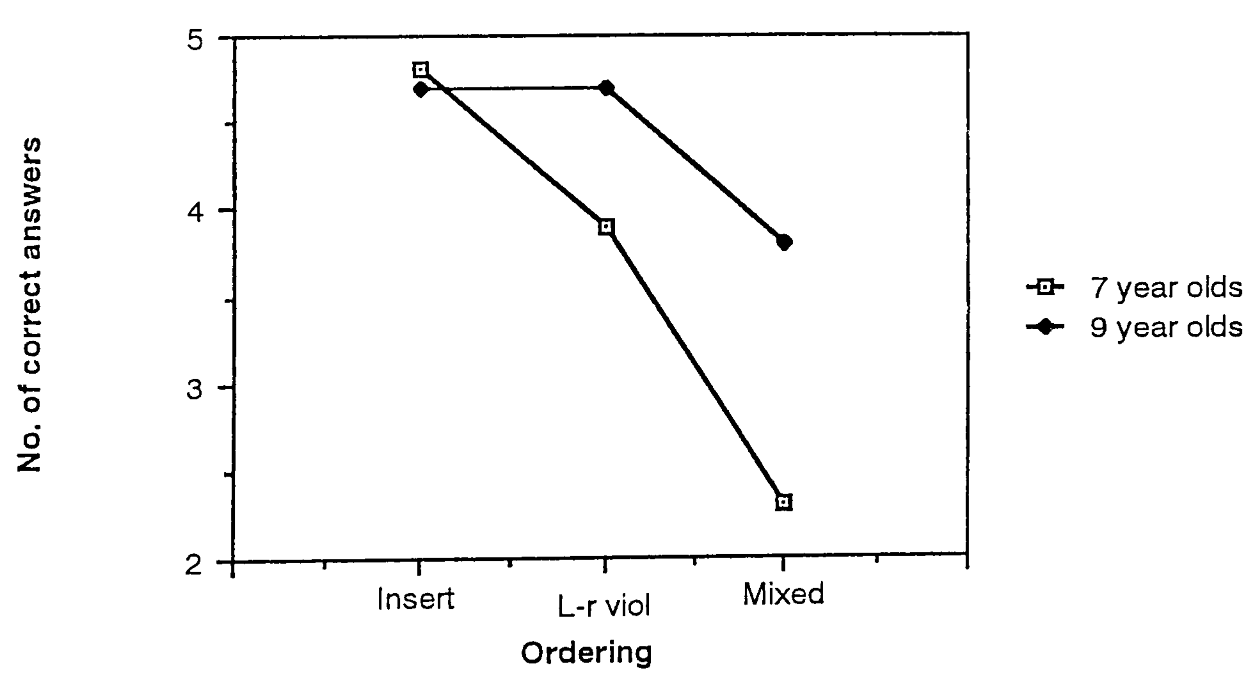
Table 6.7: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

	Ordering		
	Insert	**L-R violation**	**Mixed**
7 year olds	4.8 (0.42)	3.9 (0.88)	2.3 (1.34)
9 year olds	4.7 (0.48)	4.7 (0.48)	3.8 (0.63)

A two way ANOVA [2 (age) x 3 (ordering), both between subjects factors], was carried out on the above data and showed main effects of both age ($F_{[1, 54]}=13.444$, $p<0.01$) and ordering ($F_{[2, 54]}=26.861$, $p<0.01$) and also an interaction ($F_{[2,54]}=5.361$, $p<0.01$). Fig. 6.1 shows a graph of the interaction.

Fig. 6.1: Age x ordering (correct answers)



Further analysis showed the following significant simple main effects:

- the 9 year old children produced more correct solutions than the 7 year old children when solving the Left to Right violation problems ($F=5.33$, $p<0.05$).

- the 9 year old children produced more correct solutions than the 7 year old children when solving the Mixed problems ($F_{[1, 54]}=18.750$, $p<0.01$).

- an effect of task ordering for the 7 year old children ($F_{[2, 54]}=26.722$, $p<0.01$).

- an effect of task ordering for the 9 year old children ($F_{[2, 54]}=4.500$, $p<0.05$).

Tukey comparisons were made between the different levels of ordering for both age groups. They revealed the following significant effects:

- the 7 year old children produced more correct solutions using the Insert ordering than they did using the Left to Right violation ordering ($q=3.67$, $p<0.05$).

- the 7 year old children produced more correct solutions using the Left to Right violation ordering than they did using the Mixed ordering ($q=6.53$, $p<0.01$).

- the 7 year old children produced more correct solutions using the Insert ordering than they did using the Mixed ordering ($q=10.21$, $p<0.01$).

- the 9 year old children produced more correct solutions using the Left to Right violation ordering than they did using the Mixed ordering ($q=3.67$, $p<0.05$).

- the 9 year old children produced more correct solutions using the Insert ordering than they did using the Mixed ordering ($q=3.67$, $p<0.05$).

Time taken to complete trials (correct answers only)

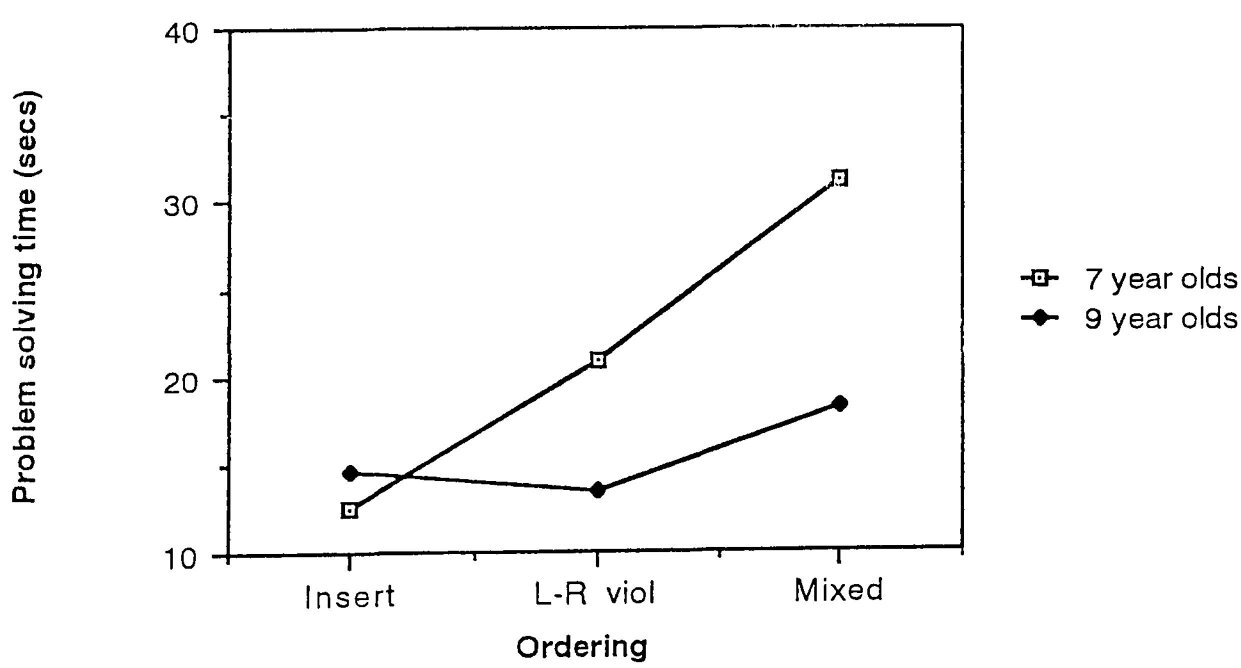
Table 6.8: Mean time taken to complete trials (in seconds)

Standard deviations are shown in parentheses

	Ordering		
	Insert	**L-R violation**	**Mixed**
7 year olds	12.56 (4.26)	20.86 (3.33)	31.19 (4.61)
9 year olds	14.58 (2.44)	13.43 (2.40)	18.22 (3.70)

A two way ANOVA [2 (age) x 3 (ordering), both between subjects factors, was carried out on the above data and showed main effects of both age ($F_{[1, 54]}$=44.359, p<0.01) and ordering ($F_{[2, 54]}$=51.333, p<0.01) and also an interaction effect ($F_{[2, 54]}$=22.787, p<0.01). Fig. 6.2 shows a graph of the interaction.

Fig. 6.2: Age x ordering (problem solving time)



Further analysis showed the following significant simple main effects:

- the 9 year old children produced correct solutions more quickly than the 7 year old children when solving the Left to Right violation problems ($F_{[1,54]}$=21.529, $p<0.01$).

- the 9 year old children produced correct solutions more quickly than the 7 year old children when solving the Mixed problems ($F_{[1, 54]}$=66.777, $p<0.01$).

- an effect of task ordering for the 7 year old children ($F_{[2, 54]}$=69.180, $p<0.01$).

- an effect of task ordering for the 9 year old children ($F_{[2, 54]}$=4.941, $p<0.05$).

Tukey comparisons were made between the different levels of ordering for both age groups. They revealed the following significant effects:

- the 7 year old children produced correct solutions more quickly using the Insert ordering than they did using the Left to Right violation ordering (q=7.35, $p<0.01$).

- the 7 year old children produced correct solutions more quickly using the Left to Right violation ordering than they did using the Mixed ordering (q=9.25, $p<0.01$).

- the 7 year old children produced correct solutions more quickly using the Insert ordering than they did using the Mixed ordering (q=16.60, $p<0.01$).

- the 9 year old children produced correct solutions more quickly using the Left to Right violation ordering than they did using the Mixed ordering ($q=4.26$, $p<0.01$).

Order of placement data

No statistical comparison has been made between the two studies in this respect. However, so that the two experiments can be compared discursively, the following tables combine the relevant data from this experiment and Experiment 3.

Table 6.9: Mean number of trials where a left to right take-up of information was used (max=5)

		Ordering		
		Insert	L-R viol.	Mixed
Age	7 years	4.8	4.1	4.4
	9 years	4.9	2.8	4.1

Table 6.10: Total number of Left to Right violation errors

The figure in parentheses shows the total number of errors in that group. The two figures in each cell are then converted to a percentage.

		Ordering			
		L-R viol.		Mixed	
Age	7 years	10 (12)	83%	40 (53)	85%
	9 years	2 (3)	66%	7 (12)	58%

Table 6.11: Total number of successful Left to Right violation self-corrections

The figure in parentheses shows the total number of self-corrections in that group. The two figures in each cell are then converted to a percentage.

		Ordering			
		L-R viol.		Mixed	
Age	**7 years**	14 (16)	88%	12 (20)	75%
	9 years	7 (9)	78%	9 (10)	90%

Table 6.12: Total number of paired item splits - (CD AB BC) tasks

The figure in parentheses shows the total number of initially correct, self-corrections and errors in that group. The two figures in each cell are then converted to a percentage.

		Init correct		Corrections		Errors	
Age	**7 years**	0 (2)	0%	8 (9)	89%	19 (19)	100%
	9 years	0 (10)	0%	5 (7)	71%	5 (8)	63%

6.1.3 Discussion

It can be seen from the Results section that the 7 and 9 year olds are behaving differently in the different levels of premise ordering. In view of this, the discussion to follow compares the two age groups' performance with each of the different types of ordering. Both types of measurements (number of correct trials and time taken to correctly order the array) show very similar patterns of results and so the following refers to both types of scores.

Insert condition - AB CD BC.

It would appear that the scores are approaching ceiling for both age groups in this condition. The mean times taken were 12.56 seconds (7 year olds) and 14.58 seconds (9 year olds). There is no significant difference between these two times. Observation of the 7 year old children showed that completion of the task was a continuous sequence of motor actions, with very few self-corrections and no time given over to reflection. An informal timing of 7 year old children who were asked to place similar 4 item arrays, but by merely copying a given ordering, gave a mean of 12.26 seconds, with a range of 10.82 to 12.98 seconds. Means of 4.7 and 4.8 correct answers out of a possible 5 are likely to be reflecting 'performance slips' rather than conceptual misunderstandings.

However, it must be remembered that the children who completed this ordering could be successful simply by applying a Left to Right strategy. Indeed the 7 year old children were remarking that the last premise (BC) was not needed. This was not the case for the 9 year old children, several of whom actually 'marked' the last presented premise with a finger, or obviously glanced at it, in order to check the already completed array. Some of these children began to ignore this premise towards the end of their block of tasks, but they usually remarked that all the tasks were the same and so they knew what the required ordering was.

It would seem, therefore, that the similar performance levels obtained by the two different age groups do not reflect similar levels of understanding. The 7 year olds success is due to the use of a low-level strategy which replicates the Left to Right ordering of the items as presented in the premise ordering. They make no attempt to verify the ordering by reference to the mid premise.

Left to Right violation condition - BC AB CD, BC CD AB

Again, the 9 year old children are approaching ceiling in this condition. The 7 year old children in Experiment 3 were significantly less successful when using this ordering, however.

This can be explained by considering the order of premise placement data, which looks in more detail at the use of the Left to Right ordering strategy. The 7 year old children in this condition were using a left to right take-up of premise information virtually all of the time (4.5 out of 5), whereas the 9 year olds left to right take-up drops to 2.8 out of 5. Thus it would seem that the older children have begun to scan the array in order to isolate an 'end-anchor', showing that they are well aware of the significance of the mid-items in the array.

However, although this ordering showed fewer errors and self-corrections for the older children, the proportion of these which were attributed to an incorrect left to right strategy application was still very high.

Thus it would seem that there is still some tendency for the 9 year old children to initally apply this strategy but they can easily identify and correct the resultant errors. This could be due to metacognitive strategies i.e. the checking of each premise item with the partially ordered array so that all the pertinent premise information is extracted and used. In order to be able to do this, it is necessary to be aware of what information is needed for the task, in other words, to possess some structural task understanding. If the 7 year olds think that the task consists of introducing the new items in the order in which they are presented in the premises, then that is the only information they will extract from the premise and rechecking will not be deemed necessary or appropriate.

Mixed condition.

Although the 9 year olds were more successful than the 7 year olds in this condition, their performance was not approaching ceiling, but was similar to the scores obtained for the 7 year olds in the L-R violation condition. It would seem therefore, that even though the older children are aware of the need to check premise information and correct for Left to Right violation errors, they still have some other constraint which is operating on performance.

It seemed to the experimenter that the 9 year old children were behaving differently with the two sub-groups of tasks included in this condition. Ordering CD BC AB appeared to present very little difficulty. There was some self-corrections, again due to left to right strategies, but far fewer actual errors. Ordering CD AB BC still seemed to be presenting problems. More of the tasks were actually correct than for the 7 year olds (10 as opposed to 1), but there were still more errors in this subgroup. However, data directly comparing the two sub-groups has not been analysed due to the possible confounding factor of presenting examples of the two subgroups alternately to each subject within the same condition. It could be that these two subgroups are sufficiently different so as to cause some initial confusion (though of course, if the children had a complete initial structural understanding this should present no problems). In view of this, a further study will directly compare the performance of 9 year old children using ordering CD AB BC with those using CD BC AB, in a between subject design.

If we look at the order of premise placement data, it can be seen that both age groups have a high proportion of paired item splits for ordering CD AB BC. Thus, although the 9 year olds are more successful than the 7 year olds in this condition, they still experience the same type of persistent error. It is only with this ordering that a left to right take-up of information (an initial strategy when task specific actions are being constructed) will lead to the formation of an array in which a pair of items needs to be re-ordered i.e. CDAB to ABCD. It seems therefore, that there is still not complete structural understanding of the task for the 9 year olds, probably due to the inability to appreciate the significance of 'given' premise information.

If an understanding of the significance of 'given' premise information is still a problem, we would expect ordering CD AB BC to be difficult, whereas ordering CD BC AB will result in performance similar to that of Left to Right violation for the 9 year old children. This is because the construction of the array with premise ordering CD AB BC usually proceeds by joining AB to CD, i.e. CDAB. An understanding of the significance of 'given' premise information, together with the viewing of the third premise (BC), would result in a correct reordering to ABCD, as the subject would realise that they needed to keep items A and B adjacent. This is because of the information contained in the first premise. The most efficient way to do this, whilst at the same time incorporating the information given in the third premise, would be to move items A and B, as a pair, to the front of item C. Of course, a split premise pair which has resulted in the incorrect reordering of CDAB to BCDA can still be corrected, but only by reference to the second premise. Thus it is still possible that children who split premise pairs will be able to raise their number of correct solutions to ceiling by using spontaneous self-corrections. However, because of the greater number of separate actions required, there may well still be some time costs.

6.2 EXPERIMENT 6 - THE EFFECT OF ORDERING ON THE INTEGRATION OF INFORMATION FOR 9 YEAR OLD CHILDREN (PART 2)

Rationale

Experiment 5 (reported above) looked at the performance of 9 year old children when ordering an array. It was found that the order in which the premises were presented significantly affected the number of correct solutions and the time taken to correctly order the arrays, such that 'Insert' and 'Left to right violation' orderings were both more successful than 'Mixed' orderings.

When the same study was ran with 7 year old children (Experiment 3), a different result was obtained. Insert ordering was the most successful. There was a significant difference between the results obtained in this condition and the L-R violation condition and also between the L-R violation and Mixed (the most difficult condition). Evidence was also obtained in support of the view that the main problem for the 7 year olds was the incorrect application of a left to right ordering strategy when trying to order the array. This left to right rule was also invariably used when the children were considering the premise information i.e. they used the information from the leftmost premise first. The strategy therefore produced errors in the L-R violation and Mixed conditions. This was not the case for the 9 year olds, however. Their performance was approaching ceiling in L-R violation, but was significantly inhibited in the Mixed condition. It would seem therefore, that there are other constraints which are having an effect here.

The Mixed condition contains two ordering types:

1. CD BC AB
2. CD AB BC

These were initially grouped together, as it was hypothesised that they both require complex, multiple ordering strategies (see Experiment 3).

However, the data collected in Experiment 5 suggested that the subjects were acting differently in the two different task types. It seemed that CD BC AB was easier than CD AB BC. However, this could not be tested statistically as the children in the study did three of each task alternately. It could be that this task ordering inhibited performance in the more 'difficult' condition, as the children were getting confused in the early stages, when trying to build a conceptual understanding of the task structure. For this reason, it has been decided to directly compare performance in the two types of ordering in a between subjects design. There are four likely groups of results:-

1. No significant difference between the groups in this study, and also no significant difference between these groups and the Mixed condition in Experiment 5. This would suggest that the difficulty experienced by the 9 year old children in Experiment 5 is due to a factor which is common to both of the sub-groups of tasks in the Mixed condition.

2. No significant difference between the groups in this study, and also no significant difference between these groups and the L-R violation condition in Experiment 5. This would suggest that the tasks themselves are inherently no more problematic than any others for the 9 year old children, as performance will be close to ceiling. Thus the former inhibition of performance will have been due to the presentation of the two tasks simultaneously.

3. A significant difference between the groups, where CD BC AB is the most difficult. In this condition the subjects will need to violate the 'left to right' heuristic twice. This type of result will throw doubt on the suggestion that the incorrect application of a left to right ordering strategy is peculiar to the 7 year olds.

4. A significant difference between the groups, where CD AB BC is the most difficult. If this is the case then there are still constraints on performance for the 9 year old children, but these are not merely due to a left to right heuristic. Inspection of the order of placement data will provide details of the precise nature of the difficulty. If the subjects in this study have a full understanding of the task structure, there should only be a small difference (due to the action of moving items A and B to the front of the

queue) between orderings AB CD BC and CD AB BC. Both orderings require inspection of the third premise in order to link the information contained in the first two premises, as they contain no common item. The children in Experiment 5 (9 year olds) appeared to recognise the necessity of checking the third premise, at least in the first few tasks, before they realised that all the tasks shared the same ordering. The 7 year olds did not share this understanding, indeed several of them remarked that the third premise was not needed.

6.2.1 Method

Design

A one factorial between subject design was used.

The two levels of premise ordering were as follows:

Condition 1	CD BC AB
Condition 2	CD AB BC

Note Serial ordering was used as a worked example.

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task. The 'order of placement' details were also recorded, for both correct and incorrect orderings.

Participants

The participants in this study were 20 mixed ability 9 year olds (mean age 9 years 3 months, range 8 years 10 months to 9 years 9 months) from a state first school, with a predominantly middle class catchment area. The subjects were from the same school as those in Experiments 3 and 4, and were randomly assigned to one of the two experimental groups.

Task description

As in Experiment 3, except that the subjects in Condition 1 all worked only with task ordering CD AB BC, and those in Condition 2 all worked only with task ordering CD BC AB.

Materials

As in Experiment 3, but using the photographs only. An example of the task materials is given at Appendix D.

Procedure

As in Experiment 3.

6.2.2 Results

Note Each subject completed six experimental trials. However, a large number of the subjects in the previous studies erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier ones the first trial from each subject has been discounted.

WISC scores

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (CD AB BC ordering)	11.7
Group 2 (CD BC AB ordering)	11.6

The average performance on the digit span test for ages between 9 years and 9 years 3 months is a score of 11 (WISC-R Manual, 1974).

Number of correct answers

Table 6.13 shows the mean number of correct answers for each experimental group.

Table 6.13: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

Ordering	
CD AB BC	**CD BC AB**
2.6 (0.52)	4.6 (0.52)

A one way ANOVA was carried out on the above data and showed a significant effect of ordering ($F_{[1, 18]}$=75.00, p<0.01). There was a significant increase in the number of correct trials for Condition 2 (CD BC AB).

Time taken to complete trials (correct answers only)

Table 6.14 shows the mean time taken (in seconds) for each experimental group.

Table 6.14: Mean time taken to complete trials (correct answers only)

Standard deviations are shown in parentheses

Ordering	
CD AB BC	**CD BC AB**
23.10 (2.27)	12.03 (1.38)

A one way ANOVA was carried out on the above data and showed a significant effect of ordering ($F_{[1, 18]}=173.49$, $p<0.01$). Again, there was a significant improvement for Condition 2 (CD BC AB).

Order of placement data

Whilst the subjects were completing each task (as before, five were counted in total) the experimenter took details of the order in which they constructed the external array. This revealed that those subjects in Condition 2 (CD BC AB) made very few errors or self corrections. This is also supported by the error and time data reported above. The subjects in Condition 1 (CD AB BC) however, were making errors by placing A after CD. This was then followed by B, forming an incorrect ordering of CDAB. They were attempting to correct these errors after having looked at the third premise (BC). A partial reordering to BCDA occurred in most instances, but then the subjects became confused and began to select any one premise at random, trying to ensure that the two items relevant to that premise were correctly ordered, but not linking the two premises together. It seems that the lack of a common item in the first two premises confused the subjects and so they were seldom successful in combining all the relationships.

Comparison with data from the 9 year old children in Experiment 5

Because the subjects in Experiment 5 (see Chapter 5) and this study were taken from the same school, and the same materials and procedure were used throughout, it was decided to combine relevant data from the two studies in a direct statistical comparison. Experiment 5 compared the effects of three different types of task ordering, where the most difficult level consisted of three of each of the two tasks which are currently being compared between subjects. The children were performing at ceiling in the other two types of ordering, one of which, together with the Mixed data, will be included in the following analysis for comparison purposes. The data from the L-R violation and Mixed conditions from Experiment 5 have therefore been treated as the third and fourth conditions.

Number of correct answers

Table 6.15: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

Ordering			
CD AB BC	**CD BC AB**	**L-R viol**	**Mixed**
2.6 (0.52)	4.6 (0.52)	4.7 (0.48)	3.8 (0.63)

A one-way ANOVA was carried out on the above data and revealed a significant ordering effect ($F_{[3, 36]}=32.314$, $p<0.01$). Tukey comparisons showed the following significant results:

- an increase in number of correct trials from CD AB BC ordering to L-R violation ordering ($q=12.30$, $p<0.01$).

- an increase in number of correct trials from CD AB BC ordering to Mixed ordering ($q=7.03$, $p<0.01$).

- an increase in number of correct trials from Mixed ordering to CD BC AB ordering ($q=4.68$, $p<0.05$).

- in agreement with previous analyses, an increase in number of correct trials from Mixed ordering to L-R violation ordering ($q=5.27$, $p<0.01$).

- in agreement with previous analyses, an increase in number of correct trials from CD AB BC ordering to CD BC AB ordering ($q=11.71$, $p<0.01$).

Time taken to complete trials (correct answers only)

Table 6.16: Mean time taken to complete trials (in seconds)

Standard deviations are shown in parentheses

Ordering			
CD AB BC	**CD BC AB**	**L-R viol**	**Mixed**
23.10 (2.27)	12.03 (1.34)	13.44 (2.39)	18.22 (3.69)

A one-way ANOVA was carried out on the above data and revealed a significant ordering effect ($F_{[3, 36]}=38.189$, $p<0.01$). Tukey comparisons showed the following significant results:

- a decrease in time taken to correctly complete trials from CD AB BC ordering to L-R violation ordering ($q=11.88$, $p<0.01$).

- a decrease in time taken to correctly complete trials from CD AB BC ordering to Mixed ordering ($q=6.01$, $p<0.01$).

- a decrease in time taken to correctly complete trials from Mixed ordering to CD BC AB ordering($q=7.61$, $p<0.01$).

- in agreement with previous analyses, a decrease in time taken to correctly complete trials from Mixed ordering to L-R violation ordering ($q=5.88$, $p<0.01$).

- in agreement with previous analyses, a decrease in time taken to correctly complete trials from CD AB BC ordering to CD BC AB ordering ($q=13.62$, $p<0.01$).

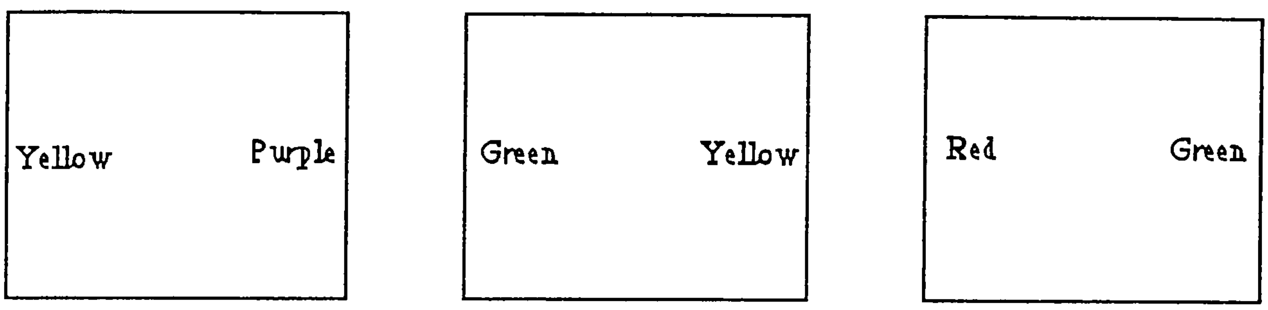
6.2.3 Discussion

It can be seen from the results that the ordering CD AB BC was significantly more difficult than ordering CD BC AB. The data from this experiment was also compared with the relevant data from Experiment 5 as the subjects used in both studies were of equivalent ages and from the same schools. This comparison showed that the children in Condition 2 of this study (ordering CD BC AB) were performing at the same success rates (approaching ceiling) as those in the Left to right violation condition of Experiment 4, and the same children were also performing better than those in the Mixed condition of Experiment 5. Additionally, the subjects in Condition 1 ordering CD AB BC) of this study experienced more difficulty than those in the Mixed condition of Experiment 5. Thus the four different types of ordering which were directly compared fell into three levels of difficulty as follows:

CONDITION		DIFFICULTY
CD AB BC		(most difficult)
Mixed condition i.e	CD AB BC and CD BC AB	(intermediate)
CD BC AB L-R violation condition i.e.	BC CD AB and BC AB BC	(least difficult)

It seems therefore, that the result obtained in Experiment 5 was not due solely to the presentation of two types of task in the same condition. If this was the case, we would have expected to see the 9 year olds performance to be at or approaching ceiling when the two types of ordering were presented in different conditions. However, when these two types of tasks were presented separately (i.e.. in the current study), there was inhibition of performance for the CD AB BC task only. Also, this task was the more difficult when compared to the condition in which the children used both types of premise presentation. Furthermore, 9 year olds working with the ordering CB BC AB were performing at or approaching ceiling. These results show that there are still constraints on performance for the 9 year old children, but only when they are ordering using the CD AB BC presentation.

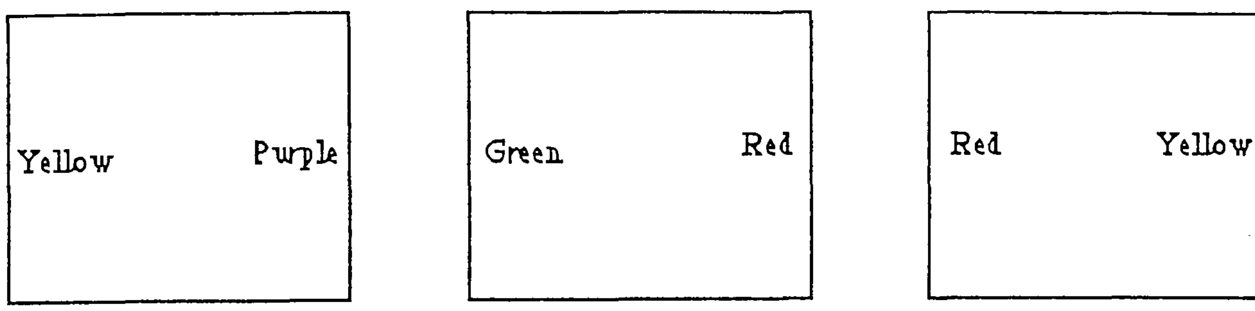
We must therefore consider what is special to this type of ordering which is still inhibiting performance for the 9 year olds. It seems that, like the 7 year olds in Experiment 3, their problems are due to a combination of the inappropriate use of a left to right ordering strategy, and working with the premises strictly in the order in which they were presented (see Chapter 5 for a full discussion of this) The data from ordering they were presented (see Chapter 5 for a full discussion of this) The data from ordering CD BC AB showed that the 9 year olds had begun to scrutinise the premises for 'end-anchor items', as they only used a left to right take-up of premise information for an average of 2.1 out of 5 trials. This is similar to the performance of the 9 year olds working with the easier orderings in Experiment 5 . Thus in the premise ordering below, the subjects would scan the array to look for the first item (red jumper) and so on, therefore building the array in a left to right manner from the outset.



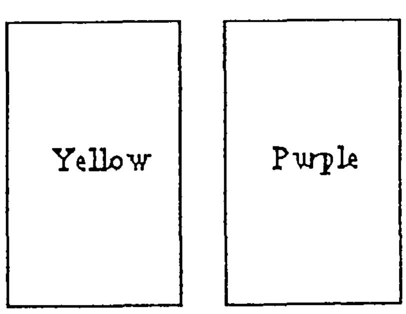
This strategy of scanning the array for the correct item means that subjects no longer have to place any items at the front of a partially ordered array, thus removing the requirement to violate the left to right ordering rule. This has therefore resulted in performance which is approaching ceiling for the children working with CD BC AB ordering. The situation is very different for those children working with ordering CD AB BC. They are still using a left to right take-up of information for an average of 4.6 out of 5 trials. This means that they are required to violate the left to right rule in the 3rd placement. This is demonstrated in Fig. 6.3 below.

Fig. 6.3: Example of CD AB BC ordering

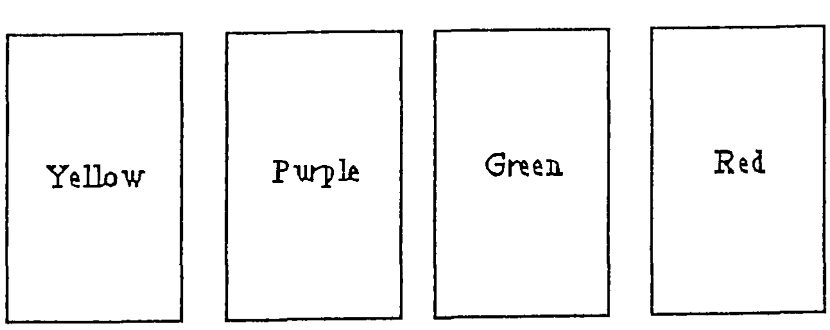
Premise ordering



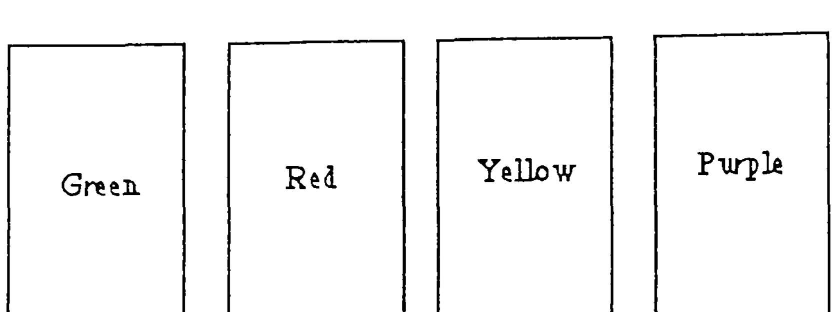
1st placement



2nd placement



3rd placement



Because the premises have been acted on in the order in which they were presented, the 3rd placement required a violation of the left to right ordering rule as the red jumper has to be placed at the front of the partially ordered array. This ordering also requires that the 2nd premise is then reconsidered, so that the green jumper is placed at the front of the array. The only way the violation can be avoided is by moving the green and red jumpers together as a premise pair.

It seems therefore that presentation of two initial unlinked premises (as in a CD AB BC as illustrated above) prevents full understanding of the task structure. Subjects begin by using a left to right take-up of information. This means that the first and second premises become erroneously linked. The subjects then realise that this situation is incorrect as they consider the third premise. However, they then only move the red jumper as they do not keep the red and green jumpers together, even though they were given as an original premise pair. The children then either believe that they have correctly completed the task, or they rescrutinise the premises and become very confused and make more errors. This has resulted in inhibition of performance for the 9 year olds working with the CD AB BC ordering.

It is interesting to note that this confusion did not occur in Experiment 5 when 9 year old children were working with ordering AB CD BC. In this case they appeared to realise that they needed to inspect the third premise in order to check the linkage of the first two premises. They did this by marking the last presented premise (BC) with a finger, or obviously glancing at it, in order to check the already completed array. The 7 year olds did not share this understanding, indeed several of them remarked that the third premise was not needed.

However, when the 9 year olds had to alter an ordering to account for this rechecking, as in ordering CD AB BC, they ran into difficulties because they did not realise the significance of information which was given to them in the premises. They did not, therefore, manipulate the premise items as a pair when reordering and so become confused.

It seems therefore that 9 year olds have a fragile conceptual understanding of the structural representation of series problems, as they appear to realise the need to check all three premises. However this understanding is not robust as it results in errors when the premise ordering requires that characters are re-ordered in pairs, dependent on information contained in the original premises.

6.3 GENERAL SUMMARY - CHAPTERS 5 AND 6

The four experiments reported in these two chapters investigated the role of premise ordering in the building of structural representations. It has been demonstrated in Experiment 3 that 7 year old children apply an inappropriate left to right ordering strategy when solving series problems. Evidence for this was provided by analysis of the error and time data and also by the order in which the subjects used the premise information and placed new items. Experiment 4 showed that this strategy was not restricted to working with unfamiliar task materials. Experiment 5 demonstrated that the inappropriate left to right ordering strategy did not occur in the 9 year old children's performance, as they were successful in ordering the tasks with required one left to right ordering violation. The 7 year olds performance on the same tasks however, did show evidence of inappropriate strategy use. The most complex ordering (CD AB BC), which was separately investigated in Experiment 6, resulted initially in attempts by the 9 year olds to order in a left to right manner, but the subjects noticed these and attempts were made to correct the resultant errors. However, successful performance was still not forthcoming as the subjects could not reorder the incorrect arrays so as to ensure that all the relational information was correctly integrated. It was suggested that this is due to a failure to appreciate and use the status of the 'given' information, i.e. that provided in the premises as part of the task. This could be due to:

1. A total lack of understanding of the relational information given in each premise.

2. A failure to remember this information at a critical decision point, probably due to limited cognitive resources.

Point 1 above is not considered likely, as the 9 year olds in Experiment 6 were **trying** to re-order based on individual premises. This demonstrates that they did understand the significance of the relational information contained in each premise. The second suggestion is much more probable, but further studies are necessary to address this.

It is also interesting to speculate whether a failure to appreciate and use the status of the 'given' information is affecting the 7 year old's performance. The third condition in Experiment 3 (mixed) was more difficult for these children than the L-R condition. If the only factor affecting the 7 year olds in these tasks is the inappropriate application of a left to right ordering strategy, we would expect these conditions to be equally difficult, as they both involve the insertion of an item into a partially ordered array at a position other than immediately to the right. However, if these children are also experiencing difficulty in dealing with the lack of a common item in the first two premises when ordering CD AB BC, then it is not surprising that the Mixed condition is posing more difficulties. A further experiment, which replicates Experiment 6 but with younger subjects, could be carried out to investigate whether 7 year old children also experience difficulty when needing to reorder a premise pair i.e. the incorrect ordering CDAB requires that the premise pair AB is moved to the front of the array.

Nonetheless, we now have strong evidence that the construction of integrated relational representation is adversely affected by the order in which premise information is presented. This encourages the application of inappropriate strategies which result in an incorrect structural representation. Chapter 8 will go to investigate the reasons why

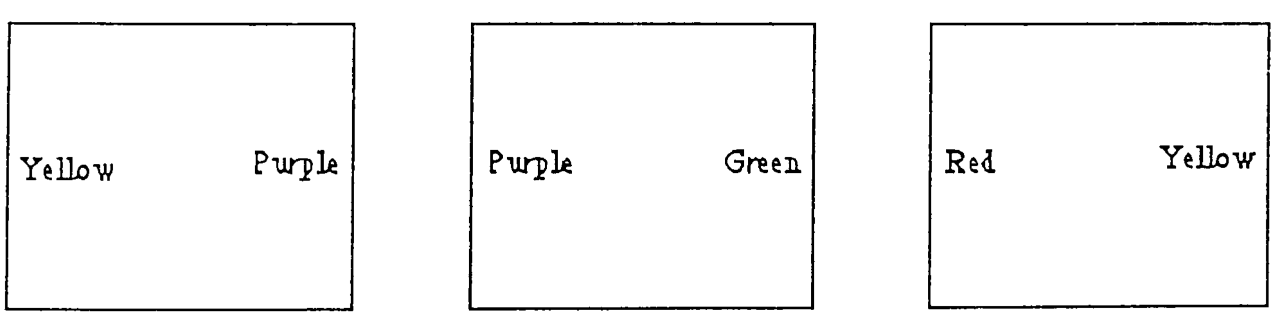
7 year olds performance is inhibited by premise ordering in these types of tasks. Before this, it is necessary to investigate any possible inhibition of performance due to the erroneous belief that duplicate tokens of the same character must be used in the integrated array (see Chapter 5 for a full discussion of this). Chapter 7 deals with this issue.

CHAPTER 7 : TASK CONSTRAINTS IN THE BUILDING OF STRUCTURAL REPRESENTATIONS (PART 3) CONCERNING THE REDUNDANCY OF ITEM INFORMATION

Experiment 1, reported in Chapter 4, identified two possible factors which might have been contributing to the difficulty which 7 year old children experience when integrating separate premises into a single, ordered array. It is the second of these factors, that of the elimination of redundant items, which is addressed in this chapter.

An experimental manipulation (reported as part of Experiment 3 in Chapter 5), has already been carried out to investigate whether or not 7 year old children were aware that duplicate characters in the original premise information needed to be amalgamated into one single character when the premises were integrated and the array was ordered. This is illustrated in Fig. 6.1 below.

Fig. 6.1: Example of random premise information.



In the example above, the yellow and purple jumpers are represented twice in separate premises, as they are the 'mid-items' of the integrated array.

It was hypothesised that the children might have been misled by the use of drawings as stimulus material. It could be that they were not aware that two similar drawings (for example the yellow and purple jumpers in Fig. 6.1 above) were actually different representations of the same character. This lack of knowledge would have prevented them from successfully ordering the array, as it is dependent on integration of premise information.

It was decided, based on work by DeLoache in 1991, to use photographs as stimulus materials, rather than drawings. This is because DeLoache showed that photographs were the easiest graphic representation for young children to understand. This is also consistent with work reported by O'Connor, Beilin and Kose (1981) who argued that 6 year olds believe in the fidelity of photographs, as they are more likely to accept illogical outcomes as true when they were represented in photographs rather than in drawings. These two studies suggest that children might find it easier to accept that two identical photographs (rather than drawings) in separate premises were of one character. This might therefore have lead them to realise the necessity to use only one photograph of a character when building the integrated array.

However, the results from Experiment 3 showed that there was no significant improvement in performance when using photographs rather than drawings as stimulus materials, therefore providing no evidence in support of the experimental hypothesis. We must now question whether this hypothesis was an adequate test of the underlying theoretical question, that is, whether or not 7 year old children understand that an integrated array necessarily entails using only one token of each character. Experiment 7, reported below, was designed to test this question using an alternative hypothesis.

7.1 EXPERIMENT 7 - THE EFFECT OF PROVIDING DUPLICATE CHARACTER TOKENS

Rationale

The question we wish to address is whether or not young children understand that the integration of separate premises into one ordered array entails that each task item (in this case individual characters) has to be represented only once.

Experiment 3 investigated this issue by comparing the use of photographs as stimulus material with that of drawings. This manipulation resulted in no significant differences in performance, thus providing no evidence in support of the experimental hypothesis. However, for both Experiments 1 and 3, the subjects were given only one token of each character with which to construct the integrated array. If the children were expecting to have duplicate character tokens to work with, then this could have had one of two possible effects:

1. The children could have become very confused and believed that completion of the task was not possible, thus resulting in inhibition of performance.

2. Due to the lack of duplicate tokens, the subjects could have been prevented from making some errors. This is because, had duplicate tokens been available, the children might have used them, thus resulting in an incorrect array. In this case the restriction of items might have lead to a facilitation of performance.

In order to distinguish between these two effects, and to reconsider the original question, the current study was designed. It compares the performance of subjects given only one token of each of the characters in the array with those given several character tokens and allowed to select the ones they need.

7.1.1 Method

Design

A one factorial between subject design was used.

The two levels of 'number of character tokens available' were one token (Condition 1) and six tokens (Condition 2).

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task. The 'order of placement' details were also recorded, for both correct and incorrect orderings.

Participants

The participants in this study were 20 mixed ability 7 year olds (mean age 7 years 2 months, range 6 years 7 months to 7 years 7 months) from a state primary school, with a predominantly middle class catchment areas. They were randomly assigned to one of the two experimental groups.

Task description

As in Experiment 3, except that all the subjects worked with the following task orderings:-

BC AB CD

BC CD AB

This corresponds to level L-R (Left to right violation) used in Experiments 3 to 5.

Materials

As in Experiment 3 but using the drawings only. Copies of these are given in Appendix C. Five more copies of the complete array were made, and then cut into individual character tokens.

Procedure

As in Experiment 3, except that the children in Condition 1 were provided with one token of each character to work with, whereas the children in Condition 2 were provided with six tokens of each character and told to use as many as they needed.

WISC-R digit recall scores were also taken from each subject.

7.1.2 Results

Note Each subject completed six experimental trials. However, a large number of the subjects in the first study erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier studies the first trial from each subject has been discounted.

WISC scores

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1	8.4

Group 2 8.6

The average performance on the digit span test for ages between 7 years and 7 years 3 months is a score of 8 (WISC-R Manual, 1974).

Number of correct answers

Table 6.1 shows the mean number of correct answers for each experimental group.

Table 6.1: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses.

Number of tokens for each character	
1 token	**6 tokens**
3.9 (0.74)	3.7 (0.95)

A one way ANOVA was carried out on the above data and showed no significant effect of number of character tokens provided ($F_{[1, 18]}=0.277$, $p>0.05$). Thus the availability of more than one token does not affect the number of successful trials.

Time taken to complete trials (correct answers only)

Table 6.2 shows the mean time taken (in seconds) for each experimental group.

Table 6.2: Mean time taken to complete trials (correct answers only)

Standard deviations are shown in parentheses

Number of tokens for each character	
1 token	**6 tokens**
21.44 (2.50)	22.93 (4.82)

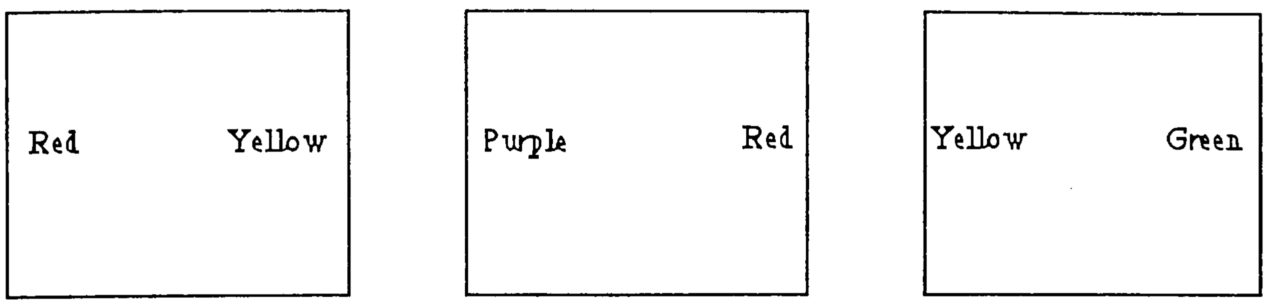
A one way ANOVA was carried out on the above data and showed no significant effect of number of character tokens provided ($F_{[1, 18]}=0.746$, $p>0.05$). Thus the availability of more than one token does not affect the time taken to successfully order the array.

Order of placement data

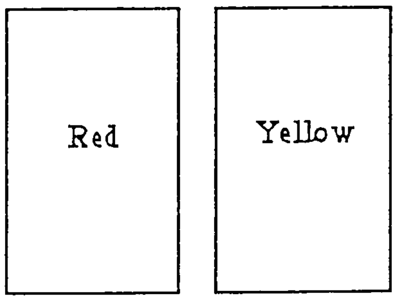
Whilst the subjects were completing each task (as before, five were counted in total), the experimenter took details of the order in which they constructed the external array. These were then examined with a view to identifying any differences in the use of left to right strategies, as reported in Experiment 3.

No differences in left to right strategies were apparent, both between the two conditions investigated in this experiment and also between the current study and the equivalent results reported in Experiment 3. Nonetheless, one other difference in the behaviour of the two experimental groups taking part in the current study was noted. The children in Condition 2 of this current experiment were **initially** using the extra tokens to replicate at least the first incidence of a mid-item character in the premises. This initial action did not effect the overall performance of the subjects in Condition 2, however, as the redundant token was quickly identified and discarded. An example of a typical construction sequence for the task BC AB CD, by a subject who had six available tokens for each character, would be as follows:

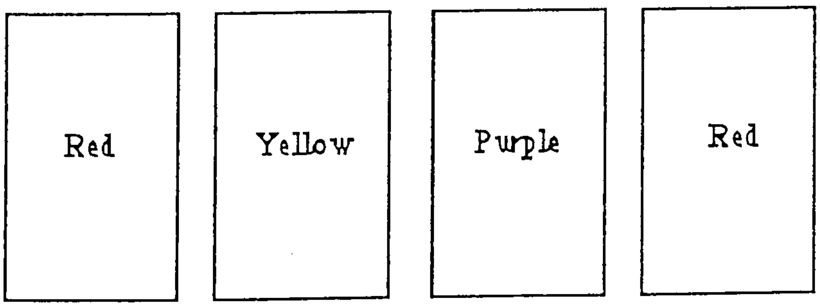
Premise presentation



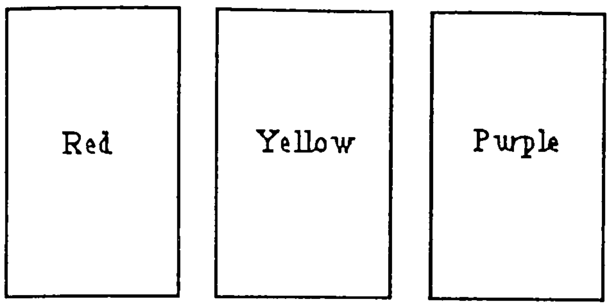
First placement



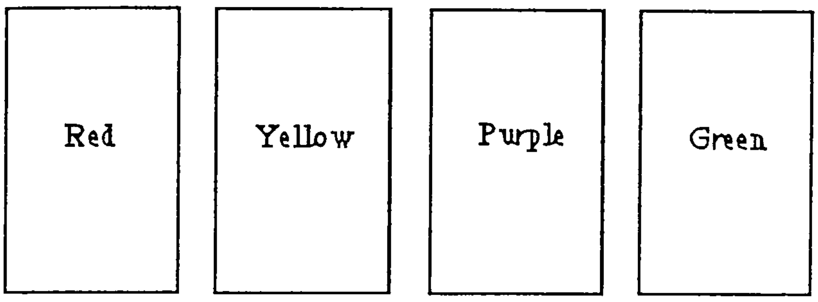
Second placement



Third placement



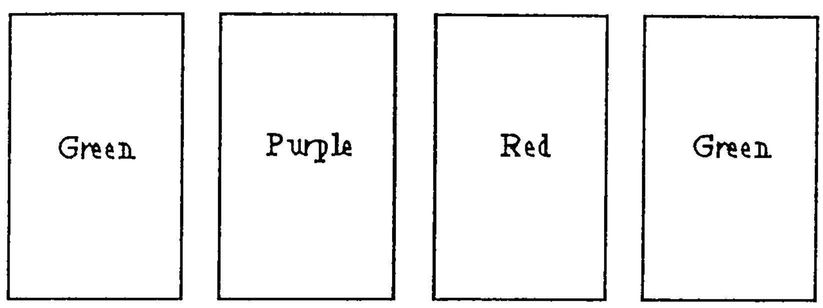
Fourth placement



7.1.3 Discussion

The lack of any significant difference between conditions means that we have no evidence in support of the experimental hypothesis, i.e. that children's performance when integrating information into a single structural representation will be affected by the number of available items from which the representation can be constructed. Experiment 3 also found no significant improvement when photographs (a more easily understood graphical representation for young children), rather than drawings, were used as the task materials. These two results mean that the theoretical claim concerning the redundancy of information remains unsubstantiated. Thus it seems unlikely that the initial difficulty for the 7 year old children reported in Experiment 1 was due to an inability to eliminate the redundant mid-items contained in the original premises. As seen in the example above, those children who had the opportunity to work with duplicate representations of the same character did make use of this initially, but it did not affect their performance, both in terms of number of errors or time taken to solve the problems.

Similar results to those from Experiment 3 were also obtained concerning the incorrect application of a left to right ordering strategy. Although the children who managed to place the identical representations of one character next to each other had no problems in eliminating one of the representations (see above), they were unable to use their knowledge concerning redundant items to aid them in forming the correct ordering. Those children who erroneously applied this strategy did so in spite of the fact that their own construction, for example



was highlighting the fact that the array needed to be ordered so that one of the identical representations representing the green jumper could be discarded. Such children knew that one of the two representations should be removed, but were very uncertain as to which one it should be, and appeared to have no means of resolving this problem. It appears that these children, although often seeming to realise that their array is somehow 'wrong', are still unable to monitor their problem solving and to correct the error. It may well be, therefore, that the left to right rule is very robust for a large proportion of 7 year old children.

Thus the work reported in this chapter has added more weight to the argument expressed in Chapters 5 and 6, i.e. that the ordering of presentation of information is a significant factor in the construction of relational representations. Chapter 8 will go on to investigate a possible cause for the 'premise ordering' effect.

However, we have obtained no evidence to support the claim that the children's difficulties were in part due to an inability to appreciate the redundancy of item information contained in the information they were given. It seems that this redundancy does not affect the integration of separate relations into a single systematic representation.

CHAPTER 8 : TASK CONSTRAINTS IN THE BUILDING OF STRUCTURAL REPRESENTATIONS (PART 4) CONCERNING DIMENSIONAL CONGRUENCE BETWEEN TASK AND INFORMATION PRESENTATION RELATIONSHIPS

Chapter 4 described a study (Experiment 1) which compared the performance of 5, 7 and 9 year old children when ordering an external array for the purposes of solving series problems. Only the 9 year old children were successful, however. The 7 year olds could order the array for some problems, but only after practice.

The ordered arrays are an external systematic representation of the relational structure inherent in series problems. It has been argued that the ability to correctly integrate relations into a single systematic representation is a necessary prerequisite for success in analogical reasoning tasks and could therefore explain the lack of consensus in the literature concerning the age at when children are able to perform analogical reasoning. Goswami (1989) provides evidence of early competence in these types of tasks, whereas there are many studies which report persistent difficulties into adulthood (Gick and Holyoak, 1980). Subsequent chapters in this thesis have therefore investigated some of the task constraints which could have been influencing the 7 year old's performance in Experiment 1.

A series of studies reported in Chapters 5 and 6 investigated the role of information ordering in the construction of an integrated array. It was found that the degree to which 7 year old children were required to violate a left to right ordering strategy significantly affected successful performance, both in terms of problem solving time and error data.

Details were also taken of the order in which the children placed the individual items. Collation of this data showed that the children relied almost exclusively on a left to right ordering strategy in the following ways:

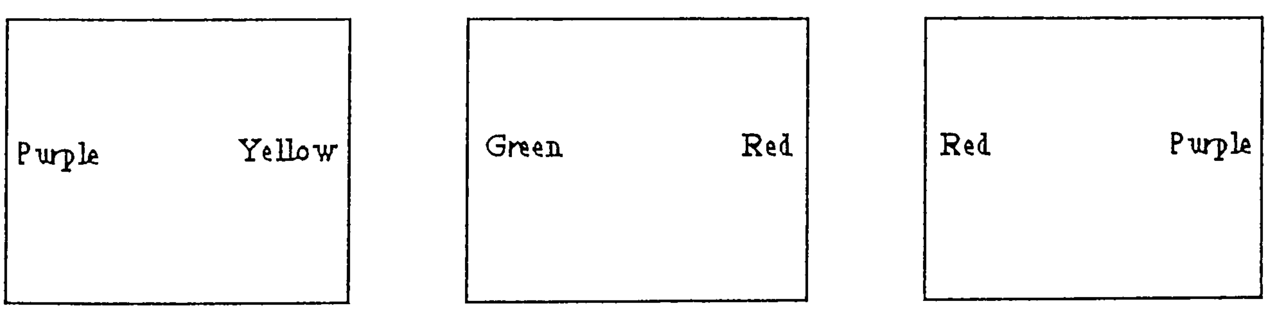
- they employed a left to right take up of information.

- spontaneous self corrections and errors involved erroneously placing an item at the end of a partially ordered array, rather than in front.

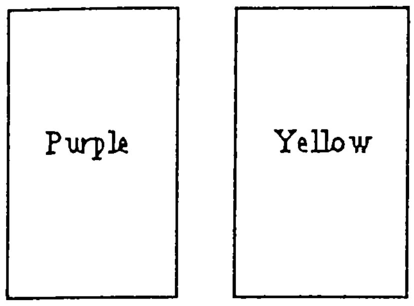
There was also evidence of 'premise splitting', particularly in CD AB BC tasks. This means that information initially given to the children in the premises was disregarded during the task solution. An example of this is given in Fig. 8.1 below.

Fig.6.1: Example of premise pair splitting (CD AB BC task).

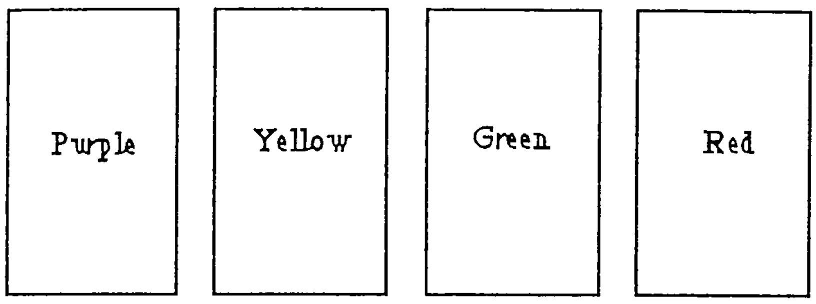
Premise information.



As subjects invariably use a left to right take-up of information, typical placements would be as follows:

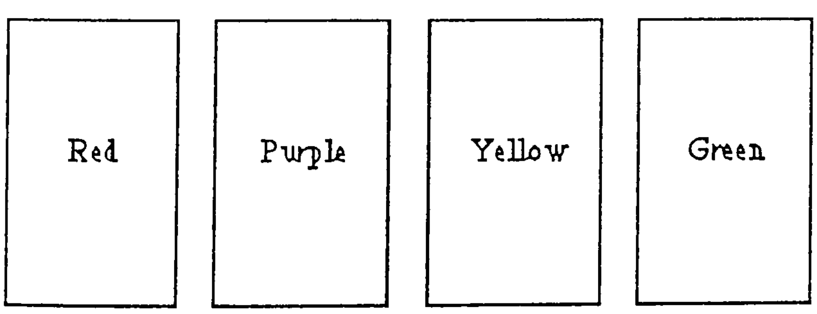


(from 1st premise)



(from 2nd premise)

At this stage, typically the subject would indicate that they had finished ordering the queue (this would of course be an incorrect ordering), or they would look at the 3rd premise and so reconsider their answer. It seemed that many of the children re-ordered the array by moving only the last character (red jumper) in order to comply with the 3rd premise. The children would then indicate that they had completed the task, with the queue being ordered as follows (still incorrect):



However, in re-ordering as above, the child has failed to take account of the information given in the 2nd premise, in that the red jumper is now no longer behind the green jumper. This type of error has been termed a 'premise pair split' because the subject has erroneously split characters which were given as immediately next to each other in the premise information.

Data from the CD AB BC task (Experiment 3) showed that 98% of the errors and self corrections involved the splitting of a 'given pair'. See Fig. 8.1 above for an example of this type of premise ordering.

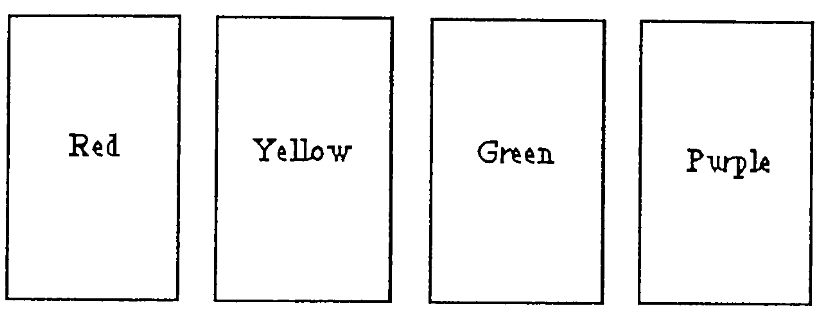
A further study in this series (Experiment 5) showed that the inappropriate left to right ordering strategy used by the 7 year olds was not affecting the performance of 9 year old children on the same tasks. These children did still have some difficulty with task ordering CD AB BC, however (see Experiment 6 for the relevant data and discussion). It has been suggested that this difficulty is due to 'premise pair splitting' (as discussed above).

There are at least two reasons for the 7 year olds use of an inappropriate left to right strategy when ordering the array. One concerns a lack of conceptual understanding of the relations between the premises. This would mean that the children have no procedural rules to guide them when linking a third item onto a premise pair. Because of this, they resort to a simple and commonplace left to right rule. Children are used to reading and writing from left to right, and so this ordering is imposed onto the current problem. This would explain the ceiling performance achieved in the **Insert** condition (AB CD BC) in Experiment 3, in that the subjects using this type of ordering to construct the array simply place their four character tokens in the same order as they occur in the premises (i.e. A B C D). This is done without appreciating the need to check the third premise (BC) as a verification that their ordering is correct.

However, a second possible reason concerns the interaction between the dimension of the relation with which the children were required to reason and the dimension in which the premises were presented. This has been mentioned briefly when discussing the results from Experiment 2 in Chapter 4.

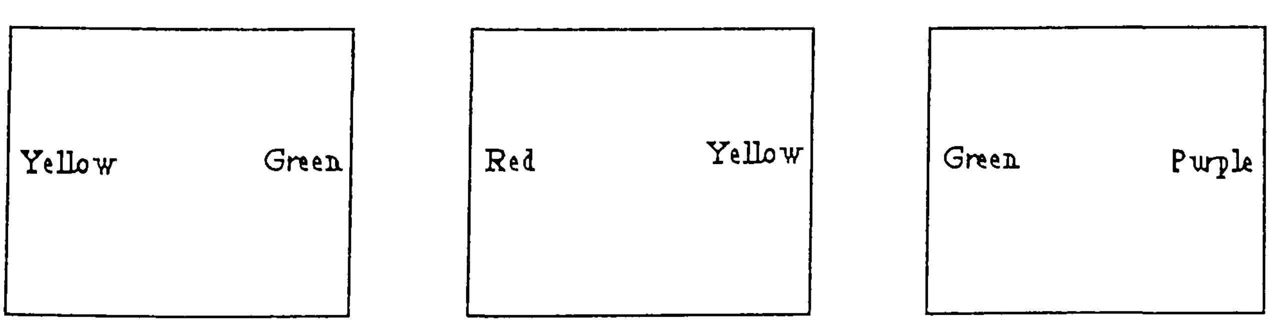
All of the studies to date (with the exception of Experiment 2) have required ordering in a horizontal dimension, and have typically used the context of people standing in a queue. Thus a typical integrated array would be as in Fig. 8.2 below.

Fig. 8.2: Example of a typical integrated array (taken from Experiment 3)



The subjects were presented with relational information from which they could construct the structural representation above. This was given in the form of paired characters. The information has always been presented to the children in a horizontal dimension, by being placed on a table in front of them. Fig. 8.3 shows a premise presentation, as seen by the subject, for the array shown in Fig. 8.2 above.

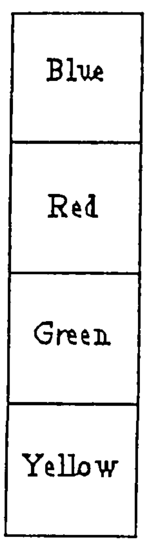
Fig. 8.3: Example of presented premise information for the array in Fig. 8.2 above.



Therefore, for all of the studies except Experiment 2, premise information has been presented horizontally and the children have been required to reason about a horizontal relationship.

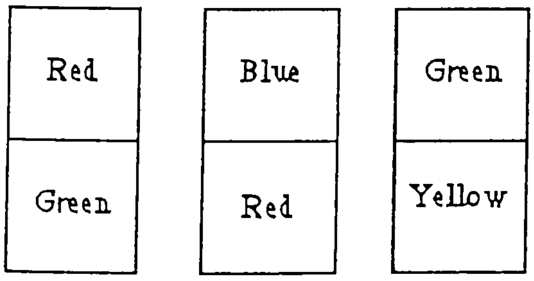
Experiment 2 was different in that the children were reasoning about the position of bricks in a tower (i.e. a vertical relationship). An example of a completed array is shown in Fig. 8.4 below.

Fig. 8.4: Example of a typical integrated array (taken from Experiment 2)



However, the premise information (in the form of paired bricks) was still presented to the children in a horizontal dimension. Fig. 8.5 shows a premise presentation for the array shown in Fig. 8.4 above.

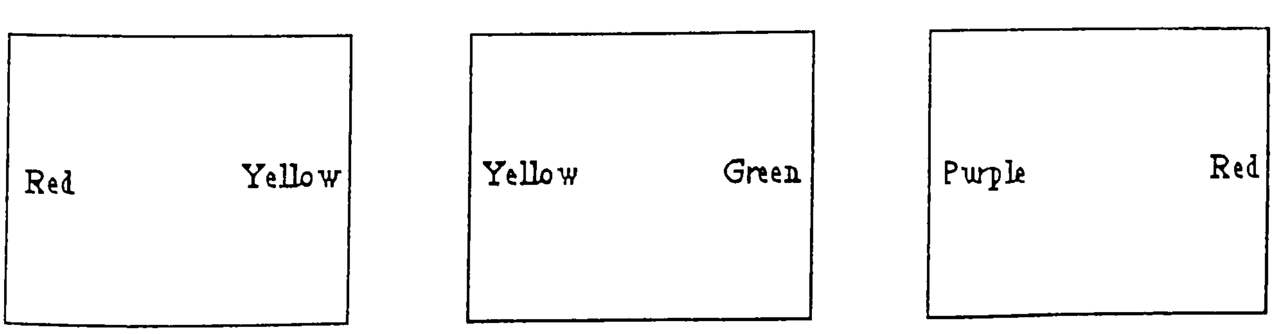
Fig. 8.5: Example of presented premise information for the array in Fig. 8.4 above

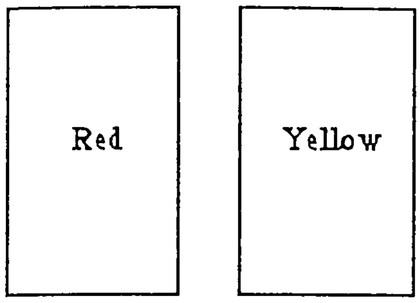


Experiment 2 was primarily carried out in order to replicate a previously reported study by Pears and Bryant (1990), so it investigated the performance of 5 year old children. Although Experiment 2 did not result in equivalent performance to that reported by Pears and Bryant, it did seem that better performance was achieved by the same age children working thc tower of bricks (vertical dimension) than with the characters in a queue (horizontal dimension used in Experiment 1). It could be that the difference in dimensions between the task relationship and the premise information presentation which existed in Experiment 2 might account for this.

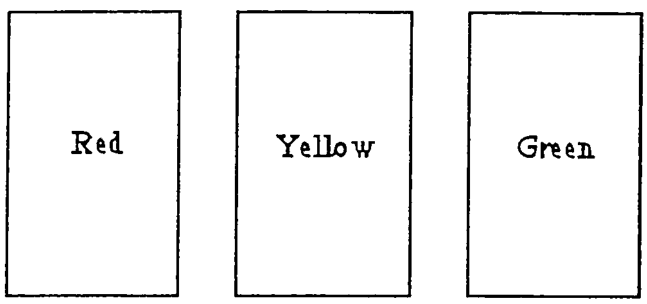
Because of the very poor performance shown by the 5 year old children in Experiment 1, the rest of this thesis has concentrated mainly on the abilities of 7 year old children. This is because my interests lie primarily with identifying inappropriate structuring strategies, rather than with developmental progression *per se*. In order for such strategies to be investigated, it was necessary to work with children who were displaying some degree of initial competence. However, it could well be that the hypothesised effects in 5 year olds of using the same task and premise presentation relationships might also result in inhibition of performance in 7 year old children. This is because the presence of congruent (i.e. the same) task and premise presentation relationships might adversely affect structural reasoning by cueing inappropriate left to right ordering strategies.

It is suggested that the reasons underlying this hypothesised effect are concerned with the psychological significance of the gap between the premise pairs. Consider the premise ordering below:

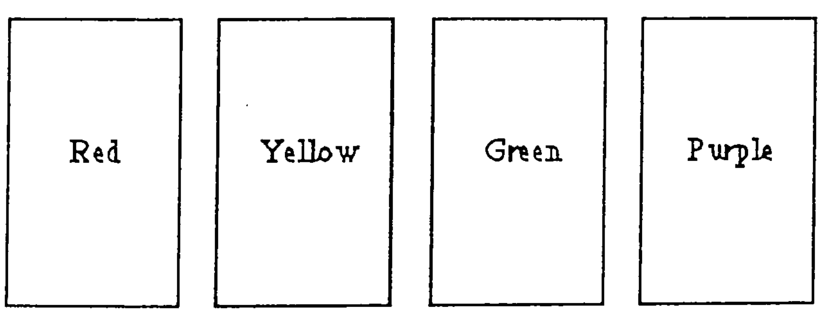




1st placement



2nd placement. This involves adding the green jumper (the novel item) onto the end of the partially ordered array. The next placement involves a decision about where to put the last item (the purple jumper). It appeared at the end of the premise ordering sequence, so is often placed at the end of the array. See below.



This incorrect placement of the purple jumper could be because the children did not realise that the gap between the 2nd and 3rd premises was salient, in that it signified two distinct and disjoined pieces of information, rather than one sequenced array.

The current chapter will begin therefore to consider the claim that it is the congruence of task and premise presentation relationships which causes inappropriate left to right ordering strategies. This occurs because the gap between premise pairs is not made salient when congruent relationships are used. In order to investigate the above, the following study has been designed to investigate the effects of making the gaps between premise pairs more salient. This will be carried out by making the spatial ordering of information incongruent with (different from) the actual dimension of the task relationship.

8.1 EXPERIMENT 8 - THE EFFECT OF VARYING THE SPATIAL DIMENSIONS OF TASK RELATIONSHIPS AND THE PRESENTATION OF PREMISE INFORMATION.

Rationale

The purpose of this study is to discover whether 7 year old children become more successful in constructing a systematic relational representation when they are forced to realise the significance of the gap between premises. This will be done by varying two dimensional factors (task relationship and premise presentation) in order to construct two congruent and two incongruent series problems. It is hypothesised that the incongruent conditions will result in facilitation of performance, as the requirement to order in a different spatial plane from that of the premise presentation will necessarily mean that the premises will be viewed as separate sections of 'paired-item' information, rather than a continuous and incremental piece of knowledge. Thus the inappropriate left to right ordering strategy observed in previous experiments will not be employed.

8.1.1 Method

Design

A two factorial between subject design was used. The two factors were as follows:

1. Type of spatial relationship used (horizontal or vertical).

2. Ordering of premise information (horizontal or vertical).

Thus there were four experimental groups.

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task.

Participants

The participants in this study were 30 mixed ability 7 year olds (mean age 7 years 2 months, range 6 years 7 months to 7 years 7 months) from state first and primary schools, with predominantly middle class catchment areas. The subjects were from the same schools as those in Experiment 3, and were randomly assigned to one of three experimental groups (relation horizontal, presentation vertical; relation vertical, presentation vertical; relation vertical, presentation horizontal).

The data for the fourth group (relation horizontal, presentation horizontal) was taken from Experiment 3 (Level L-R, Drawings). This is because the two experiments shared common methodologies, experimenters and subject pools.

Thus there were ten subjects in each experimental group.

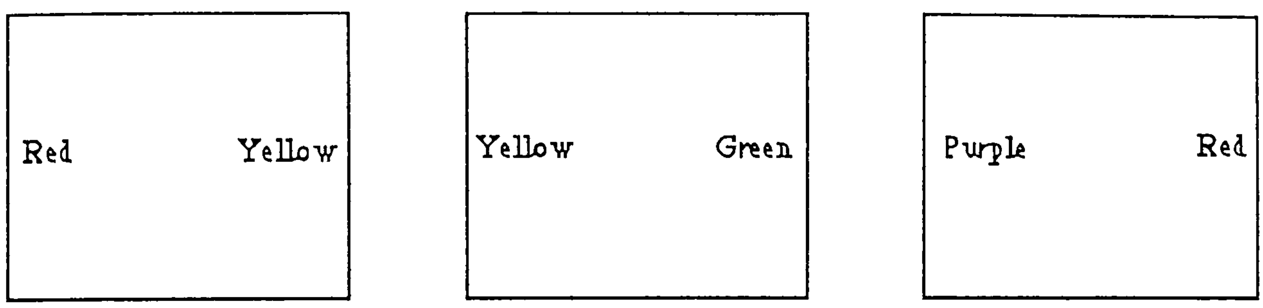
Task description

Horizontal spatial relationship

As in Experiment 3, except that all the subjects worked with Level L-R task ordering. Copies of the task materials are given in Appendix C.

The subjects in the horizontal presentation group were presented with the premises laid out as follows :- BC CD AB. An example of this is given in Fig. 8.6 below.

Fig. 8.6: Example of horizontal spatial relationship and horizontal presentation



The subjects in the vertical presentation group were presented with the premises laid out as follows :-

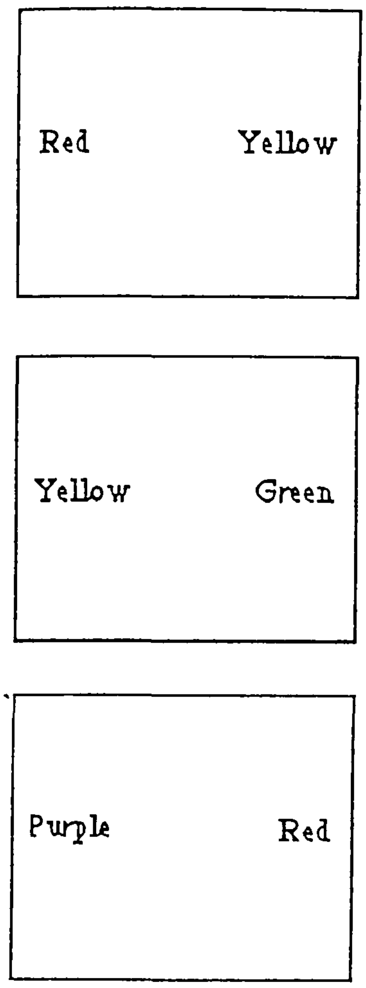
BC

CD

AB

An example of this is given in Fig. 8.7 below.

Fig. 8.7: Example of horizontal spatial relationship and vertical presentation



Vertical spatial relationship

The relation used was 'on top of'. This was based on work done by Pears and Bryant (discussed in Chapters 3 and 4). As in the horizontal spatial relationship, 4 item orderings were used. All the items were front views of girls, distinguishable from each other by wearing different coloured T-shirts. Thus a complete vertical ordering would be 4 girls standing on each others shoulders, all facing forwards (the subjects were told that the orderings were of children learning to be acrobats). As in the horizontal relationship, the orders of colours used were randomly varied, so as to give 6 different orderings of four colours (yellow, green, red, purple). Copies of the task materials are given in Appendix F.

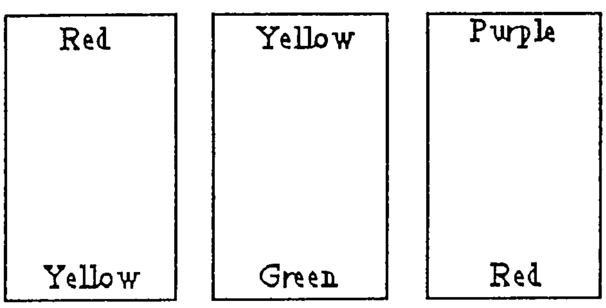
Each subject completed six tasks. The tasks were ordered in the same way as those in the horizontal relationship condition.

The subjects in the horizontal presentation group were presented with the premises laid out as follows :- B C A

C D B

An example of this is given in Fig. 8.8 below.

Fig. 8.8: Example of vertical spatial relationship and horizontal presentation



The subjects in the vertical presentation group were presented with the premises laid out as follows :-

B

C

C

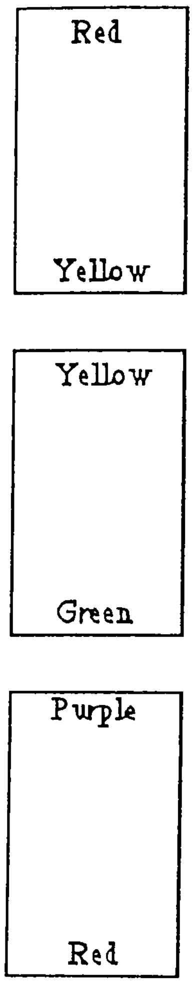
D

A

B

An example of this is given in Fig. 8.9 below.

Fig. 8.9: Example of vertical spatial relationship and vertical presentation



Materials

As in Experiment 3, but only drawings were used (either of women viewed 'side-on' in a queue, or acrobats standing on each others shoulders). Copies of the task materials are given in Appendices C and F.

Procedure

As in Experiment 3 except that reference was made to acrobats standing on each others shoulders, rather than people in a queue, where appropriate. As previously described, the premises were presented in a horizontal array for half of the subjects working with each of the spatial relationships, and in a vertical array for the other half.

The WISC-R digit span subtest was administered to the participants.

8.1.2 Results

Note Each subject completed six experimental trials. However, a large number of the subjects in the previous studies erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier ones the first trial from each subject has been discounted.

WISC scores

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (relation horizontal, presentation vertical)	8.9
Group 2 (relation vertical, presentation vertical)	8.6
Group 3 (relation vertical, presentation horizontal)	8.8
Group 4 (relation horizontal, presentation horizontal)	9.2

The average performance on the digit span test for ages between 7 years and 7 years 3 months is a score of 8 (WISC-R Manual, 1974).

Number of correct answers

Table 8.1 shows the mean number of correct answers for each experimental group.

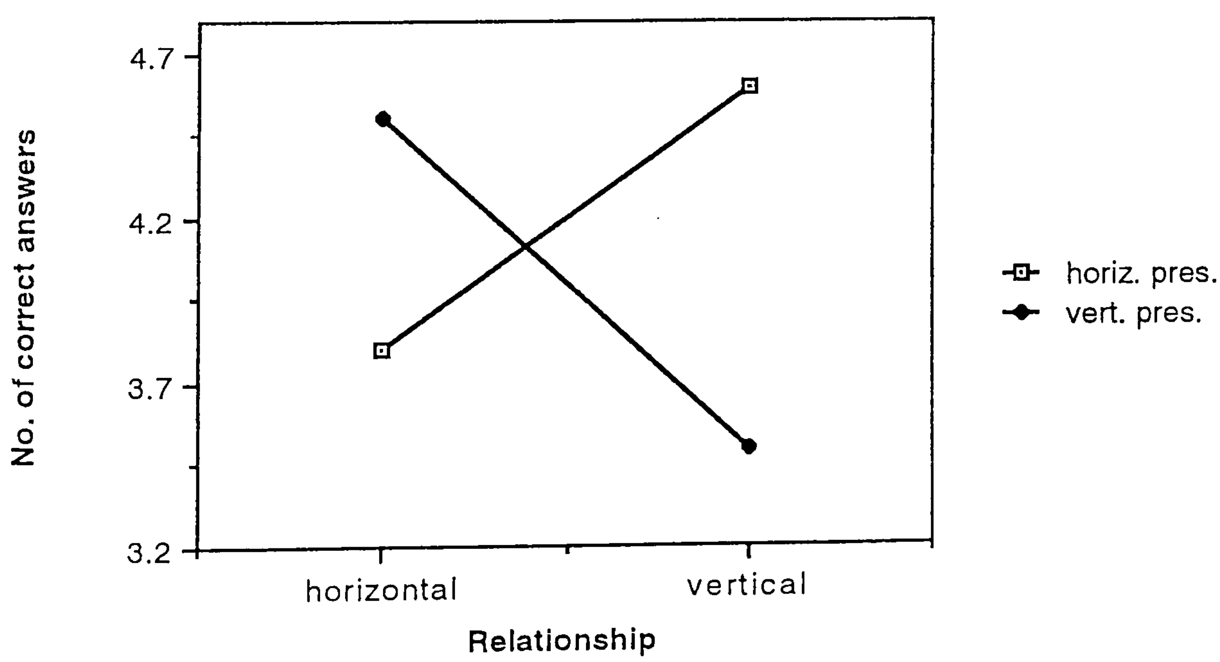
Table 8.1: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

		Presentation	
		Horizontal	Vertical
Relationship	Horizontal	3.8 (1.23)	4.5 (0.53)
	Vertical	4.6 (0.52)	3.5 (0.97)

An ANOVA [2 (relationship) x 2 (presentation), both between subjects factors] was carried out on the above data. This revealed a significant interaction between the two factors ($F_{[1, 36]}$=10.80, p<0.01 - see Fig. 8.10), with no main effects.

Fig. 8.10: Relationship x presentation interaction (number of correct answers)



Further analysis of the above data revealed the following significant simple main effects:

- the subjects working with the vertical relation and horizontal premise presentation scored more correct answers than those working with the vertical relation and vertical premise presentation ($F_{[1, 36]}=8.067$, $p<0.01$).

- the subjects working with the vertical premise presentation and the horizontal relationship scored more correct answers than those working with the vertical premise presentation and the vertical relationship ($F_{[1, 36]}=6.667$, $p<0.05$).

- the subjects working with the horizontal premise presentation and the vertical relationship scored more correct answers than those working with the horizontal premise presentation and the horizontal relationship ($F_{[1, 36]}=4.267$, $p<0.05$).

Note The difference between the number of correct answers from those subjects working with the horizontal relation and vertical premise presentation and those working with the horizontal relation and horizontal premise presentation approached significance ($F_{[1, 36]}=3.267$, $p=0.0791$).

Time taken to complete trials (correct answers only)

Table 8.2 shows the mean time taken (in seconds) for each experimental group.

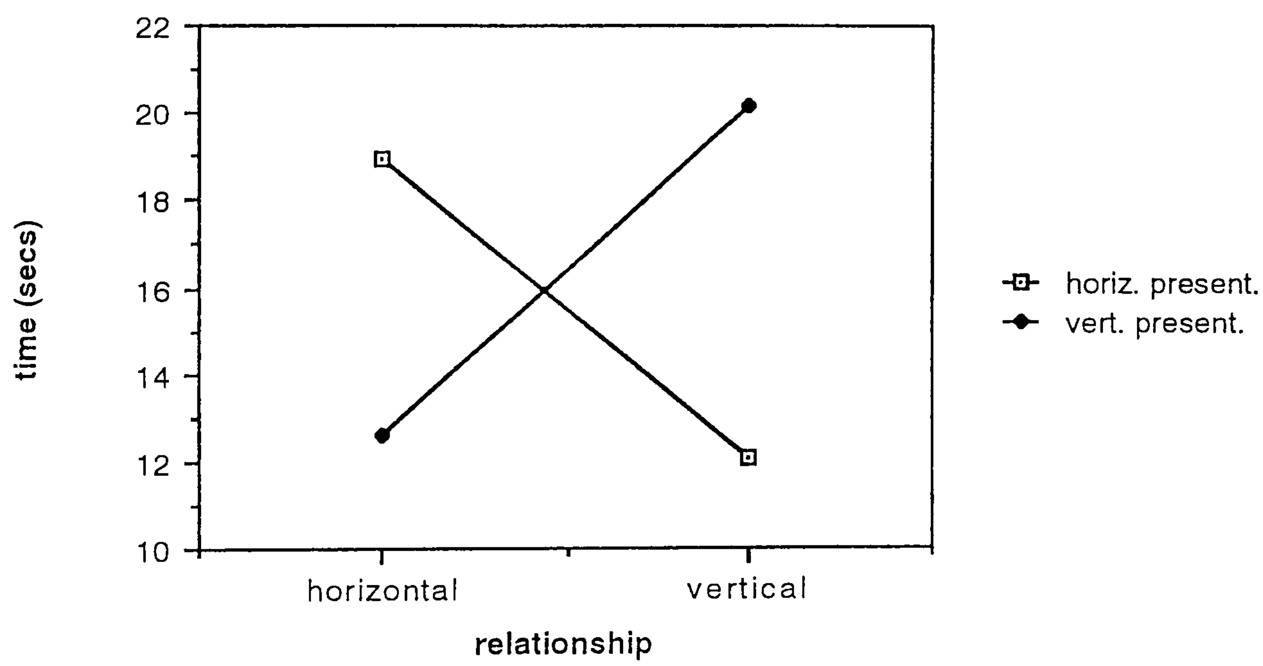
Table 8.2: Mean time taken to complete trials (correct answers only)

Standard deviations are shown in parentheses

		Presentation	
		Horizontal	**Vertical**
Relationship	**Horizontal**	18.92 (3.94)	12.59 (1.50)
	Vertical	11.97 (1.41)	20.16 (1.70)

An ANOVA [2 (relationship) x 2 (presentation), both between subjects factors] was carried out on the above data. Again, this revealed a significant interaction between the two factors ($F_{[1, 36]}=93.322$, $p<0.01$ - see Fig. 8.11, with no main effects.

Fig. 8.11: Relationship x presentation interaction (time taken to correctly solve tasks)



Further analysis of the above data revealed the following significant simple main effects:

- the subjects working with the horizontal relation and vertical premise presentation produced correct solutions more quickly than those working with the horizontal relation and horizontal premise presentation ($F_{[1, 36]}=35.518$, $p<0.01$).

- the subjects working with the vertical relation and horizontal premise presentation produced correct solutions more quickly than those working with the vertical relation and vertical premise presentation ($F_{[1, 36]}=59.321$, $p<0.01$).

- the subjects working with the horizontal premise presentation and the vertical relationship produced correct solutions more quickly than those working with the horizontal premise presentation and the horizontal relationship ($F_{[1, 36]}=42.749$, $p<0.01$).

- the subjects working with the vertical premise presentation and the horizontal relationship produced correct solutions more quickly than those working with the vertical premise presentation and the vertical relationship ($F_{[1, 36]}=50.744$, $p<0.01$).

Comparison with data from the Insert condition (Experiment 3)

Because the subjects in Experiment 3 and this study were taken from the same schools, and the same type of materials and procedure were used throughout, it was decided to combine the relevant data from the two studies in a direct statistical comparison. Due to results obtained in Experiments 3 and 5, the scores obtained for both measurements in the Insert condition have been deemed to be at or approaching ceiling, and so a direct comparison of these 'base-line' scores with the current study will provide evidence as to whether the experimental manipulation has facilitated performance to 'ceiling' levels. Experiment 3 was a two factor experiment, where one of the factors involved a comparison between using photographs and drawings as stimulus material. No significant differences were found, and so this has not been investigated further. Only the data from the children who used the drawings in the **Insert** condition has been used in this comparison, so as to equate with the stimulus material used in the current study.

Number of correct answers

Table 8.3: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

Ordering		
rel. horiz./pres. vert.	**rel.vert./pres. horiz**	**Insert (Exp. 3)**
4.5 (0.53)	4.6 (0.52)	4.9 (0.32)

A one way ANOVA was carried out on the above data and showed no significant effect of ordering ($F_{[1, 36]}=2.017$, $p>0.05$).

Time taken to complete trials (correct answers only)

Table 8.4: Mean time taken to complete trials (in seconds)

Standard deviations are shown in parentheses

Ordering		
rel. horiz./pres. vert.	**rel.vert./pres. horiz**	**Insert (Exp. 3)**
12.59 (1.50)	12.06 (1.44)	10.23 (1.77)

A one way ANOVA was carried out on the above data and showed a significant effect of ordering ($F_{[2, 27]}=6.137$, $p<0.05$). Tukey comparisons showed that the Insert ordering (Experiment 3) resulted in correct trials being completed more quickly than both the relation horizontal/presentation vertical condition ($q=4.72$, $p<0.01$) and the relation vertical/presentation horizontal condition ($q=3.67$, $p<0.05$).

8.1.3 Discussion

The results obtained from this experiment have shown that the poor performance demonstrated by the 7 year old children in Experiments 1 and 3 can be facilitated by manipulating the dimension of the task and presentation relations. When these two factors are incongruent, that is one is in a vertical and one in a horizontal dimension, performance is significantly better than when the factors are congruent, that is both are in either a vertical or a horizontal dimension. This is portrayed diagrammatically in Fig. 8.12 below.

Fig. 8.12: Diagrammatic representation of findings (current experiment)

		Task relationship	
		Vertical	**Horizontal**
Premise presentation	**Vertical**	**B C C D A B**	**BC CD AB**
	Horizontal	**B C A C D B**	**BC CD AB**

Note The double border denotes incongruent dimensional relations and successful performance.

When the task was presented using the two incongruent conditions the error rate was not significantly different from that of the **Insert** condition in Experiment 3. The time which subjects took to correctly solve incongruent problems was higher than that of the

Insert condition, however. This is not surprising if we consider the difference between the two types of orderings. Insert ordering consists of AB CD BC. Thus the child can straightaway order using the first two premises and needs to use the third premise only as confirmation. L-R ordering (that used in the current experiment) is either BC CD AB or BC AB CD. These both require an item to be placed at the front of the array. This still results in a slight tendency for the children to follow an inappropriate left to right heuristic. This is easily realised and corrected, as can be seen by the equivalent error rates, but it takes a few seconds longer, hence the difference in problem solving time.

It seems therefore that when children are not confused by task features, they are able to integrate premises successfully. This shows that they must have knowledge that separate relations can be linked together to form a systematic representation. They are misled, however by task information which does not make the critical features of a problem salient. Thus performance was inhibited by a lack of appreciation of the significance between premises. As soon as these gaps were made salient by altering the dimension of the premise information, the children were successful.

This result is interesting, especially if we consider it in terms of the literature describing the use of 'scaffolding' in both formal and informal education (Wood, Bruner and Ross, 1976). This has been reviewed more fully in Chapter 2. To summarise, Wood (1988) defines the term 'scaffolding' as "the breaking down of a task into a sequence of smaller tasks which children can manage to perform" and "highlighting things they need to take account of" (p. 80). An initial consideration of this literature might suggest that the placement of premise information in the 'correct' spatial dimension for task reasoning would help the subject, as this means that some of the work towards achieving the task goal will have been carried out. However, the results from the current study demonstrate that the 'smaller tasks' must be carefully selected so as not to cue the children into using inappropriate strategies. This is an important issue, and will be discussed more fully in Chapter 10.

We must also consider the generalisability of these results. The goal in this study was to achieve the correct ordering of either women or children in a queue or a tower. It could be said that these type of tasks are very unfamiliar to young children, and that perhaps the cause of their difficulties stems from this unfamiliarity.

When presented with a novel or unfamiliar task, an often sensible action is to import a well-used strategy from a similar domain. Children of this age spend a lot of their time in school focusing on the basic skills of reading, writing and arithmetic, where 'add to end' orderings are extremely common. For example, in the domain of addition, children solve problems of the form $3 + 2 = ?$, where the action they are required to perform comes at the end of a string of symbols. Also, when writing a story, young children usually approach the task in a sequential manner, and the story is constructed by more text being added to the end of what has already been written. The domains of reading, writing and arithmetic have some features in common with the series problems used in the current study. All are concerned with the ordering of symbols, and they were all carried out in a formal educational setting (the experiments were carried out in a school classroom, in close proximity to the child's usual work-area). Due to the unfamiliarity of the experimental task, the children may have imported the familiar 'add to end' ordering which is usually appropriate for the task requirements which they encounter during formal education.

There is ample evidence (Donaldson, 1978) that children's performance in unfamiliar laboratory tasks can often be significantly improved by the use of equivalent contextualised 'real-world' tasks which are meaningful and motivating. A previous experiment (Experiment 4) found that performance was not improved by the use of a familiar task when both task relationship and premise presentation used a horizontal dimension. Thus we have already shown that there is no 'familiarity effect' when both task and premise presentation relationships are congruent (both in a horizontal dimension). However, because the results obtained from the current study are both interesting and important, it was considered necessary to investigate the effects of reasoning with a familiar task when both task and premise presentation relationships use a vertical relationship and also when the two relationships differ from each other (i.e. one is in a vertical dimension and the other horizontal).

8.2 EXPERIMENT 9 - THE EFFECT OF REASONING WITH CONGRUENT AND INCONGRUENT CONDITIONS USING A FAMILIAR TASK.

Rationale

This study has been designed to ascertain whether the difference obtained between congruent and incongruent task relationships and premise presentations are also present when children reason with a familiar 'real-world' task. If the comments above, concerning the effects of unfamiliar tasks on children's tendencies to import an inappropriate strategy are valid, then we might expect to see no significant differences between congruent and incongruent conditions when reasoning with a familiar task. This is because incongruence will then only be of consequence for unfamiliar tasks. On the other hand, if the use of an inappropriate ordering strategy occurs also when the task is familiar, then we will obtain a similar pattern of results as that obtained in Experiment 8.

After discussion with the class teachers, it was decided to use the 'LEGO' train (see Experiment 4) for the horizontal task and a toy crane with which the children were familiar for the vertical condition. Both toys were well used by the children, and they were happy to insert carriages or crates into an ordering for various reasons, whilst playing with the toys.

8.2.1 Method

Design

A two factorial between subject design was used. The two factors were as follows:

1. Type of spatial relationship used (horizontal or vertical).

2. Ordering of premise information (horizontal or vertical).

Thus there were four experimental groups.

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task.

Participants

The participants in this study were 40 mixed ability 7 year olds (mean age 6 years 11 months, range 6 years 6 months to 7 years 5 months) from a state and primary school, with a predominantly middle class catchment area. The subjects were from one of the schools used in Experiment 2, and were randomly assigned to one of four experimental groups (relation horizontal, presentation horizontal; relation horizontal, presentation vertical; relation vertical, presentation vertical; relation vertical, presentation horizontal).

Task description

Horizontal spatial relationship

As in Experiment 8, except that the task was to order different coloured carriages from a toy train, using appropriately coloured drawings of the carriages as premise information. Copies of the task materials are given in Appendix E. The order of the four colours used (red, yellow, blue and green) was randomly varied, so as to give six different orderings. Examples of the two different premise orderings (horizontal and vertical) are given below.

Fig. 8.13: Example of horizontal spatial relationship and horizontal presentation.

Blue Red	Green Blue	Red Yellow

Fig. 8.14: Example of horizontal spatial relationship and vertical presentation.

Blue Red
Green Blue
Red Yellow

Vertical spatial relationship

As in Experiment 8, except that the task was to order different coloured crates from a toy crane, using appropriately coloured drawings of the crates as premise information.

Copies of the task materials are given in Appendix G. Again, the order of the four colours used (red, yellow, blue and green) was randomly varied, so as to give six different orderings. Examples of the two different premise presentations are given below.

Fig. 8.15: Example of vertical spatial relationship and horizontal presentation.

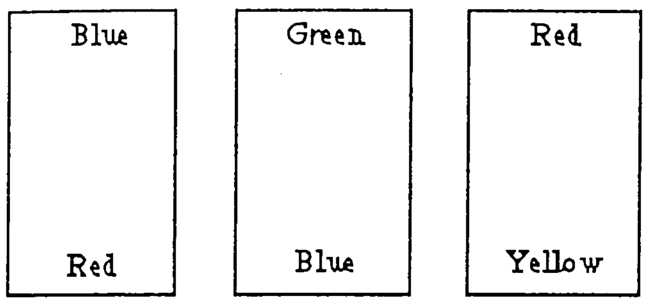
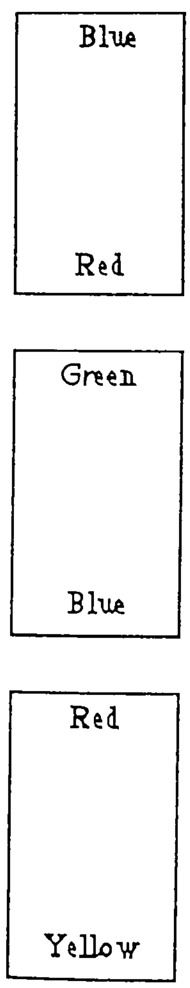


Fig. 8.16: Example of vertical spatial relationship and vertical presentation.



Materials

Horizontal task relationship

A toy 'LEGO' train, with an engine and four different coloured carriages, with a character facing towards the front of each carriage, was used as the 'ordered array'. The children were handed the carriages and asked to put them in the right order. Premise information was presented using 3 drawings of pairs of carriages, coloured appropriately. Appendix E shows an example of premise information andthe apropriately ordered array.

Vertical task relationship

A toy crane with four different coloured crates which could be suspended in any order from the main hook was used as the 'ordered array'. The children were handed the crates and asked to hang them in the right order. Premise information was presented using 3 drawings of pairs of crates, coloured appropriately. Appendix G shows an example of premise information and the apropriately ordered array.

Procedure

As in Experiment 8, except that reference was made to carriages being placed in order behind the engine for the horizontal task relationship conditions and crates being hung from the crane in the correct order for the vertical task relationship conditions. As previously described, the premises were presented in a horizontal array for half of the subjects working with each of the spatial relationships, and in a vertical array for the other half.

8.2.2 Results

Note Each subject completed six experimental trials. However, a large number of the subjects in the previous studies erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier ones the first trial from each subject has been discounted.

WISC scores

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (relation horizontal, presentation vertical)	8.3
Group 2 (relation vertical, presentation vertical)	8.8
Group 3 (relation vertical, presentation horizontal)	8.2
Group 4 (relation horizontal, presentation horizontal)	8.4

The average performance on the digit span test for ages between 6 years 8 months and 7 years is a score of 8 (WISC-R Manual, 1974).

Number of correct answers

Table 8.5 shows the mean number of correct answers for each experimental group.

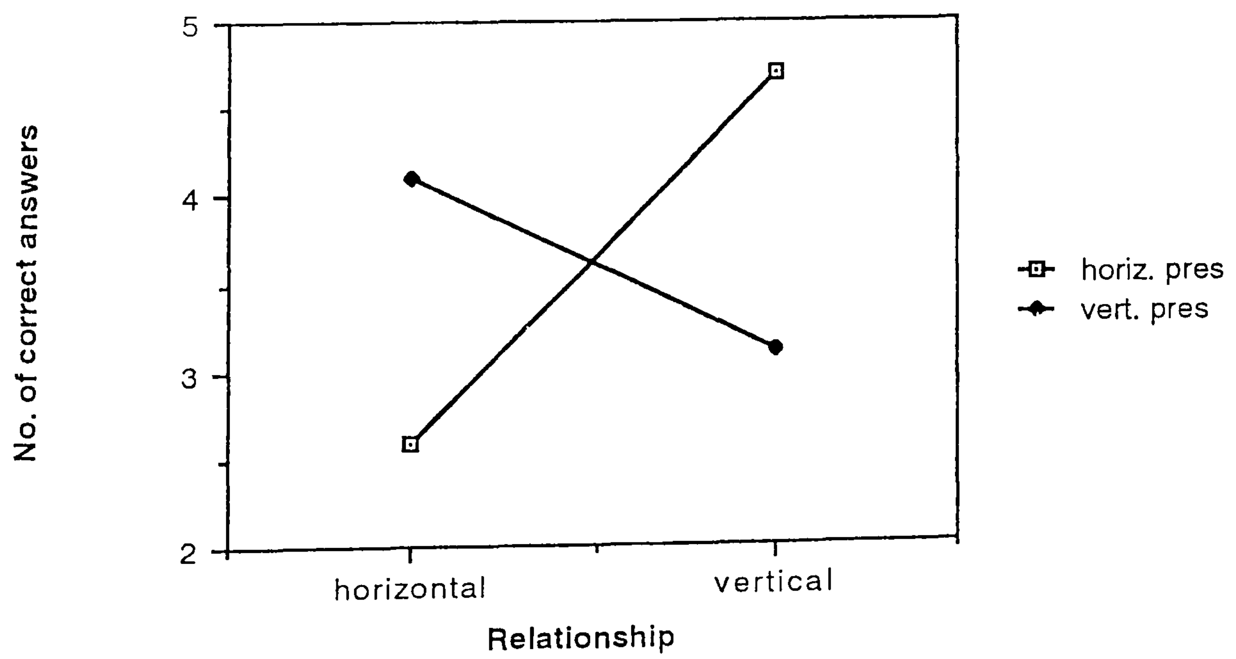
Table 8.5: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

		Presentation	
		Horizontal	**Vertical**
Relationship	**Horizontal**	3.3 (0.82)	4.4 (0.7)
	Vertical	4.7 (0.48)	3.1 (0.88)

An ANOVA [2 (relationship) x 2 (presentation), both between subjects factors] was carried out on the above data. This revealed a significant interaction between the two factors ($F_{[1, 36]}$=33.646, p<0.01 - see Fig. 8.17), with no main effects.

Fig. 8.17: Relationship x presentation interaction (number of correct answers)



Further analysis of the data revealed the following significant simple main effects:

- the subjects working with the horizontal relation and vertical premise presentation produced more correct solutions than those working with the horizontal relation and horizontal premise presentation ($F_{[1, 36]}=11.169$, $p<0.01$).

- the subjects working with the vertical relation and horizontal premise presentation produced more correct solutions than those working with the vertical relation and vertical premise presentation ($F_{[1, 36]}=23.631$, $p<0.01$).

- the subjects working with the horizontal premise presentation and the vertical relationship produced more correct solutions than those working with the horizontal premise presentation and the horizontal relationship ($F_{[1, 36]}=18.092$, $p<0.01$).

- the subjects working with the vertical premise presentation and the horizontal relationship produced more correct solutions than those working with the vertical premise presentation and the vertical relationship ($F_{[1, 36]}=15.600$, $p<0.01$).

Thus, for all comparisons, incongruent task and premise presentation relationships were significantly better than congruent ones.

Time taken to complete trials (correct answers only)

Table 8.6 shows the mean time taken (in seconds) for each experimental group.

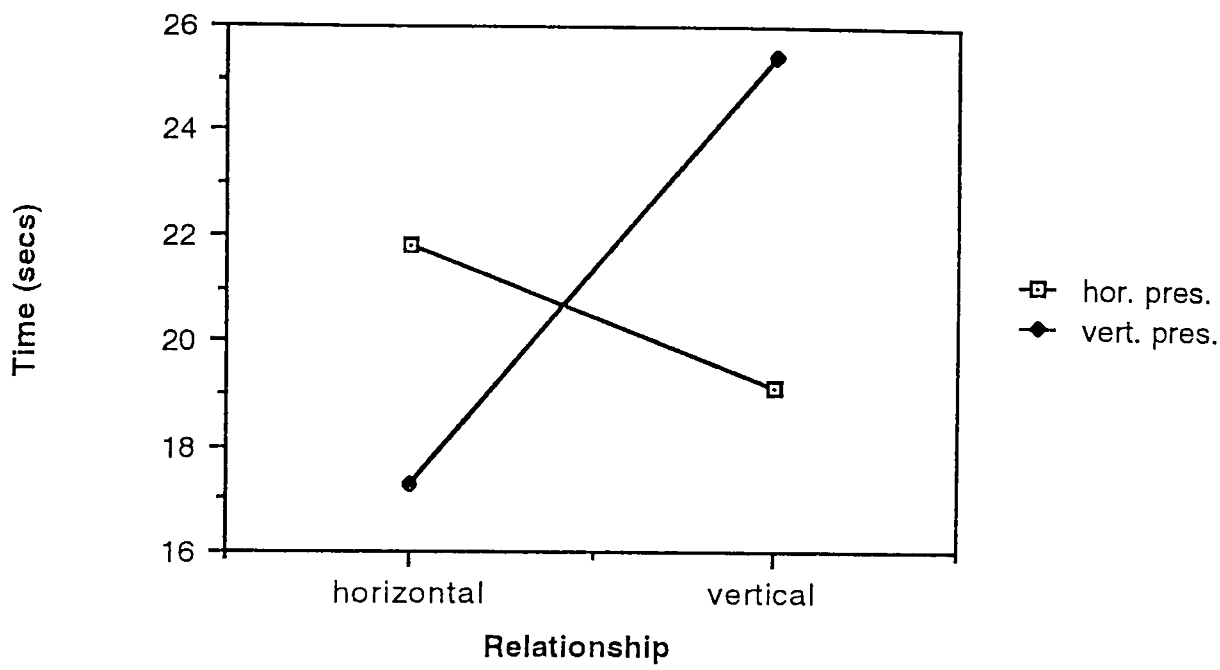
Table 8.6: Mean time taken to complete trials (correct answers only)

Standard deviations are shown in parentheses

		Presentation	
		Horizontal	**Vertical**
Relationship	**Horizontal**	21.79 (2.48)	17.26 (1.17)
	Vertical	19.10 (1.26)	25.40 (2.45)

An ANOVA [2 (relationship) x 2 (presentation), both between subjects factors] was carried out on the above data. Again, this revealed a significant interaction between the two factors ($F_{[1, 36]}$=77.875, p<0.01 - see Fig. 8.18). There was also a main effect of task relationship ($F_{[1, 36]}$=19.757, p<0.01) such that those children working with the horizontal relationship produced correct solutions more quickly.

Fig. 8.18: Relationship x presentation interaction (time taken to correctly solve tasks)



Further analysis of the data revealed the following significant simple main effects:

- the subjects working with the horizontal relation and vertical premise presentation produced correct solutions more quickly than those working with the horizontal relation and horizontal premise presentation ($F_{[1, 36]}$=27.280, p<0.01).

- the subjects working with the vertical relation and horizontal premise presentation produced correct solutions more quickly than those working with the vertical relation and vertical premise presentation ($F_{[1, 36]}$=52.663, p<0.01).

- the subjects working with the horizontal premise presentation and the vertical relationship produced correct solutions more quickly than those working with the horizontal premise presentation and the horizontal relationship ($F_{[1, 36]}$=9.591, p<0.01).

- the subjects working with the vertical premise presentation and the horizontal relationship produced correct solutions more quickly than those working with the vertical premise presentation and the vertical relationship ($F_{[1, 36]}=88.041$, $p<0.01$).

Again, for all comparisons, incongruent task and premise presentation relationships were significantly better than congruent ones.

8.2.3 Discussion

The results obtained from this study replicate those from Experiment 8, with one exception i.e. there was a main effect of task relationship for one of the dependent variables. This showed that the subjects produced correct solutions more quickly when working with the horizontal relationship. There was however no main effect of task relationship for the number of correct solutions produced.

It is suggested that the apparent increase in problem solving time for the children working with the vertical task (toy crane) is actually due to the design of the toy used. Hooks and eyes were used to enable the crates to be linked together (a hook was situated on the end of the crane arm and on the bottom of each crate, together with an eye on the top of each crate). However, this meant that the top and bottom of the blocks were not interchangeable, but that the differences between them were not sufficiently salient. The experimenter noticed a marked tendency for the children to become very confused when they couldn't link adjacent blocks as they had the two eyes next to each other. This was not a problem for the children working with the toy train, as the front of the carriage was very clearly marked by the use of a model person facing forwards.

In other aspects, this study replicates Experiment 7. For both measures, and for all comparisons, incongruent task and premise presentation relationships were significantly better than congruent ones.

It could be that this effect is due to an inappropriate choice of 'real task', and that with different objects to work with, performance would not have been inhibited by congruent relationships. This appears unlikely, however. As in Experiment 4, the children who used the toys certainly seemed more 'at home' with the task, and there were no puzzled expressions or inappropriate questions, as there were with the children who used the queue and tower tasks. It could be that the presentation of premise information using real objects, rather than drawings, would have an effect, although it is not obvious exactly how this would facilitate performance. On the contrary, asking the children to order using duplicates of the objects used in the premises might cause some confusion.

It seems, therefore, that the 'congruence constraint' observed in Experiment 7 is not due to familiarity or representational factors. This means that the construction of an integrated array is inhibited when children are required to reason using the spatial relationship as the one in which the task information is presented (i.e. the two relationships are congruent). Performance can be improved by using incongruent relationships, thus making the gaps between different premise pairs salient. This 'congruence constraint' experienced by 7 year old children is still operational when they work with familiar task materials which are real objects rather than representations of objects. The evidence from the current study, together with that from Experiment 8, suggests that this constraint is very salient for a large proportion of 7 year olds.

In summary, Experiments 8 and 9 demonstrated the existence of a 'congruence constraint' in the solving of series problems using dimensional relationships. Congruent task and premise presentation relationships significantly inhibited performance when compared with incongruent ones. As a result of these findings it has been suggested that this constraint is preventing the successful building of structural representations of problems.

Both of these studies have used 7 year old children, as the level of complexity of the task was considered appropriate for this age group. However, the first study in this thesis (Experiment 1) looked at the performance of 5 year old children when building a structured array very similar to that used in the later studies. The subjects in this study found the task extremely difficult (performance was approaching 'floor'), and for that reason all the following experiments used only older children (7 and 9 year olds). The task used with the 5 year old children in Experiment 1 was to order a five person array using the horizontal relationship 'behind' and horizontal premise presentation.

Based on the results from Experiments 8 and 9 we can now suggest two potential manipulations which might improve the 5 year old's performance :

1. Number of items in the array.

   It has always seemed likely that a reduction in the number of array items will facilitate performance, by reducing the amount of working memory required to solve the task. This is discussed more fully in Chapter 4. However, the constraints which are operating in this situation can be divided into two types i. e. those arising primarily from the problem solving capabilities of the child and those arising primarily from the task presentation, though obviously these will interact with each other. As the emphasis of this research programme has been to look at task constraints, a reduction in the number of array items has not been previously investigated.

2. Removal of the 'congruence constraint'

Congruent task and premise presentation relationships significantly inhibited performance for the 7 year olds on a very similar task, so it could be that the removal of this constraint for younger children will also improve their performance.

In view of the above, it was decided to replicate Experiment 8 (a comparison of congruent task relationships and premise presentations with incongruent), but using a younger age group (5 year olds rather than 7 year olds). However, because of the likely working memory constraint, it was decided to use 3 item arrays rather 4 items. This was to give the younger children a much better chance of achieving some success (i.e. not performing at 'floor').* Nonetheless, if we see significant facilitation of performance in the congruent conditions, we can conclude that the 'congruence constraint' is not limited to 7 year old children.

8.3 EXPERIMENT 10 - THE EFFECT OF REASONING WITH CONGRUENT AND INCONGRUENT CONDITIONS IN 5 YEAR OLD CHILDREN

Rationale

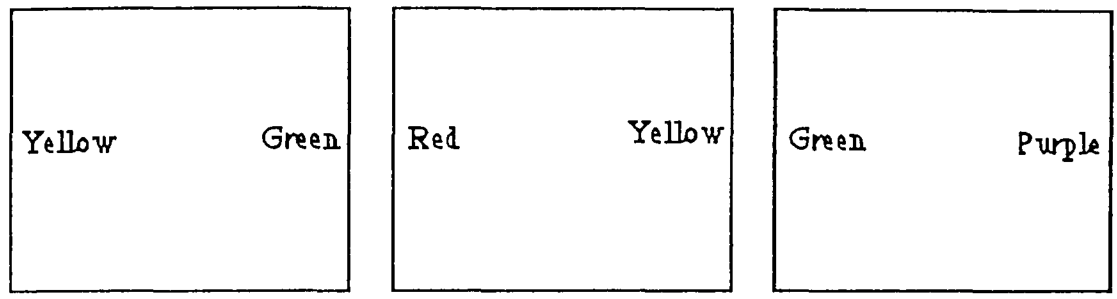
The aims of this study are to investigate whether the poor performance demonstrated by the 5 year old children when building the structured array in Experiment 1 can be facilitated by the removal of the 'congruence constraint'. This manipulation has been showed to improve performance in 7 year old children when solving a similar task (see Experiment 8 above). The following experiment will enable us to ascertain whether the 'congruence constraint' identified in Experiment 8 is limited to 7 year olds. The term

* a pilot study has indicated that 5 year olds have some success with ordering 3 item arrays.

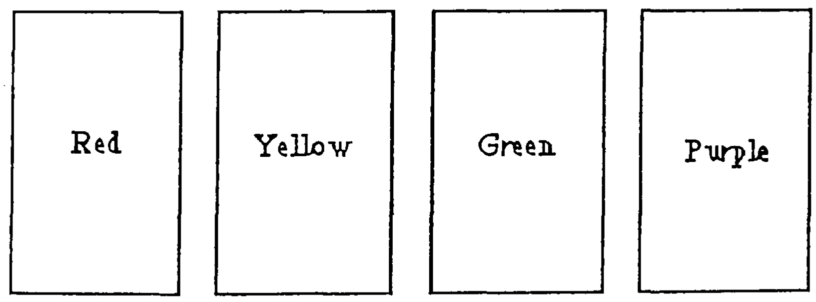
'congruence constraint' refers to the congruent dimension which was used for both the task relationship and for the presentation of the premise information. This is illustrated in Fig. 8.19 below.

Fig. 8.19: Example of congruent dimensions for task relationship and premise information presentation

Premise presentation



Correctly completed array (described to the subjects as a bus queue)



In the figure above, the premises are presented in a horizontal manner, and the relationship between individual characters which the subject is required to reason about is also horizontal.

8.3.1 Method

Design

A two factorial between subject design was used. The two factors were as follows:

1. Type of spatial relationship used (horizontal or vertical).

2. Ordering of premise information (horizontal or vertical).

Thus there were four experimental groups.

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task.

Participants

The participants in this study were 40 mixed ability 5 year olds (mean age 5 years 1 month, range 4 years 11 months to 5 years 4 months) from a state nursery school, with a predominantly middle class catchment area. The subjects were from the one of the same schools used in Experiment 3, and were randomly assigned to one of four experimental groups (relation horizontal, presentation horizontal; relation horizontal, presentation vertical; relation vertical, presentation vertical; relation vertical, presentation horizontal).

Task description

As in Experiment 8, except that all the subjects worked with 3 item arrays instead of 4, thus only three different colours were used throughout (red, green and yellow). This number of items used resulted in two possible premise orderings : AB BC and BC

AB. All the trials used ordering BC AB, so that some manipulation of premise item orderings would be required in order to correctly build the array. Three trials were completed by each subject*.

The premises were presented as in Experiment 8, again depending on whether horizontal or vertical task premise presentation was applicable.

Materials

As in Experiment 8, but drawings of 3 item arrays were used (either of women viewed 'side-on' in a queue, or acrobats standing on each others shoulders). Copies of the task materials (but with 4 items) are given in Appendix C.

Procedure

The WISC-R digit span subtest was administered to the participants.

As in Experiment 8 except that the children worked through three trials, rather than six, after they had been showed the worked example. Also, all the references to four women/acrobats were changed to three.

As previously described, the premises were presented in a horizontal array for half of the subjects working with each of the spatial relationships, and in a vertical array for the other half.

---

* This was because an informal pilot study in the same nursery school had indicated that a greater number of trials resulted in loss of interest.

8.3.2 Results

Note Each subject completed three experimental trials in total. However, a large number of the subjects in the previous studies erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier ones the first trial from each subject has been discounted.

WISC scores

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (relation horizontal, presentation vertical)	6.3
Group 2 (relation vertical, presentation vertical)	6.4
Group 3 (relation vertical, presentation horizontal)	6.2
Group 4 (relation horizontal, presentation horizontal)	6.3

Number of correct answers

Table 8.7 shows the mean number of correct answers for each experimental group.

Table 8.7: Mean number of correct trials (max=2)

Standard deviations are shown in parentheses

		Presentation	
		Horizontal	**Vertical**
Relationship	**Horizontal**	1.2 (0.42)	1.7 (0.48)
	Vertical	1.7 (0.48)	1.1 (0.32)

An ANOVA [2 (relationship) x 2 (presentation), both between subjects factors] was carried out on the above data. This revealed a significant interaction between the two factors ($F_{[1, 36]}=16.254$, $p<0.01$ - see Fig. 1), with no main effects* .

Fig. 8.20 below shows the interaction between relationship and premise presentation for the mean number of correct answers.

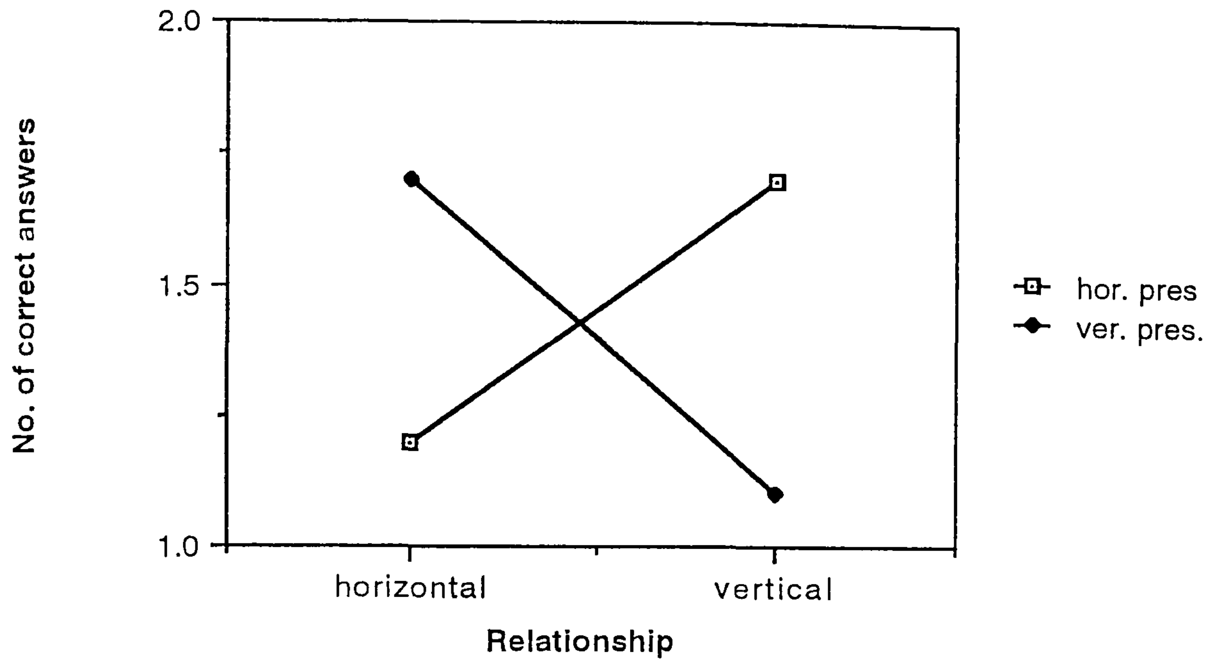
---

* The data is not normally distributed. This is because all the subjects scored either 1 or 2 correct answers. In view of this, analysis using the equivalent non parametric test (Fishers exact), was also carried out. This showed that the distribution of correct answers between the four experimental groups was significantly different from chance ($p<0.01$) The frequencies of correct answers for each of the conditions are shown in the table below. There were 10 subjects in each condition and each subject could score a maximum of 2. Thus the maximum number of correct answers for each condition is 20.

<u>Frequencies of correct trials (max=20)</u>

		Presentation	
		Horizontal	**Vertical**
Relationship	**Horizontal**	12	17
	Vertical	17	11

Fig. 8.20: Relationship x presentation interaction (number of correct answers)



Further analysis of the data revealed the following significant simple main effects:

- the subjects working with the horizontal relation and vertical premise presentation scored more correct answers than those working with the horizontal relation and horizontal premise presentation ($F_{[1, 36]}=6.716$, $p<0.05$).

- the subjects working with the vertical relation and horizontal premise presentation scored more correct answers than those working with the vertical relation and vertical premise presentation ($F_{[1, 36]}=9.672$, $p<0.01$).

- the subjects working with the horizontal premise presentation and the vertical relationship scored more correct answers than those working with the horizontal premise presentation and the horizontal relationship ($F_{[1, 36]}=6.716$, $p<0.05$).

- the subjects working with the vertical premise presentation and the horizontal relationship scored more correct answers than those working with the vertical premise presentation and the vertical relationship ($F_{[1, 36]}=9.672$, $p<0.01$).

Thus, for all comparisons, incongruent task and premise presentations were significantly better than congruent ones.

Time taken to complete trials (correct answers only)

Table 8.8 shows the mean time taken (in seconds) for each experimental group

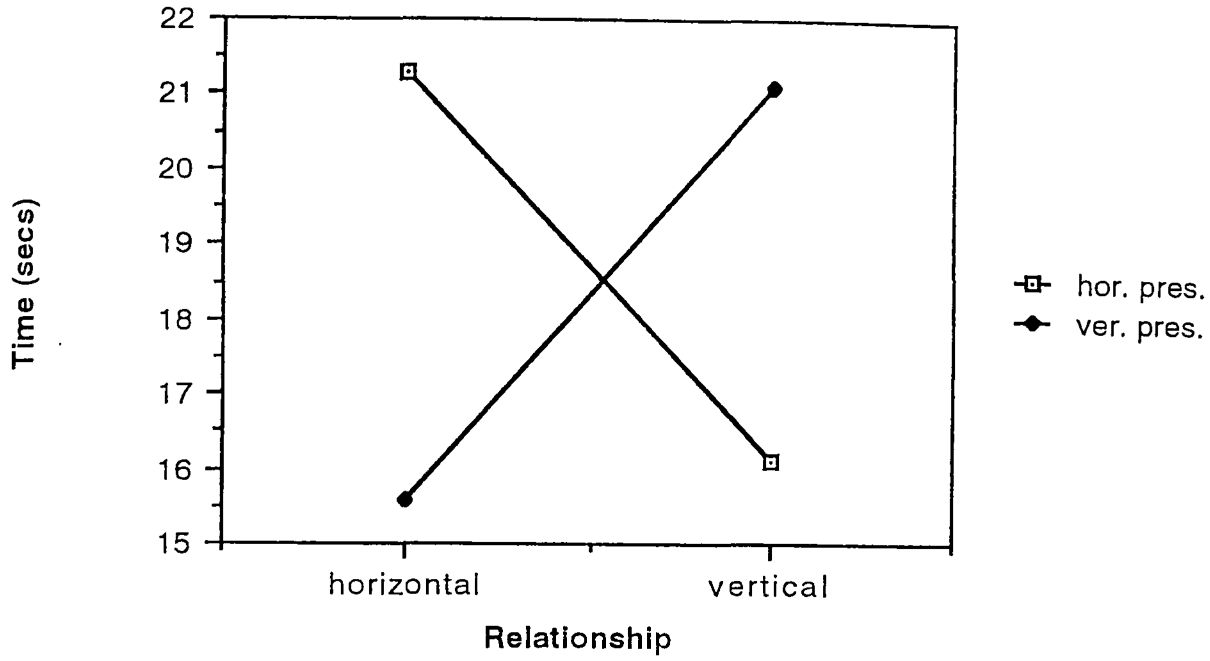
Table 8.8: Mean time taken to complete trials (correct answers only)

Standard deviations are shown in parentheses

		Presentation	
		Horizontal	**Vertical**
Relationship	**Horizontal**	21.29 (2.18)	15.58 (0.92)
	Vertical	16.12 (1.28)	21.13 (1.38)

An ANOVA [2 (relationship) x 2 (presentation), both between subjects factors] was carried out on the above data. Again, this revealed a significant interaction between the two factors ($F_{[1, 36]}=125.253$, $p<0.01$ - see Fig. 8.21), with no main effects.

Fig. 8.21: Relationship x presentation interaction (time taken to correctly solve tasks)



Further analysis of the data revealed the following significant simple main effects:

- the subjects working with the horizontal relation and vertical premise presentation produced correct solutions more quickly than those working with the horizontal relation and horizontal premise presentation ($F_{[1, 36]}=71.140$, $p<0.01$).

- the subjects working with the vertical relation and horizontal premise presentation produced correct solutions more quickly than those working with the vertical relation and vertical premise presentation ($F_{[1, 36]}=54.665$, $p<0.01$).

- the subjects working with the horizontal premise presentation and the vertical relationship produced correct solutions more quickly than those working with the horizontal premise presentation and the horizontal relationship ($F_{[1,36]}=58.206$, $p<0.01$).

- the subjects working with the vertical premise presentation and the horizontal relationship produced correct solutions more quickly than those working with the vertical premise presentation and the vertical relationship ($F_{[1, 36]}=67.208$, $p<0.01$).

Again, for all comparisons, incongruent task and premise presentations were significantly better than congruent ones.

8.3.3 Discussion

This study replicates the results from Experiment 7, but with a younger age group. For both measures, and for all comparisons, incongruent task and premise presentation relationships were significantly better than congruent ones. Thus, although the task was made easier so that the 5 year olds were not performing at 'floor', the 'congruence constraint' was still significantly inhibiting performance.

We now have evidence that the 'congruence constraint' operates with simpler tasks and a younger age group. It is possible that the removal of this constraint might raise the level of performance of the 5 year olds to that of the 7 year olds. This would question the assumption that working memory limitations are having a significant effect on the integration of premises into a single ordered array. In order to test for this we would need to compare the performance of 5 year olds working with 4 item arrays and incongruent relationships to that of 7 year olds.

8.4 GENERAL SUMMARY

The three experiments reported in this chapter have demonstrated the existence of a task constraint which is affecting the construction of systematic relational representations in children. These representations were constructed by integrating separate items of relational information (known as premises) When the dimensions of the task relationship and premise information were congruent, performance was inhibited. This has been termed the 'congruence constraint'. Examples of congruent and incongruent tasks are given below.

Fig. 8.22: Example of a congruent task (both the task relationship and the premise presentation are horizontal.

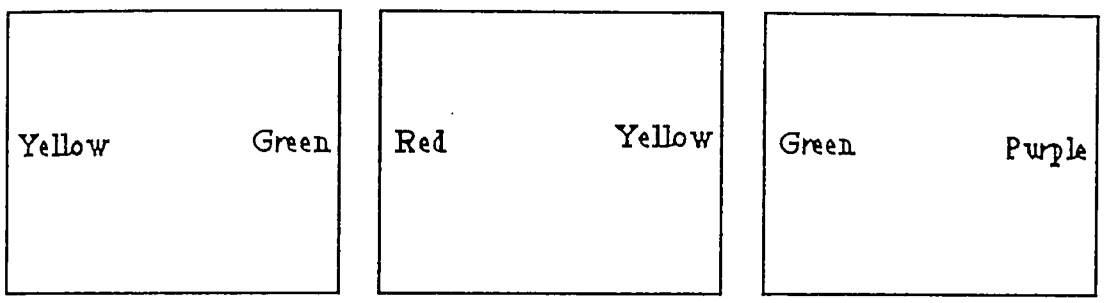
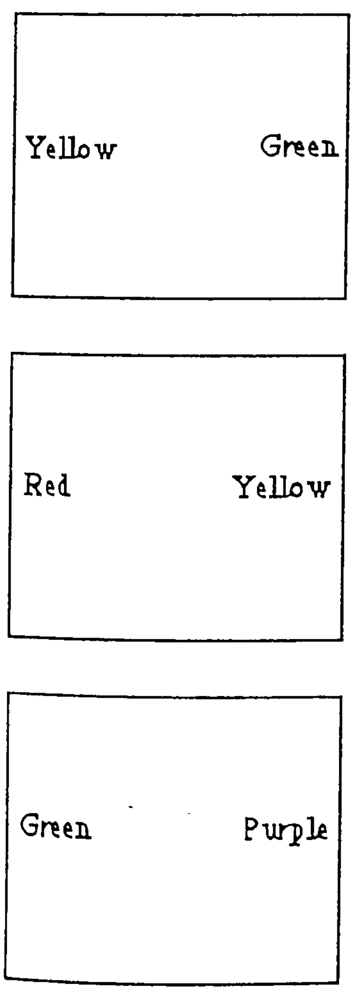


Fig. 8.23: Example of an incongruent task (the task relationship is horizontal, whereas the premise presentation is vertical.



Studies have shown that the performance of both 5 and 7 year old children using drawings of items was affected, as well as that of 7 year olds using familiar toys. It is likely that this constraint operates by withdrawing or reducing the significance of the gap between premise pairs. As a result of this, the children are encouraged to apply an inappropriate left to right rule, which often results in incorrect ordering and so a correct structural task representation cannot be built.

It is relevant here to consider the results reported by Pears and Bryant (1990) and also those obtained in Experiment 2 (which was a replication of Pears and Bryant). Both of these two studies used a vertical task relationship, with an incongruent premise presentation relationship, in that the children (5 year olds in both cases) were required to order coloured bricks in a tower (a vertical relationship) whilst the premises (in the form of pairs of bricks) were presented to the child horizontally (see Figs. 8.24 and 6.25 below)

Fig. 8.24: Example of premise presentation - Experiment 2 and Pears and Bryant (1990)

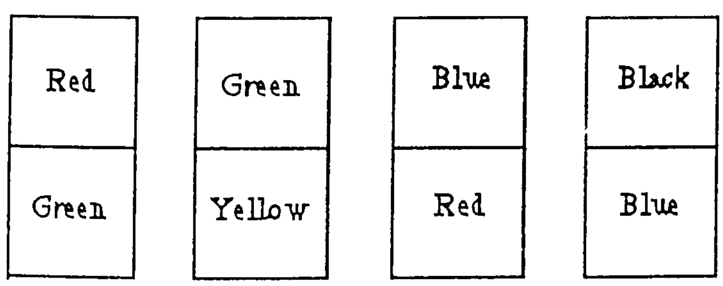
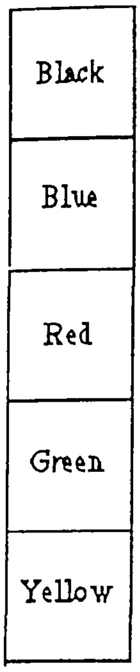


Fig. 8.25: Example of fully ordered array for the above premises- Experiment 2 and Pears and Bryant (1990)



Although no direct comparisons between congruent and incongruent conditions were made, it did seem that performance in both of these studies was better than that shown in Experiment 1, where the same aged children worked with a similar task, except that they were ordering in a horizontal dimension (family members were placed in a queue). Premises were also presented horizontally in Experiment 1, resulting in congruent task and premise presentation relationships. An example of this is given in Figs. 8.26 and 8.27 below.

Fig. 8.26: Example of premise presentation - Experiment 1

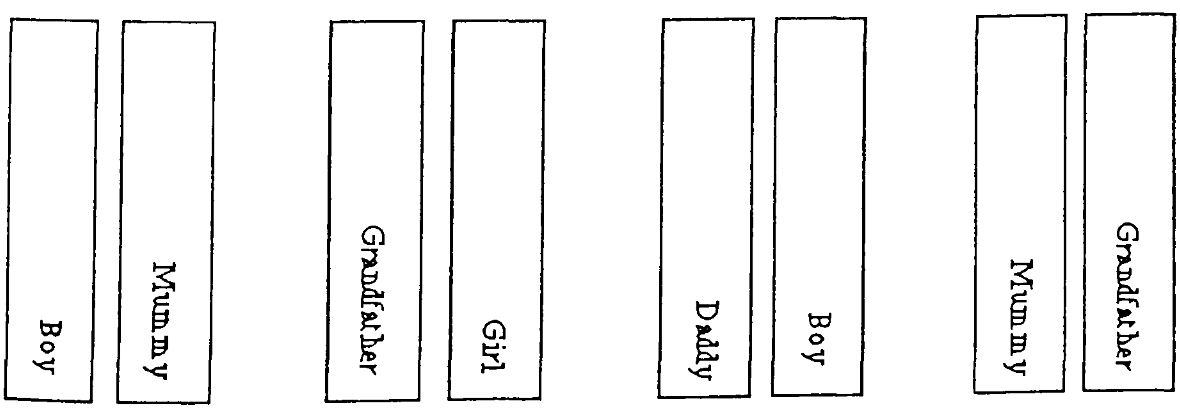
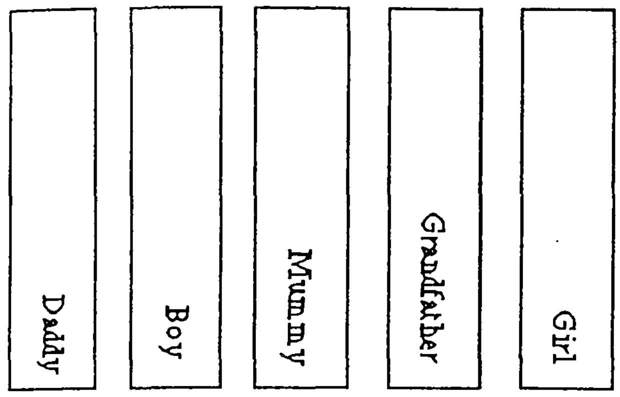


Fig. 8.27: Example of fully ordered array for the above premises- Experiment 1



There are two possible reasons for the improved performance demonstrated in these two studies; the dimension of the actual relationship (vertical rather than horizontal) and the incongruence of task and premise presentation relationships. However, none of the three experiments in the current chapter found a main effect of task relationship when vertical and horizontal conditions were compared. Both real objects (toys) and drawings were used, and the performance of 5 and 7 year olds was investigated. It is suggested therefore, that the performance found by Pears and Bryant and also reported in Experiment 2, was due to the incongruence of task and premise presentation relationships, rather than because a vertical relationship was used. This is interesting, particularly if we consider work reported by Handel, DeSoto and London (1968) which claimed that people find the vertical representation of relationships easier than the horizontal. This is an important point, and will be returned to in Chapter 10.

It has been argued in earlier chapters that children solve series problems (the tasks used in this thesis to study structural representations) by performing an analogical mapping with orderings which they have abstracted from everyday life (Halford, 1992). This abstraction has been termed a generalised ordering schema, after Cheng and Holyoak, 1985.

We have evidence from the studies reported in this chapter that successful mapping is prevented in 'near' analogies due to the 'congruence constraint' described above. A 'near' analogy is one in which surface features are common to both base and target structures (Gentner, 1989). In the case of series problems, near analogies have been defined as those in which both structures share the same actual relation. Thus, in the problems studied here, the children have been required to reason about actual spatial relationships, both in the horizontal dimension ('behind') and in the vertical one ('on top of' and 'underneath'). These are the same relations as the one which Halford has used in the abstracted ordering schema (the base structure).

It is now necessary to see whether the 'congruence constraint' also applies to 'far' analogies, i.e. those which have no surface features in common with the generalised schema. Chapter 9 will address this issue.

CHAPTER 9 : STRUCTURAL CONSTRAINTS AND 'FAR' ANALOGICAL MAPPING

Introduction

The primary goal of this thesis is to explore the role of structural task representations in the development of analogical reasoning. This has been investigated using series problems, as Halford (1992) has claimed that these types of problems are solved by analogy to an 'internalised ordering schema' which has been abstracted from interaction with orderings which occur during everyday play etc. Because the solving of series problems is heavily dependent on the construction of an internal spatial array (Trabasso *et al*, 1975; Riley, 1976), they are well suited to an exploration of the way in which children actually construct a task representation where the relevant objects and their relationships to each other are depicted.

Experiments 2 to 10 have studied the effects of task constraints on series problems which use a spatial task relationship (either vertical or horizontal). These have been termed 'near' analogies (after Gentner, 1989) because they have surface features in common with the ordering schema put forward by Halford. Fig. 9.1 below reproduces this schema. It can be seen from this that the actual relationship used is a spatial one. This constitutes the common surface feature shared with spatial series problems.

Fig. 9.1: Series problem mapped into an ordering schema (Halford, 1992)

		above		above	
base schema	**TOP**	________	**MIDDLE**	________	**BOTTOM**
target schema	**JOHN**	________	**PETER**	________	**TOM**
		taller than		taller than	

From Halford (1992)

The studies referred to above have demonstrated that there are at least two task constraints operating on primary aged children's ability to combine separate spatial relations into a single integrated representation. These have been identified as follows:

1. An effect of information ordering (Chapters 5 and 6).

2. An effect of congruity between task and information presentation relations (Chapter 8).

These constraints are therefore preventing successful mapping from the 'base' domain (the ordering schema) to the 'target' domain (the actual problem).

Halford claimed that children solve all types of series problems, including those using more complex, abstract relationships, by making analogies to the 'ordering schema'. Abstract relationships are 'far' analogies (Gentner, 1989) as they have only the relational structure in common with the 'ordering schema'. It is now necessary to ascertain whether the effects listed above are present when children solve abstract series problems.

Experiment 1 demonstrated very little success when young children were asked to solve abstract series problems. However, a similar lack of success was demonstrated in spatial series problems, that is, those where mapping relations over from spatial to abstract domains was not required. These results were therefore suggesting that there were no differences in 7 year old children's performance between spatial ('near') and abstract ('far') problems. This evidence is interesting when we consider Gentner's (1989) account of the 'relational shift' in analogical reasoning. She claims that there is a developmental shift from reliance on surface similarity to the use of relational structure. We might therefore expect that young children would find spatial series problems easier than abstract ones, as the former have surface features (the actual task relation) in common with the ordering schema.

Later studies in this thesis have shown that performance in concrete (spatial) problems was facilitated by consideration of the factors listed above. If the initial results from Experiment 1 are correct and there is no difference between spatial and abstract series problems, then we would expect both domains to result in similar patterns of performance when congruent and incongruent task and premise presentation relationships are compared. Experiment 8 reported these results for spatial problems. On the other hand, if Gentner is correct and young children find 'near' analogies easier, then the facilitative effect observed for incongruent conditions when reasoning with spatial series problems might not be replicated for abstract ones. The initial study reported in this chapter will therefore investigate the 'congruence constraint' using abstract series problems.

This will enable us to ascertain whether the previously observed facilitation effects are generalisable to non-spatial domains and also to observe whether children can use 'spatial' bases to reason about abstract relationships. Comparison with the data from Experiment 8 will allow us to conclude whether or not children find abstract series problems more difficult than spatial ones.

9.1 EXPERIMENT 11 - THE EFFECT OF CONGRUENT AND INCONGRUENT BASE AND PREMISE PRESENTATION RELATIONSHIPS WHEN REASONING WITH 'FAR' ANALOGIES.

Rationale

Chapter 8 indicated that, in some circumstances, primary aged children were more successful in solving spatial series problems when the task and premise presentation relationships were incongruent, that is, they both employed either a vertical or horizontal dimension. This effect, termed the 'congruence constraint' has been demonstrated to significantly affect 7 year old children's performance when integrating spatial relational information from both drawings and real objects, and also 5 year old's performance when integrating this information from drawings.

The current study was designed to ascertain whether this effect is also present when children solve abstract series problems. It has previously been demonstrated that children use spatially ordered arrays to solve evaluative 'non-spatial' series problems e.g. Riley, 1975; Experiment 1, this thesis). In view of this, and as in the previous experiments, the subjects will be encouraged to order the abstract relational information by using a spatial array. However, we have no clear evidence which satisfactorily discriminates between vertical and horizontal arrays. The studies reported in Chapter 8 found no differences in performance when comparing the direction of ordering required by the task. However, work by DeSoto, London and Handel (1965), discussed in

Chapters 3 and 4, suggested that some abstract evaluative relations are spatially tied to a vertical cognitive array. They claimed that these are relationships which have an everyday linguistic linkage to words which describe vertical positions. The example they used was 'better-worse', stating that the dictionary definition of the word 'high' shows definite vertical ties to this relationship.

The significance of their proposition for this thesis will be fully investigated in Experiment 12. For the purposes of this experiment, it was decided to compare both vertical and horizontal spatial arrays when reasoning about a relation which appeared to have no obvious tie to a spatial dimension. This will enable us to gain some insights concerning the 'congruence constraint' in abstract series problems without any possible confound caused by the implied 'vertical spatial tie' for some evaluative relations.

9.1.1 Method

Design

A two factorial between subject design was used. The two factors were as follows:

1. Dimension of base relationship (vertical or horizontal).

2. Ordering of premise presentation (vertical or horizontal).

Thus there were four experimental groups.

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task.

Participants

The participants in this study were 40 mixed ability 7 year olds (mean age 6 years 11 months, range 6 years 5 months to 7 years 3 months) from state first and primary schools, with predominantly middle class catchment areas. The subjects were from the same schools used in Experiment 3, and were randomly assigned to one of the four experimental groups.

Task descriptions and materials

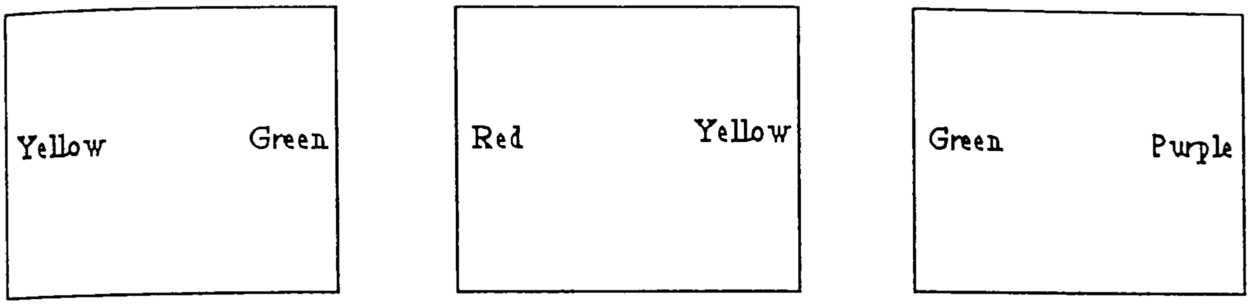
These were the same as reported in Experiment 8, except that the subjects were told to use the concrete arrays to reason about the neutral abstract relationship 'louder than' (see Procedure section). All the conditions used 'left to right' premise ordering (BC CD AB and BC AB CD). Further details of the task descriptions are given below.

Horizontal base relationship

The four-item array of women viewed 'side-on', used for Experiments 3, 5, 6 and 8 formed the horizontal base array. Copies of the task materials are given in Appendix C.

The subjects in the horizontal presentation group were presented with the premises laid out as follows : BC CD AB. An example of this is given in Fig. 9.2 below.

Fig. 9.2: Example of horizontal spatial relationship and horizontal presentation.



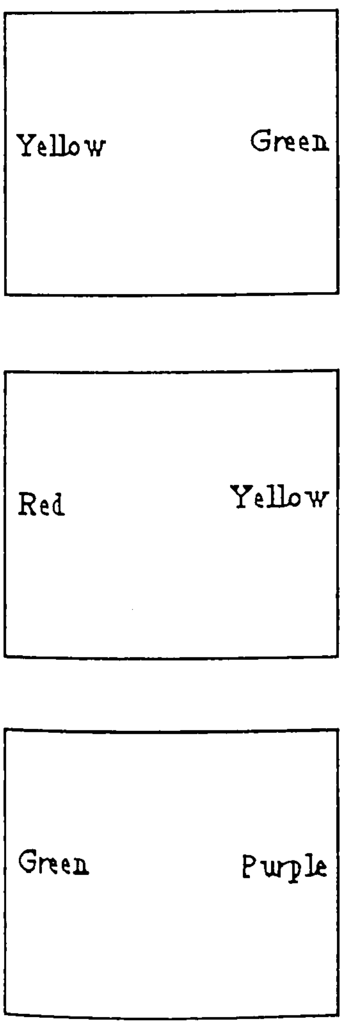
The subjects in the vertical presentation group were presented with the premises laid out as follows : BC

CD

AB

An example of this is given in Fig. 9.3 below.

Fig. 9.3: Example of horizontal spatial relationship and vertical presentation.



Vertical spatial relationship

The four-item array of acrobats in a 'human tower,' used for Experiment 8 formed the vertical base array. Copies of the task materials are given in Appendix F.

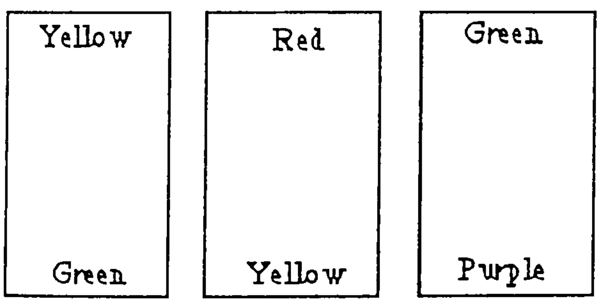
The subjects in the horizontal presentation group were presented with the premises laid out as follows :

B C A

C D B

An example of this is given in Fig. 9.4 below.

Fig. 9.4: Example of vertical spatial relationship and horizontal presentation.



The subjects in the vertical presentation group were presented with the premises laid out as follows:

B

C

C

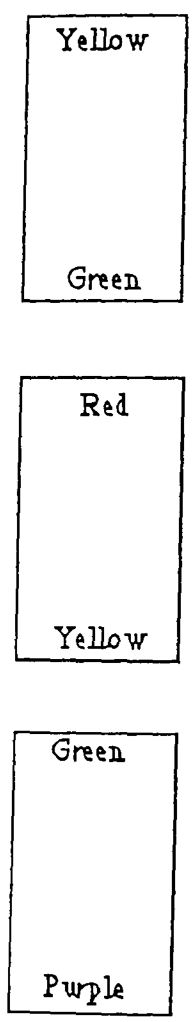
D

A

B

An example of this is given in Fig. 9.5 below.

Fig. 9.5: Example of vertical spatial relationship and vertical presentation.



Procedure

This was as similar as possible to Experiment 8. However, because a large part of the instructions given to the subjects related to the mapping of the actual task relationship into the concrete base relationship, the procedure is described in full below.

Subjects were tested individually in a quiet room. The experimenter first explained the task, using a complete photograph and serial ordering of premises as a worked example. The children then worked through the six examples, using copies of the individual 4 items to build a concrete array. Final ordering, time taken and order of item placement were recorded by the experimenter.

Wording of Instructions

"I've done some drawings of four people standing in a queue/children standing on each others shoulders...Look, like this one... Can you see that they're all wearing different coloured jumpers/tee-shirts? You see that the person wearing the <yellow> jumper/tee-shirt is first/on top, then the <red> and then the <green> and last of all/at the bottom the <purple> one. I also made some copies....they were exactly the same. I cut one of the copies down/across the middle, so that the first/top two people/children in the queue/tower were in the first/top half and the last/bottom two people/children were in the second/bottom half.......look...like this.....I then got another copy and cut off the first/top and last/bottom people/children in the queue/tower and threw them away, so that I only had the middle two people/children left in this part...look...like this.....The other drawings have got the same people/children in, but they're in a different order in each one. I've got some puzzles for you to do, where you have to work out who has the loudest voice, and then the next loudest, and then the next loudest, and then finally the person with the least loudest voice. I want you to use these part drawings as clues to help you do the puzzle because you won't be able to look at the whole drawing until after you've finished.....look, I'll show you how........<Use serial ordering> Here are four separate drawings of the four people/children. We can use these to build up our own queue/tower from the clues. Look at these two people/children here. The drawing shows us that the person wearing the green jumper/tee-shirt has a louder voice than the person wearing the red jumper/tee-shirt...OK? So you use your little drawings to start to build the queue/tower.....now look at this next one.. Who's got a louder voice than the person/child with the green jumper/tee-shirt?....OK, so now you use your drawings again.....Now use them again when you've worked out what this last drawing is showing you......Right, now you've put all of the people/children in the order which these three part drawings told you about and so you've finished this puzzle. Well done.....There are six more for you to do. They all use the same people/children but the order of the people/children will be different and the order of the drawings which I give you might make the puzzles a little bit harder."

As previously described, the premises were presented horizontally for half of the subjects working with each of the spatial relationships, and vertically for the other half.

WISC-R digit recall scores were also taken from each subject.

9.1.2 Results

Note Each subject completed six experimental trials in total. However, a large number of the subjects in the previous studies erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier ones the first trial from each subject has been discounted.

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (base relation vertical, presentation vertical)	8.7
Group 2 (base relation vertical, presentation horizontal)	8.5

Group 3 (base relation horizontal, presentation vertical) 8.8

Group 4 (base relation horizontal, presentation horizontal) 8.5

The average performance on the digit span test for ages between 6 years 8 months and 7 years is a score of 8 (WISC-R Manual, 1974).

Number of correct answers

Table 9.1 shows the mean number of correct answers for each experimental group.

Table 9.1: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

		Presentation	
		Horizontal	**Vertical**
Base Relationship	**Horizontal**	2.8 (0.62)	2.6 (0.84)
	Vertical	4.1 (0.57)	2.9 (0.88)

An ANOVA [2 (base relationship) x 2 (presentation), both between subjects factors] was carried out on the above data. This revealed a significant interaction between the two factors ($F_{[1, 36]}=4.128$, $p<0.05$ - see Fig. 9.6). The two main effects were also significant:

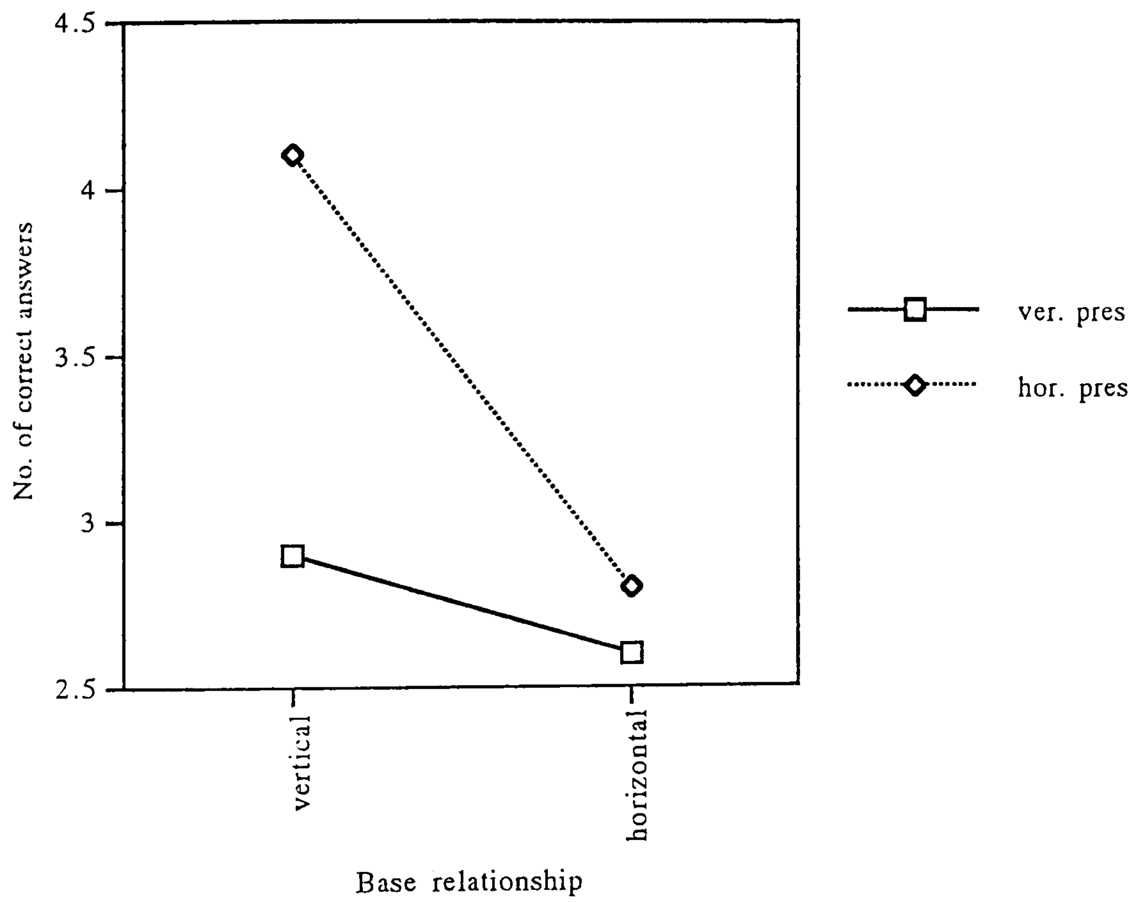
1. Base relationship.

Those children working with the vertical base relationship produced a greater number of correct trials overall than those working with the horizontal base relationship ($F_{[1, 36]}=10.569$, $p<0.01$).

2. Dimension of premise presentation.

Those children working with the horizontal premise presentation produced a greater number of correct trials overall than those working with the vertical premise presentation ($F_{[1, 36]}=8.092$, $p<0.01$).

Fig. 9.6: Base relationship x premise presentation interaction



Further analysis of the above data revealed the following significant simple main effects:

- the subjects working with the vertical base relationship and the horizontal premise presentation produced a greater number of correct solutions than those working with the vertical base relationship and the vertical premise presentation ($F_{[1, 36]}=11.890$, $p<0.01$).

- the subjects working with the horizontal premise presentation and the vertical base relationship produced a greater number of correct solutions than those working with the horizontal premise presentation and the horizontal base relationship ($F_{[1, 36]}=13.954$, $p<0.01$).

Time taken to complete trials (correct answers only)

Table 9.2 shows the mean time taken (in seconds) to correctly complete the trials for each experimental group.

Table 9.2: Mean time taken to complete trials (correct answers only

Standard deviations are shown in parentheses

		Presentation	
		Horizontal	**Vertical**
Base Relationship	**Horizontal**	17.57 (1.47)	17.26 (1.69)
	Vertical	12.73 (0.81)	16.96 (1.28)

An ANOVA [2 (base relationship) x 2 (presentation), both between subjects factors] was carried out on the above data. This revealed a significant interaction between the two factors ($F_{[1, 36]}=28.135$, $p<0.01$ - see Fig. 9.7). The two main effects were also significant:

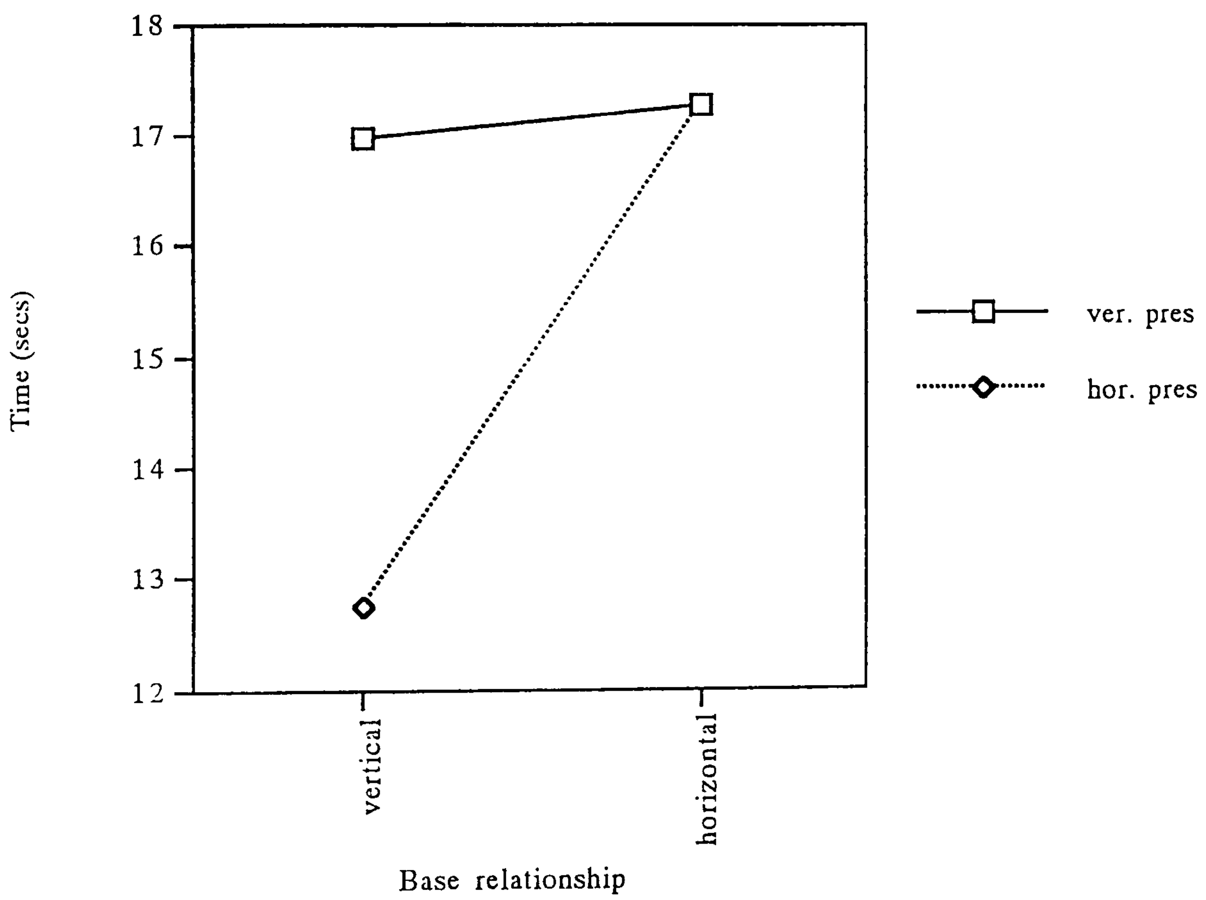
1. Base relationship.

Those children working with the vertical relationship produced correct solutions more quickly overall than those working with the horizontal relationship ($F_{[1,36]}=35.834$, $p<0.01$).

2. Premise presentation.

Those children working with the horizontal premise presentation produced correct solutions more quickly overall than those working with the vertical premise presentation ($F_{[1, 36]}$=20.980, p<0.01).

Fig 7.7: Base relationship x premise presentation interaction (time taken to correctly solve tasks)



Further analysis of the above data revealed the following significant simple main effects:

- the subjects working with the vertical base relationship and the horizontal premise presentation produced correct solutions more quickly than those working with the vertical base relationship and the vertical premise presentation ($F_{[1, 36]}$=48.853, p<0.01).

- the subjects working with the horizontal premise presentation and the vertical base relationship produced correct solutions more quickly than those working with the horizontal premise presentation and the horizontal base relationship ($F_{[1, 36]}=63.737$, $p<0.01$).

Comparison with data from Experiment 8 (incongruent conditions)

Because the subjects in Experiment 8 and this study were taken from the same schools, and the same type of materials and procedure were used throughout, it was decided to combine the relevant data from the two studies in a direct statistical comparison. If we consider just the two incongruent conditions (i.e. using both horizontal and vertical task relationships) the only difference between the studies is the task relationship. The subjects in Experiment 8 used the ordering of the array to reason about a concrete dimensional relationship ('in front' or 'on top of') whereas those in this study used the same array to reason about an abstract relationship ('louder than').

Number of correct answers

Table 9.3: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

		Task relationship	
		Concrete	**Abstract**
Base relationship	**Horizontal**	4.5 (0.53)	2.6 (0.84)
	Vertical	4.6 (0.52)	4.1 (0.57)

An ANOVA [2 (base relationship) x 2 (task relationship), both between subjects factors] was carried out on the above data. This revealed a significant interaction between the two factors ($F_{[1, 36]}=12.423$, $p<0.01$ - see Fig. 9.8). The two main effects were also significant:

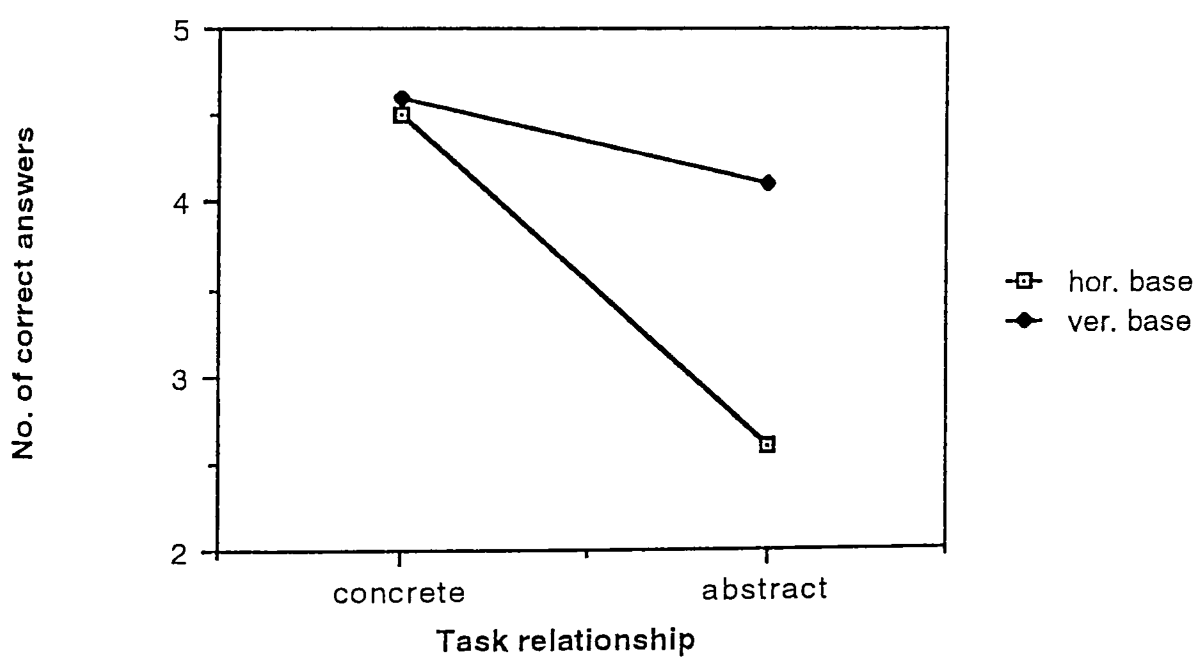
1. Base relationship.

Those children working with the vertical relationship produced a greater number of correct trials overall than those working with the horizontal relationship ($F_{[1,36]}=16.225$, $p<0.01$).

2. Task relationship.

Those children working with the concrete relationship produced a greater number of correct trials overall than those working with the abstract relationship ($F_{[1, 36]}=36.507$, $p<0.01$).

Fig 7.8: Base relationship x task presentation interaction (number of correct answers)



Further analysis of the above data revealed the following significant simple main effects:

- the subjects working with the horizontal base relationship and the concrete task relationship produced more correct solutions than those working with the horizontal base relationship and the abstract task relationship ($F_{[1, 36]}=45.761$, $p<0.01$).

- the subjects working with the abstract task relationship and the vertical base relationship produced more correct solutions than those working with the abstract task relationship and the horizontal base relationship ($F_{[1, 36]}=28.521$, $p<0.01$).

Time taken to complete trials (correct answers only)

Table 9.4 shows the mean time taken (in seconds) for each experimental group.

Table 9.4: Mean time taken to complete trials (correct answers only)

Standard deviations are shown in parentheses

		Task relationship	
		Concrete	**Abstract**
Base relationship	**Horizontal**	12.59 (1.50)	17.26 (1.7)
	Vertical	12.06 (1.44)	12.73 (0.81)

An ANOVA [2 (base relationship) x 2 (task relationship), both between subjects factors] was carried out on the above data. Again, this revealed a significant interaction between the two factors ($F_{[1, 36]}=20.426$, $p<0.01$ - see Fig. 9.9). Also, the main effects were again significant:

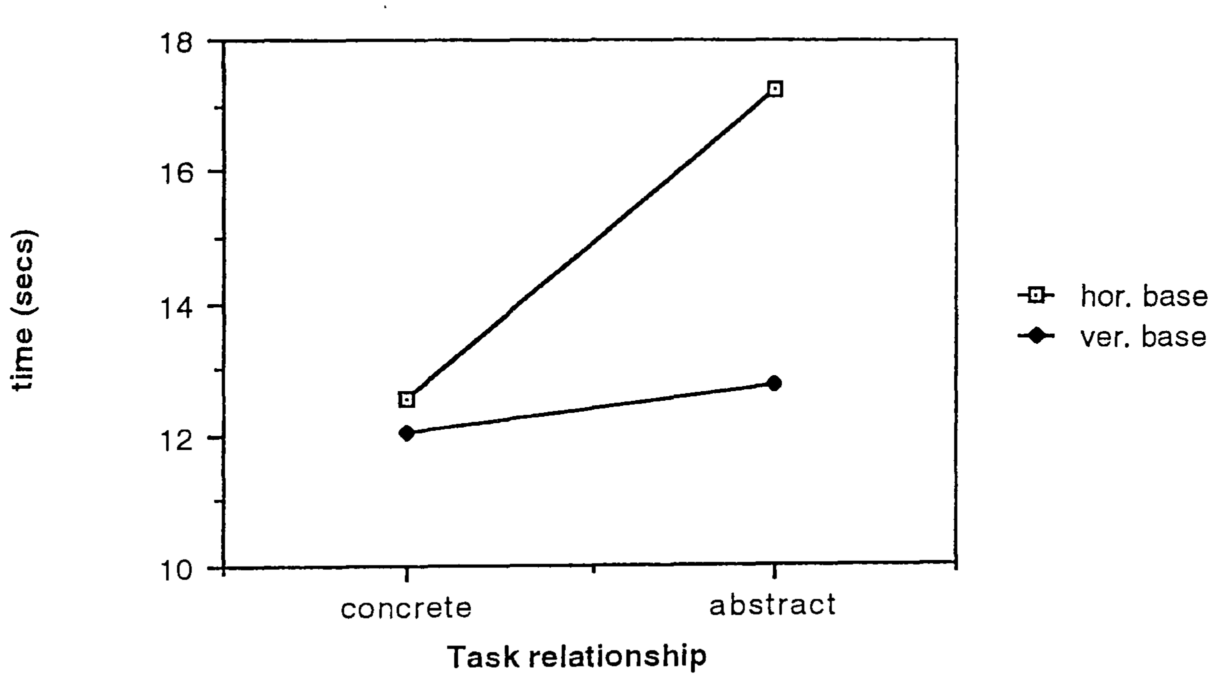
1. Base relationship.

Those children working with the vertical relationship produced correct solutions more quickly overall than those working with the horizontal relationship ($F_{[1,36]}=32.499$, $p<0.01$).

2. Task relationship.

Those children working with the concrete relationship produced correct solutions more quickly overall than those working with the abstract relationship ($F_{[1, 36]}$=36.314, p<0.01).

Fig 7.9: Base relationship x task presentation interaction (time taken to correctly solve tasks)



Further analysis of the above data revealed the following significant simple main effects:

- the subjects working with the horizontal base relationship and the concrete task relationship produced correct solutions more quickly than those working with the horizontal base relationship and the abstract task relationship ($F_{[1,36]}$=55.605, p<0.01).

- the subjects working with the abstract task relationship and the vertical base relationship produced correct solutions more quickly than those working with the abstract task relationship and the horizontal base relationship ($F_{[1,36]}=52.228$, $p<0.01$).

9.1.3 Discussion

Both dependent variables (number of correct trials and time taken to correctly order the array) show very similar patterns of results and so the following discussion refers to both measurements.

The aim of this experiment was to ascertain whether the 'congruence constraint' identified for spatial series problems ('near' analogies) also has an effect on performance when children are ordering arrays concerning abstract relationships ('far' analogies). It can be seen from the interaction that incongruent task and premise presentation relationships did have a facilitative effect on performance, but only for the vertical base and horizontal premise presentation relationships. There was no difference between the congruent (horizontal premise presentation) and incongruent (vertical premise presentation) conditions for the horizontal base relationships. Similarly, there was no difference between congruent (vertical base relationship) and incongruent (horizontal base relationship) conditions for the vertical premise presentation. Thus it appears that, when mapping 'far' analogies, the facilitative effect demonstrated for incongruent task and premise presentation relationships is only present for the vertical base.

It was decided to directly compare performance using spatial and abstract tasks, to see whether this claim could be substantiated. The current study and Experiment 8 both employed the same design, but whereas the former used an abstract task, Experiment 8 used a spatial one. Both incongruent conditions (horizontal base relationship, vertical premise presentation and vertical base relationship, horizontal premise presentation) from the two studies were therefore compared in one analysis. This showed both an interaction and main effects of base and task relationships. It was found that, whilst there was no effect of the task dimension for concrete problems, inhibition of performance was caused by the use of a horizontal base when reasoning with the abstract task. Also, the performance of subjects in this condition was significantly worse than that of the subjects who used the vertical base for reasoning about the abstract task. These results are therefore consistent with those obtained by the analysis carried out solely on the current study, referred to above. Ordering arrays using a horizontal base and an abstract task relationship means that performance cannot be facilitated by the use of incongruent task and premise presentation relationships. On the other hand, the use of a vertical base, or a concrete relationship, or both, does result in such facilitation.

It would seem therefore, that we have some evidence to support the claim that incongruent task and premise presentation relationships boost performance for 'near' analogies, irrespective of the dimension of the task relationship. However, when this relationship is used as a base for mapping 'far' analogies, facilitation of performance only occurs when a vertical base is used. It could therefore be that abstract ('far') analogies are, in some circumstances at least, more difficult than ('near') analogies. Thus, even when the task is made easier by the use of incongruent base and premise presentation relationships, children are unable to take advantage of this unless a vertical base relationship is used.

At first sight, this appears inconsistent with the results from Experiment 1, which showed that 7 year old children were equally successful using either concrete or abstract relationships in a similar type of task. However, it must be remembered that the subjects in Experiment 1 were using a random selection of all levels of difficulty of premise orderings with a five-item array, and were also working with congruent task and premise presentation relationships. It could well be that these difficulties were swamping any differences which we might have found between 'near' (spatial) and 'far' (abstract) mappings.

We must also consider the actual relationships chosen for these studies. Experiment 1 used 'happier than', whereas this study used 'louder than'. The latter relation was originally chosen because it was thought to be neutral. However, it could be that this is in fact easier to map onto a vertical relationship than a horizontal one. Although the relationship actually compared volume ('louder than'), it could be that the children were more successful when using a vertical array as they have imported terminology from the 'high-low' continuum which is used to describe differences in pitch. This is consistent with work reported by Handel, DeSoto and London (1968), which suggested that, for adults, some relations are 'spatially tied' to a vertical dimension. Their subjects (undergraduate students) were first asked to give spatial assignments to relation words. Each subject was given sheets containing a statement, for example, Tom is better than Bob, and also a diagram, with vertical and horizontal lines crossing at 90 degrees. A box was situated at both ends of either line. The subjects were asked to write the name of the statement items in whichever two of the four boxes seemed most appropriate. They found that the relationships 'better-worse', 'father-son' and 'more-less' were assigned to the vertical dimension for 71% to 78% of the time. Relations such as 'faster-slower' and 'farther-nearer' were assigned to the horizontal dimension for about 50% of the time, though there were also inconsistent assignments for 25% and 43% of the time respectively. Following this, they presented their subjects with three term series problems and found that relationships having similar patterns of spatial assignment also had similar patterns of premise combination difficulty (these combinations were formed by varying the position of opposite relations in each task - for example, 'better' and 'worse') They concluded from this that adults did use spatial arrays to order evaluative relationships. Furthermore, based on the results from the spatial assignment task, they also claimed that it was easier for adult subjects to represent relations vertically when ordering series problems. However, it must be remembered that this related to written series problems solved without the ordering of any external arrays.

Nonetheless, there is a suggestion that it might be easier for people to map relations vertically, rather than horizontally. DeSoto, London and Handel (1965), working with series problems with no explicit spatial relationship, suggest that this could be due to linguistic conventions, as there is often a tie to the vertical dimension for many evaluative relations. This is demonstrated by metaphorical phrases such as 'upper-crust', 'high status', 'top dog' etc. It could be that this preference for vertical ordering is also affecting the children in the current study, such that they can only map 'far' analogies onto vertical orderings.

However, it is also considered that a horizontal ordering could well be salient for 7 year old children. They are used to 'lining-up' in order in the classroom (a horizontal ordering) and also to 'lining-up' long strings of toy cars or animals during play. In a similar vein, Johnson (1987) when asserting that spatial structure is mapped onto mental structures, places equal emphasis on vertical and horizontal schemas. Indeed, it might be suggested that young children are more used to horizontal rather than vertical orderings as they have more opportunity for the former. These can be constructed on any table or floor, whereas the opportunity to order toys vertically requires surfaces at different heights. These may not always be available.

It seems that we are left with two possibilities. Either young children conform to the suggestion made by Handel *et al* (1968), and show a definite preference for vertical mappings, or they can show a preference for either dimension, depending on the actual task relation. If horizontal relations are salient for young children, then it would seem that there are many relations for which these would be more appropriate. Comparisons such as 'nearer to' or 'faster than' or perhaps 'wider than' could well be viewed as more similar to horizontal orderings which, it is argued, children encounter everyday.

It could also be that there are some relationships which are not semantically tied to either dimension. However, 'louder than' does not appear to be one of them! If this is the case, then we might expect to find either that 7 year old children can use either dimension to map such relations, or that they have difficulty in using any.

In view of the above, it was decided to investigate the effect of the dimension of the base relationship on success in mapping various different target relations.

9.2 EXPERIMENT 12 - THE EFFECT OF THE DIMENSIONALITY OF THE BASE STRUCTURE ON THE MAPPING OF 'FAR' RELATIONAL ANALOGIES.

Rationale

Experiment 11 (reported above) suggested that, for young children, some types of non-spatial series problems might be spatially tied to either vertical or horizontal orderings. Thus, when using a spatially ordered array as an analogical base for reasoning about an abstract series problem (a 'far' analogy), we might see differences in performance between vertical and horizontal arrays, depending on the abstract task relationship used. It was therefore necessary to identify some abstract relationships which could vary according to the spatial dimension to which they are tied.

The study discussed above (Handel *et al*, 1968) reported that, for adults, the relation 'better/worse' was tied to a vertical ordering for 78% of the time. They did not however claim that any evaluative relations were tied to horizontal orderings. Nonetheless, their results showed that the relation most consistently assigned to a horizontal ordering was 'faster/slower' (48% of the time).

It was therefore decided to use these two relationships to investigate the claim that, for 7 year olds, evaluative relations are 'spatially tied' to whatever dimension of ordering they have closer semantic links to. This is contrary to Handel *et al's* position, which was that evaluative relations are tied to vertical orderings. For comparison purposes, it was also decided to identify a possible neutral relationship (one which is not spatially tied). This might enable us to see whether such relations are more comparable to situations where the base and target (task) relations are compatible (spatially tied) or to those situations where they are not compatible (because the target relation is spatially tied to the dimension opposite to that used for the base relationship). The research reported by Handel *et al* could not be drawn on here. The relation which had the most inconsistent spatial assignment was darker hair/lighter hair. This obviously could not be used, as the subjects will be required to order using drawings, where actual hair colour will be apparent. The relation 'older than' was chosen. This is because the women in the drawings used as task materials do not appear to differ in age, so visual cues cannot be used. It is also considered likely that this relation is not closely linked with either a vertical or horizontal base relationship.

This study will therefore test the hypothesis that the relationship 'better than' is tied to a vertical ordering for 7 year old children, whereas the relationship 'faster than' is tied to a horizontal one.* No definite predictions can be made about the relationship 'older than'.

9.2.1 Method

Design

A two factorial between subject design was used. The two factors were as follows:

1. Dimension of base relationship (vertical or horizontal).

2. Task (target) relationship (best, fastest or oldest).

* An informal pilot study attempted to carry out Handel et al's spatial assignment task (described on page 16) However, this was unsuccessful as the children did not understand what was required of them.

Thus there were six experimental groups.

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task.

Participants

The participants in this study were 60 mixed ability 7 year olds (mean age 6 years 10 months, range 6 years 5 months to 7 years 6 months) from two state primary schools with predominantly middle class catchment areas. The subjects were randomly assigned to one of the six experimental groups.

Task descriptions and materials

These were the same as reported in Experiment 11, except that the subjects were told to use the concrete arrays to reason about the abstract relationship 'better than', 'faster than' or 'older than' (see Procedure section). All premises were presented in the opposite dimension to that of the base relationship (i.e. they were always incongruent). Further details of the task descriptions are given below.

Horizontal base relationship

The four-item array of women viewed 'side-on', used for Experiments 3, 5, 6, 8 and 11 formed the horizontal base array. Copies of the task materials are given in Appendix C.

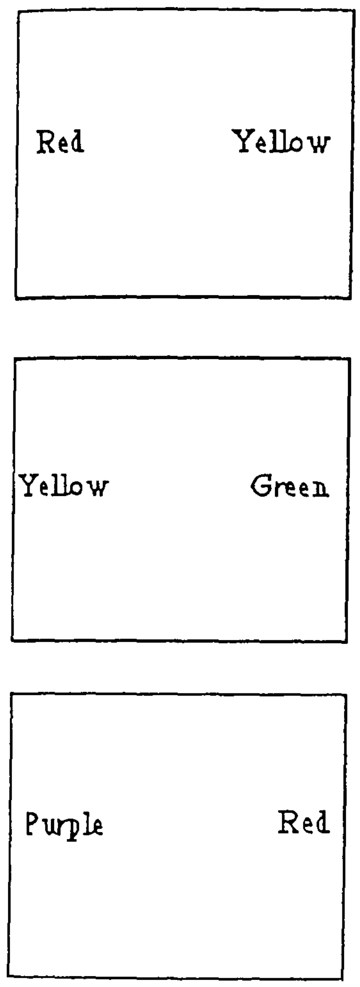
The premises were laid out as follows : BC

CD

AB

An example of this is given in Fig. 9.10 below.

Fig. 9.10: Example of horizontal base relationship.



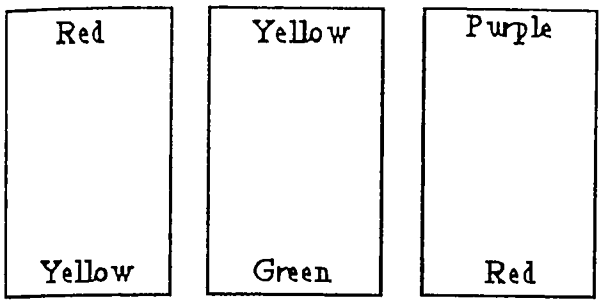
Vertical spatial relationship

The four-item array of acrobats in a 'human tower,' used for Experiments 8 and 11 formed the vertical base array. Copies of the task materials are given in Appendix F. The premises were laid out as follows :

B	C	A
C	D	B

An example of this is given in Fig. 9.11 below.

Fig. 9.11: Example of vertical base relationship.



Procedure

This was identical in all respects to Experiment 11, with one exception. Instead of referring to the person /child with the loudest voice, and then the next loudest etc., the appropriate task relationship was used. Thus:

- for those subjects using the relationship 'best', reference was made to the person/child who was the best at sums, and then the next best etc.

- for those subjects using the relationship 'fastest', reference was made to the person/child who was the fastest, and then the next fastest etc.

- for those subjects using the relationship 'oldest', reference was made to the person/child who was the oldest, and then the next oldest etc.

WISC-R digit recall scores were also taken from each subject.

9.2.2 Results

Note Each subject completed six experimental trials in total. However, a large number of the subjects in the previous studies erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier ones the first trial from each subject has been discounted.

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (base relation vertical, task relation best)	8.3
Group 2 (base relation vertical, task relation fastest)	8.1
Group 3 (base relation vertical, task relation oldest)	8.5
Group 4 (base relation horizontal, task relation best)	8.5
Group 5 (base relation horizontal, task relation fastest)	8.8
Group 6 (base relation horizontal, task relation oldest)	8.2

The average performance on the digit span test for ages between 6 years 8 months and 7 years is a score of 8 (WISC-R Manual, 1974).

Number of correct answers

Table 9.5 shows the mean number of correct answers for each experimental group.

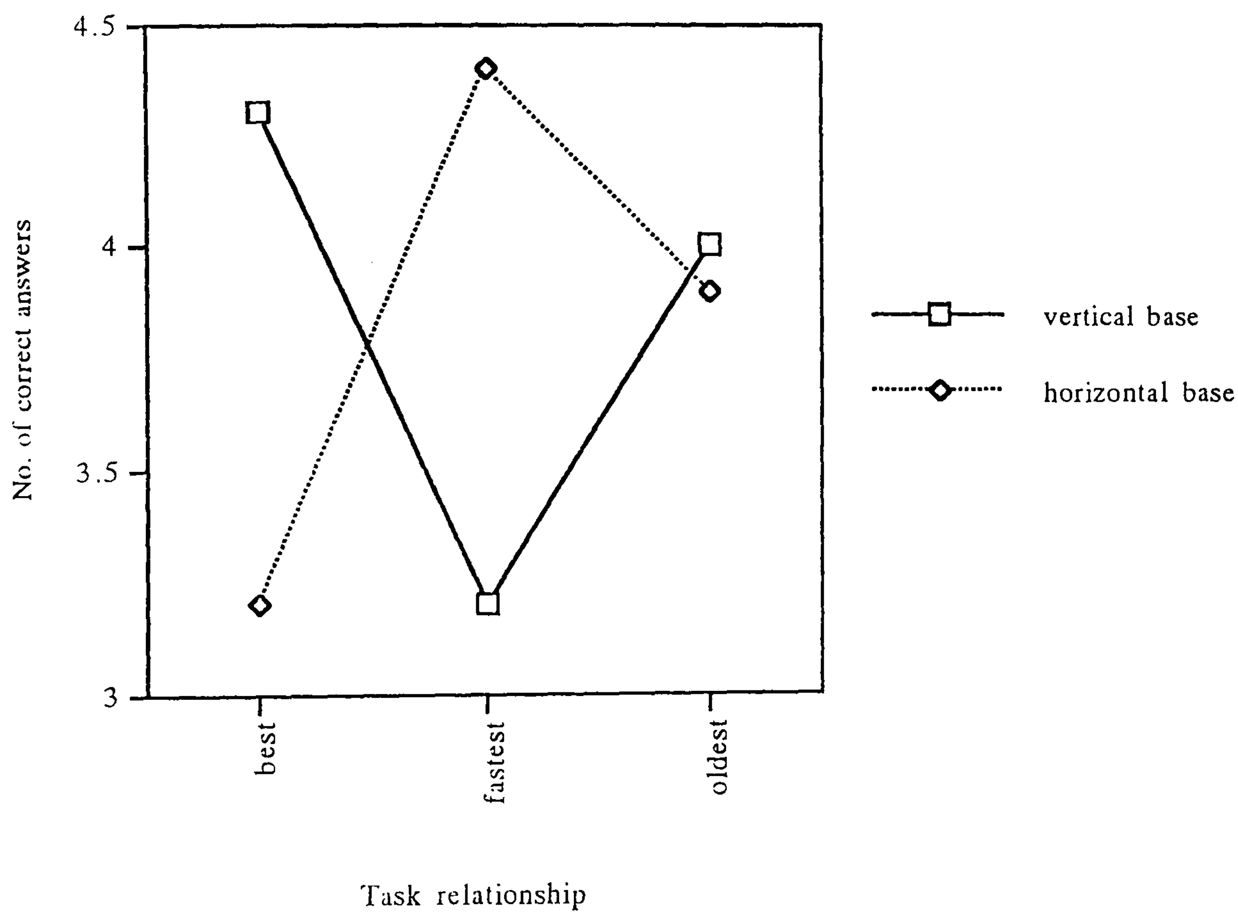
Table 9.5: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

		Task relationship		
		Best	**Fastest**	**Oldest**
Base Relationship	**Horizontal**	3.2 (0.92)	4.4 (0.70)	3.9 (0.99)
	Vertical	4.3 (1.06)	3.2 (1.03)	4.0 (0.67)

An ANOVA [2 (base relationship) x 3 (task relationship), both between subjects factors] was carried out on the above data. This revealed a significant interaction between the two factors ($F_{[2, 54]}=8.052$, $p<0.01$ - see Fig. 9.12).

Fig. 9.12: Base relationship x task relationship interaction.



Further analysis of the above data revealed the following significant simple main effects:

- the subjects ordering the task relationship 'best' and using the vertical base relationship scored more correct answers than those ordering the task relationship 'best' and using the horizontal base relationship ($F_{[1, 54]}=7.325$, $p<0.01$).

- the subjects ordering the task relationship 'fastest' and using the horizontal base relationship scored more correct answers than those ordering the task relationship 'fastest' and using the vertical base relationship ($F_{[1, 54]}=8.717$, $p<0.01$).

- a significant simple main effect of the task relationship when the horizontal base relationship was used ($F_{[2, 54]}=4.399$, $p<0.05$).

- a significant simple main effect of the task relationship when the vertical base relationship was used ($F_{[2, 54]}=3.915$, $p<0.05$).

Tukey comparisons were then carried out between the different levels of task relationship for both the horizontal and vertical base relationships. These revealed the following significant effects:

- when using the horizontal base relationship, subjects who were ordering the task relationship 'fastest' scored more correct answers than those ordering the task relationship 'best' ($q=4.18$, $p<0.01$).

- when using the vertical base relationship, subjects who were ordering the task relationship 'best' scored more correct answers than those ordering the task relationship 'fastest' ($q=3.8$, $p<0.05$).

Time taken to complete trials (correct answers only)

Table 9.6 shows the mean time taken (in seconds) to correctly complete the trials for each experimental group.

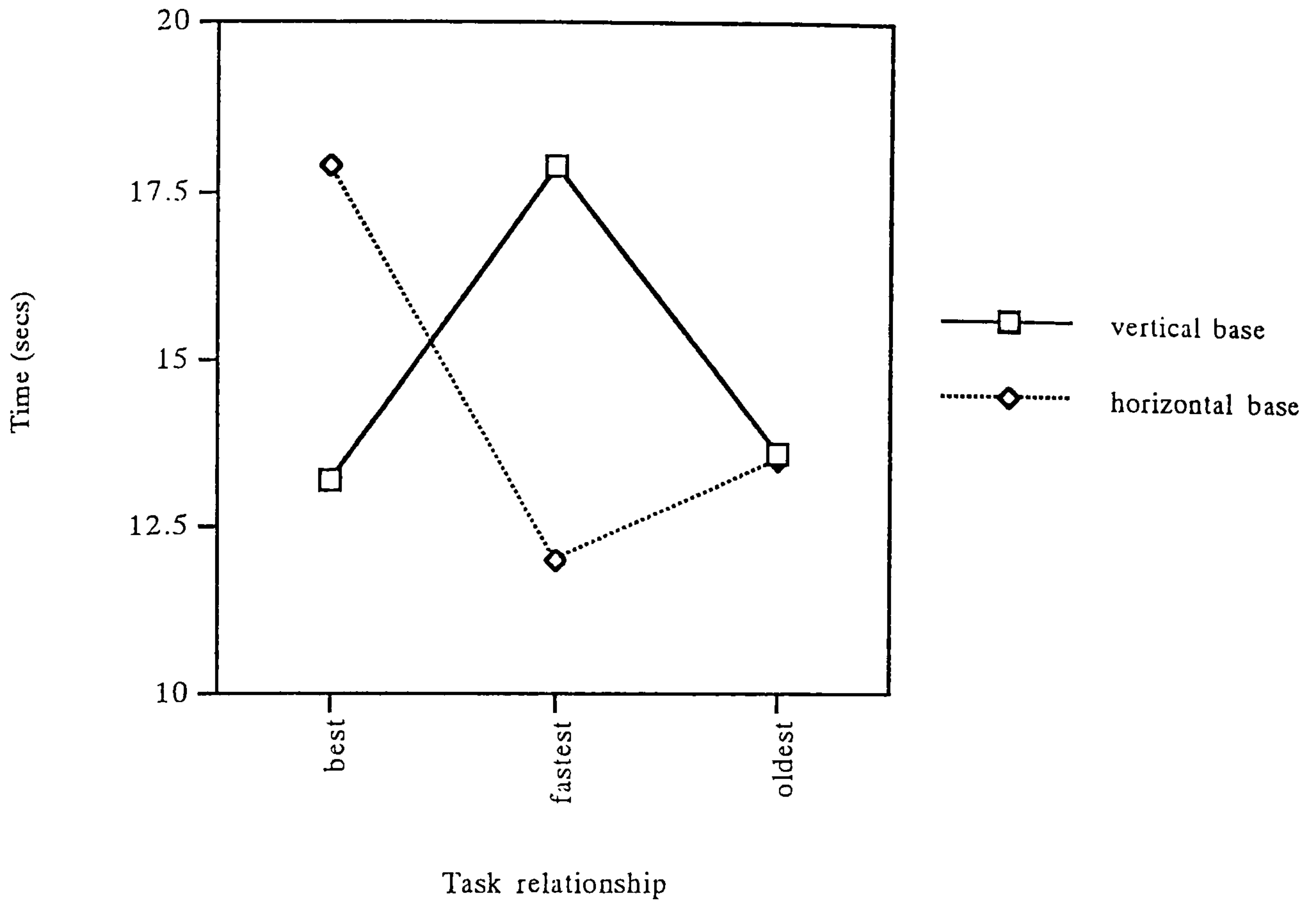
Table 9.6: Mean time taken to complete trials (correct answers only

Standard deviations are shown in parentheses

		Task relationship		
		Best	**Fastest**	**Oldest**
Base Relationship	**Horizontal**	17.99 (2.31)	12.00 (1.02)	13.47 (1.42)
	Vertical	13.19 (0.87)	17.90 (1.93)	13.58 (1.74)

An ANOVA [2 (base relationship) x 2 (task relationship), both between subjects factors] was carried out on the above data. This revealed a significant interaction between the two factors ($F_{[2, 54]}$=54.269, p<0.01 - see Fig. 9.13).

Fig 7.13: Base relationship x premise presentation interaction



Further analysis of the above data revealed the following significant simple main effects:

- the subjects ordering the task relationship 'best' and using the vertical base relationship produced correct solutions more quickly than those ordering the task relationship 'best' and using the horizontal base relationship ($F_{[1,54]}$=43.631, $p<0.01$).

- the subjects ordering the task relationship 'fastest' and using the horizontal base relationship produced correct solutions more quickly than those ordering the task relationship 'fastest' and using the vertical base relationship ($F_{[1, 54]}$=65.805, $p<0.01$).

- a significant simple main effect of the task relationship when the horizontal base relationship was used ($F_{[2, 54]}$=36.883, $p<0.01$).

- a significant simple main effect of the task relationship when the vertical base relationship was used ($F_{[2, 54]}=25.845$, $p<0.01$).

Tukey comparisons were then carried out between the different levels of task relationship for both the horizontal and vertical base relationships. These revealed the following significant effects:

- when using the horizontal base relationship, subjects who were ordering the task relationship 'fastest' produced correct solutions more quickly than those ordering the task relationship 'best' ($q=11.65$, $p<0.01$).

- when using the horizontal base relationship, subjects who were ordering the task relationship 'oldest' produced correct solutions more quickly than those ordering the task relationship 'best' ($q=11.80$, $p<0.01$).

- when using the vertical base relationship, subjects who were ordering the task relationship 'best' produced correct solutions more quickly than those ordering the task relationship 'fastest' ($q=9.16$, $p<0.01$).

- when using the vertical base relationship, subjects who were ordering the task relationship 'oldest' produced correct solutions more quickly than those ordering the task relationship 'fastest' ($q=8.40$, $p<0.01$).

9.2.3 Discussion

Both dependent variables have provided us with evidence to support the primary hypothesis. It is the case that performance using the relationship 'better than' is facilitated by the use of a vertical base relationship, whilst performance using the relationship 'faster than' is facilitated by the use of a horizontal base relationship. Thus it seems that these two relations are spatially tied, the former to a vertical ordering and the latter to a horizontal ordering. The claim made by DeSoto *et al* (1968), concerning a preference for only vertical orderings is therefore not upheld, at least for 7 year old children.

The position concerning the relation 'older than' is not so straightforward. If we consider the number of correct orderings, there are no significant differences between the dimension of the base relationship for that relation. There are also no differences between the 'older than' relationship and either of the two others ('better than' and 'faster than'). This applies to both vertical and horizontal base relationships. Thus it seems that children are just as likely to order the evaluative relation 'older than' correctly with either dimension (horizontal or vertical) of base relationship. Also, for both base dimensions, 'older than' seems to be at an intermediate level of difficulty when compared to 'better than' and 'faster than'.

A similar pattern of results concerning the dimension of the base relationship when ordering with 'older than' is also apparent for the time taken to correctly complete the trials. However, there are also some inconsistencies between the two dependent variables. The 'time' data showed that the relationship 'older than' was easier than 'better than' when using the horizontal base, and was also more difficult than 'faster than'. Considering the data from the vertical base, 'older than' was easier than 'faster than', but there were no significant differences between it and 'better than'.

The situation concerning 'neutral' relations therefore appears inconclusive. The error data demonstrates that children show no preference concerning the dimension of the base relationship, and find neutral relations to be at an intermediate level of difficulty. The 'time' data suggest that, whilst no preferences are demonstrated concerning the dimension of the base relationship, performance with the neutral relation 'older than' is equivalent to that using the relation 'better than' when reasoning with the vertical base. It would seem therefore that perhaps performance is facilitated when using a vertical dimension, at least when we consider the time taken to correctly order the arrays. However, further studies are necessary to tease out the difference in performance patterns between the two dependent variables, and also to ascertain whether the effect can be replicated for other relationships. This is beyond the scope of this thesis, but the design of possible studies, together with their implications, will be discussed in the concluding chapters.

It is interesting to speculate on the extent to which this facilitation is dependent on the similarity of the base to the abstracted schema, hypothesised by Halford (1992). Very few details are given concerning this, other than that spatial orderings are used (probably 'above' or 'on top of'). Halford does however state that orderings are abstracted from the child's everyday life, in which spatial orderings will be frequently be encountered. The child's common experience of orderings is with objects which can plausibly be ordered vertically. They will therefore be items such as toy bricks or other 'piling toys', or else will be placed such that they are visibly supported, for example on shelves or in dolls houses.

The studies in this thesis have used objects which can plausibly be ordered vertically. We have made use of towers of bricks, of acrobats and of crates hanging from a crane. Thus the objects are very similar to the child's everyday experience and presumably the children have had no problems in conceiving these objects as able to be ordered vertically. If their 'generalised ordering schema' is still closely tied to everyday experience, then it could be that only plausible vertical orderings will be able to be used as bases for ordering evaluative relationships. On the other hand, if the schema has been fully abstracted, then the actual situation in which the base ordering occurs will have no effect on performance.

In view of this, it was decided to conduct a study which looked at the effect of mapping from plausible and implausible base relationships onto 'far' series problems.

9.3 EXPERIMENT 13 - THE EFFECT OF IMPLAUSIBLE BASE RELATIONSHIP SITUATIONS ON THE MAPPING OF 'FAR' ANALOGIES

Rationale

This study is being carried out to ascertain whether 7 year old children are able to order abstract relationships using implausible base relationships. This will give us some idea as to the extent to which children are still relying on their everyday experience when performing 'far' analogical mappings. It was decided to use drawings of task objects, as before. Also, in view of the results obtained from previous studies in this thesis, incongruent task and premise relationships were used and the subjects worked with the relationship 'better than' and ordered using a vertical base. This was to ensure the best situation possible for successful performance using an abstract relationship (a 'far' analogy). The hypothesis being tested is that children will perform better when ordering with items in plausible base relationship situations.

9.3.1 Method

Design

A one factorial between subject design was used. The two levels of situation of base relation were as follows:

Condition 1 Plausible

Condition 2 Implausible

Measures were taken of the number of correct versus incorrect orderings achieved, and also the time taken to successfully complete each task.

Participants

The participants in this study were 20 mixed ability 7 year olds (mean age 7 years 9 months, range 7 years 3 months to 8 years 1 month) from state first and primary schools, with predominantly middle class catchment areas. The children had all previously participated in a series problem experiment, reasoning with a horizontal spatial relationship. This had occurred between 5 and 7 months previously. The subjects were assigned to the experimental groups according to the length of time which had elapsed since the previous study. This resulted in an even split of recency of experience between the two experimental groups.

Task descriptions and materials

4 item arrays were used throughout. The base relationship used was always 'on top of'. Premise were always presented horizontally, resulting in incongruous base and premise presentation relationships. For all conditions, the children were asked to reason about the relationship 'better than'.

Implausible base situations

The task items were side views of different coloured cars. The orders colours used were randomly varied, so as to give 6 different orderings of four colours (yellow, green, red, purple). The subjects were also given a rectangular piece of cardboard, stuck to the desk in front of them in a vertical manner, by means of adhesive putty. Appendix H shows an example of a fully ordered array.

Plausible base situations

The task items were identical to those used in the implausible situation. However, the children were told that the cars were in a multi-storey car park (see Procedure section) and a piece of cardboard, divided vertically into equal sections, was stuck onto the desk with adhesive putty. The subjects were told to use this when working out the answer to the puzzle (see Procedure section for a full description of this). Appendix H shows an example of a fully ordered array.

Procedure

This was as similar as possible to Experiment 12. However, because a large part of the instructions given to the subjects related to the mapping of the actual task relationship into the base relationship, the procedure is described in full below.

Subjects were tested individually in a quiet room. The experimenter first explained the task, using a complete photograph and serial ordering of premises as a worked example. The children then worked through the six examples, using copies of the individual 4 items to build a concrete array. Final ordering, time taken and order of item placement were recorded by the experimenter.

Implausible base situations

Wording of Instructions

"Can you remember the puzzles you did when you worked with me last time?.....They were a bit like this.......here's some little drawings of 2 people standing in a bus queue.....you can use them as clues to work out the order of all four people in the queue....Do you remember how we did it?.......Have a go with these...<child then worked through serial ordering of horizontal spatial relationship...any who were unsuccessful were excluded>.......Now, here are some drawings of cars......

I've got some puzzles for you to do, where you have to work out which car is the best, and then the next best, and then the next best, and so on. I want you to use these part drawings as clues to help you do the puzzle because you won't be able to look at the whole drawing until after you've finished.... Here are four separate drawings of the four cars......see if you can use these to solve the puzzles, like you did before.....I want you to use this piece of cardboard to put the drawings on.......There are six puzzles for you to do. They all use the same cars. The order they're supposed to be in will be different and the order of the drawings which I give you might make the puzzles a little bit harder."

Plausible base situations

Wording of Instructions

"Can you remember the puzzles you did when you worked with me last time?.....They were a bit like this.......here's some little drawings of 2 people standing in a bus queue.....you can use them as clues to work out the order of all four people in the queue....Do you remember how we did it?.......Have a go with these...<child then worked through serial ordering of horizontal spatial relationship...any who were unsuccessful were excluded>.......Now, here are some drawings of cars in a multi-storey car park......I've got some puzzles for you to do, where you have to work out which car is the best and then the next best, and then the next best, and so on. I want you to use these part drawings as clues to help you do the puzzle because you won't be able to look at the whole drawing until after you've finished.... Here are four separate drawings of the four cars.....See if you can use these to solve the puzzles, like you did before.....I want you to use this piece of cardboard to put the drawings on, because it's the car park.....There are six puzzles for you to do. They all use the same cars. The order they're supposed to be in will be different and the order of the drawings which I give you might make the puzzles a little bit harder."

WISC-R digit recall scores were also taken from each subject.

9.3.2 Results

<u>Note</u>

1. Each subject completed six experimental trials in total. However, a large number of the subjects in the previous studies erroneously tried to complete the first trial by recalling the demonstration ordering. In order to make the analyses comparable with earlier ones the first trial from each subject has been discounted.

2. Two subjects in total were unsuccessful in ordering the horizontal array used as an example and so were excluded from the experiment.

The results from the WISC digit recall test showed that the mean scores were as follows:

Group 1 (plausible situation)	8.5
Group 2 (implausible situation	8.6

The average performance on the digit span test for ages between 7 years 9 months and 8 years is a score of 8 (WISC-R Manual, 1974).

Number of correct answers

Table 9.7 shows the mean number of correct answers for each experimental group.

Table 9.7: Mean number of correct trials (max=5)

Standard deviations are shown in parentheses

Base relationship situation	
Implausible	**Plausible**
3.3 (0.82)	4.3 (0.67)

An one-way ANOVA was carried out on the above data. This revealed a significant effect of the base relationship situation, such that the subjects working with the plausible situation scored more correct answers than those working with the implausible situation ($F_{[1, 18]}=8.824$, $p<0.01$)

Time taken to complete trials (correct answers only)

Table 9.8 shows the mean time taken (in seconds) to correctly complete the trials for each experimental group.

Table 9.8: Mean time taken to complete trials (correct answers only

Standard deviations are shown in parentheses

Base relationship situation	
Implausible	**Plausible**
13.24 (0.89)	11.96 (0.77)

An one-way ANOVA was carried out on the above data. This revealed a significant effect of the base relationship situation, such that the subjects working with the plausible situation produced correct solutions more quickly than those working with the implausible situation ($F_{[1, 18]}=11.855$, $p<0.01$).

9.3.3 Discussion

The results discussed above have provided evidence in support of the experimental hypothesis. It was shown that 7 year old children are more successful at integrating abstract relational information into a single ordered array when they reason using a plausible base relationship, rather than an implausible one.

It would seem from this that the 'abstracted base schema' (Halford, 1992) is still closely tied to the everyday situations in which spatial orderings are encountered. 7 year old children seem unable to fully abstract the ordering relationship from these situations and they still require a familiar contextual situation in order that successful mapping will take place.

9.4 GENERAL SUMMARY

The three experiments reported in this chapter have investigated the existence of structural constraints on the mapping of abstract series problems ('far' analogies). It was shown that some evaluative relationships used in these types of tasks are tied to either a vertical or a horizontal array for 7 year old children. Furthermore, their performance when integrating separate pieces of this type of relational information can be significantly improved if the appropriate dimensional array is used as a base for analogical mapping. The children are then able to take advantage of the incongruence of task and premise presentation relationships which facilitates performance in spatial series problems. It seems that some evaluative relations are not spatially tied, but further research is necessary to substantiate this claim.

It is unclear from these studies whether the children are actually mapping from their internalised ordering schema (as hypothesised by Halford, 1992) to the abstract target domain, or whether they are mapping across from the external array (the drawings provided as part of the experimental materials) to the target domain. This will be fully discussed in Chapter 10. However, work reported by Riley, 1975 (reviewed in Chapter 4), claims that children use the external array purely as a memory aid, and that it is equivalent to the internalised array. If we accept this, then it seems from the last experiment in this chapter that the internalised array is still closely linked to the child's everyday experience of ordering. When the array is manipulated such that it represents an implausible situation, performance becomes worse.

The current chapter has presented the last of the empirical work undertaken for this thesis. It has shown that, whilst the situation is more complex, the constraints on integrating spatial information identified in earlier chapters also exist for abstract tasks. The following chapter will present and discuss a theoretical integration of the research reported in Chapters 4 to 9. The limitations of this work, together with suggestions for further work to follow up the findings, will also be discussed.

CHAPTER 10 : DISCUSSION AND CONCLUSIONS

10.1 OVERVIEW OF THESIS

This thesis has been concerned with the role of structural task representations in the performance of children on analogical reasoning tasks.

Recent work in the area of the development of analogical reasoning has shown that children are able to recognise and use the relational similarity constraint to solve both classical and problem analogies from a very early age. Many experiments (for example, Goswami and Brown, 1989) have provided evidence that 4-5 year old children can answer problems using relational reasoning rather than by using associative reasoning, as long as they possess the necessary domain knowledge. This research contradicts earlier claims in the Piagetian tradition concerning the late development of analogical reasoning ability.

Although there is now ample evidence from studies such as those discussed by Goswami, 1992, that young children can, in optimum conditions, map individual and systems of relationships from base to target domains, we have little knowledge of what might affect this ability. Obviously, as these studies have shown, relevant domain knowledge is crucial. However, this thesis has argued that there are other factors which will affect whether children can recognise and map relational similarity. In classical analogies, the relational structure of the task is very salient to the subjects, in that it is presented as part of that task. Provided that the child knows how the A & B items are related to each other, they can apply this same relationship to the C item and thus solve the problem. In many situations however, the task structure is not as obvious and there may be more than one way to represent the task objects and relationships between them. Because of this, the claim was made that the ability to construct a structural task representation will affect the ability to reason analogically, and that for many types of analogical problems, representation of salient task objects and the relations between them is a necessary precursor to the recognition and use of the relational similarity constraint. For this reason, the research here has concentrated on how children build up a representation of the system of relationships in a problem, which can be used for analogical mapping. This has been carried out by building on Halford's (1992) theory that children solve series problems by analogy to a generalised ordering schema (a base domain) which has been abstracted from everyday life.

Experiment 1 was mainly exploratory, and was designed to look at differences in the use of the generalised schema when children reason with either 'far' or 'near' analogical target problems. However, regardless of the 'nearness' of base and target domains, it was found that 5 and 7 year old children were unable to integrate task items into a single external task representation and were also unable to answer questions about the relationships existing between task objects. It was therefore concluded that these children were experiencing difficulty in constructing an appropriate structural task representation and that this difficulty was affecting their ability to map relations between base and target domains.

There are of course two problems with this conclusion:

1) The observed performance of the children when constructing an external integrated representation of the relational information may not reflect similar performance when constructing a similar internal representation.

   This is a claim which can never be directly tested, as we have no direct access to internal representations. However, previous research by Riley, discussed in Chapter 3, showed that children working both with and without external task representations displayed very similar patterns of results when solving series problems. Thus it is argued that we are justified in using subject's performance in working with an external array as an indication of how they are constructing a mental representation of the task.

2) There may be no direct causal link between the children's inability to correctly represent the relational structure of a series problem and their inability to answer the inferential questions.

The link between structural task representations and the successful answering of inferential questions was not directly tested by the manipulation of the presence or absence of integrated relational information. However, there is a wealth of evidence (reviewed in Chapter 2) which shows that children and adult novices solve series problems by the construction of integrated internal array. If we accept this evidence, then it follows that the lack of the ability to construct such a representation will result in poor performance. Experiments 1 and 2 provided us with more evidence in support of this claim, in that the subjects were unable both to answer the inferential questions and to correctly integrate the relational information into an ordered array.

Because of the evidence obtained from Experiments 1 and 2, that is, that children were unable to integrate separate items of information such that relations between objects were correctly represented, it was decided to carry out a series of further studies which would investigate some factors which, it was hypothesised, were contributing to this inability. Because our aim was to look at the development of analogical reasoning, rather than the solving of successful problems, the children were no longer asked to answer inferential questions. We were interested purely in how the subjects constructed a structural task representation such that the relational information encoded in that representation could be successfully mapped from base to target domains.

Experiments 3-13, with 3 exceptions, were all carried out with 7 year old children. This was because the aim of the research was to investigate what factors might affect young children's ability to reason analogically, rather that to develop an age-based account of how children's ability to construct structural task representations might develop. Thus the approach adopted is to consider constraints inherent in the task which are having a significant effect on a crucial cognitive skill; that of recognising and using relational similarity. It could well be that these constraints are having an effect because of the children's lack of experience in structural reasoning, rather than because of an age based difference *per se*.

It could well be that there are other age related differences which are also having an effect; the suggested limited capacity of working memory is an obvious candidate. However, it is not within the scope of this thesis to give any comments on this. Nonetheless, we have evidence from Experiments 3 to 13 that there are pragmatic task constraints which are affecting analogical mapping from base to target, because they are affecting the construction of structural task representations.

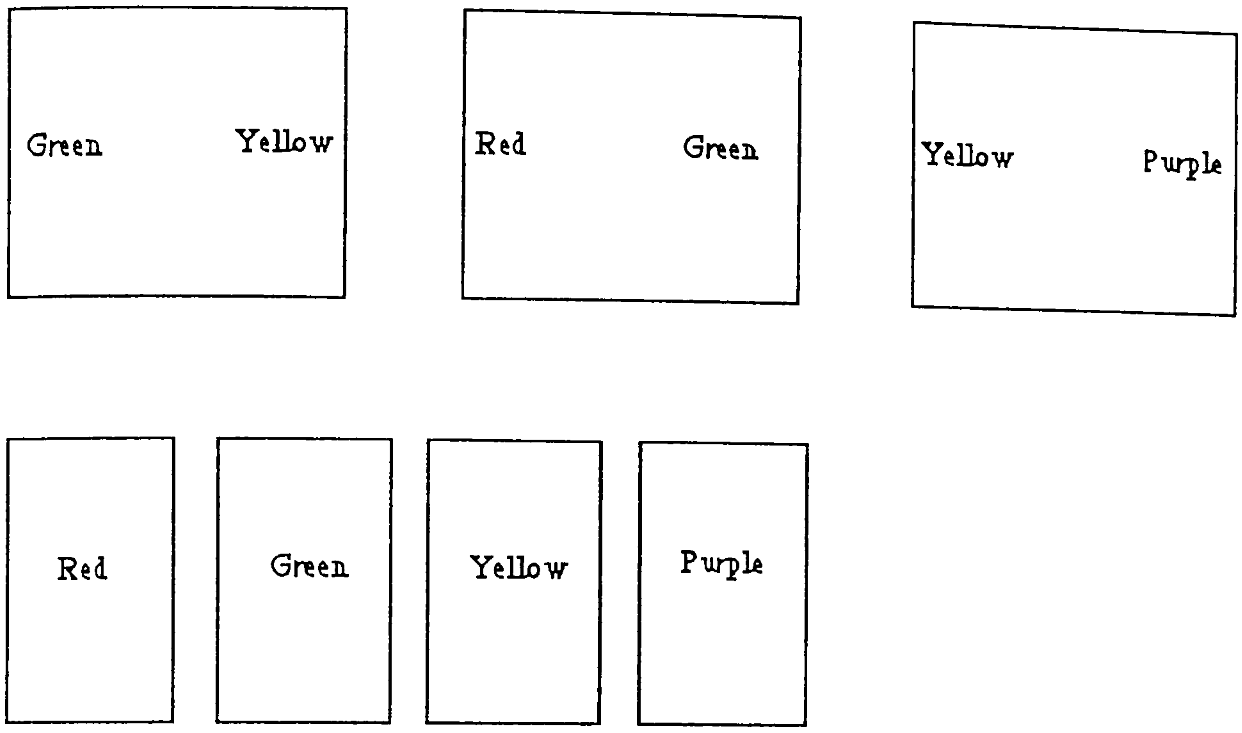
The results reported in this thesis could have occurred because young children are relatively inexperienced at some types of analogical reasoning. Thus, when they are presented with a task for which the requirement to represent it at the structural level is both necessary and nontrivial, their ability to use the relational similarity constraint is significantly affected. Individuals with more experience will probably be able to overcome difficulties caused by the task features, as their ability to represent task objects and their relations to each other will be better developed and practised. However, it could be that more complex tasks will result in similar effects from task constraints in adult problem solvers.

The point of this thesis is not that children of a particular age are able or unable to construct task representations, but that there are factors external to the individual (i.e. in the task) which can have an effect on the ability to reason analogically. Experiments 3 to 13 investigated two possible reasons inherent in the task for the inability of young children to construct an integrated relational task representation. The following section will take each of the hypothesised task constraints in turn and discuss the empirical evidence which was collected.

10. 2 TASK CONSTRAINTS AFFECTING THE CONSTRUCTION OF AN INTEGRATED EXTERNAL REPRESENTATION.

The two exploratory studies described in Chapter 4 indicated that 5 and 7 year old children were unable to correctly integrate separate premises containing relational information into a single array. Because of this, Experiment 3 investigated two possible factors, information ordering and type of stimulus material, which might have been having an effect on performance. An example of a task presentation, together with the corresponding correctly ordered array, is given in Fig. 10.1 below.

Fig. 10.1: Presentation of relational information and correctly ordered array.



The results from Experiment 3 to 10 will now be discussed in terms of two underlying factors. One of these, that concerning an inability to recognise that mid items (i.e. the 'red' and 'green' items in the example above) are represented twice in the relational information such that there is some redundancy of information, was explicitly raised in the introduction to Experiment 3. The second factor has emerged gradually during the implementation of the experimental work, as a means of explaining several of the results. It will be suggested that the common explanatory feature for the results in Experiments 3 to 6 and 8 to 10 concerns a failure to understand the significance of the gap between separate pieces of information.

1. The redundancy of mid item information.

It was suggested that the reason that the subjects were unsuccessful in constructing a single task representation was because they did not realise
that this required an integration of separate items of relational information, such

that the mid items in the array (which were represented twice) needed to be amalgamated into a single instantiation of each item. If the children did not appreciate that mid items were represented twice only so that their relationships with both adjoining items could be represented, then they would not understand that forming a single integrated array necessarily meant that only one of each mid item would be required. This claim was tested in two ways. The first involved contrasting the effects of using photographs and drawings as stimulus material (Experiment 3). There was no significant difference between these two conditions. We could not therefore conclude that the use of photographs, rather than drawings, will aid children in realising that two similar tokens are in fact of the same character and need to be amalgamated. Thus previous evidence concerning the use of photographs as easy graphic representations for young children, which was used as the rationale for this experiment, did not aid us in facilitating children's performance. A second experiment (Experiment 7) directly tested the hypothesis that 7 year old children are unaware that only one token for each mid item needs to be present in the completed single representation. This was done by presenting different groups of children with either one of each item to work with whilst constructing the array, or six of each item.

Again, there is no evidence that this had an effect on performance, as the experimental manipulation resulted in no significant difference between the different groups of subjects. The children appeared to be aware that they only needed one of each item as duplicate items, when available, were quickly discarded.

Given the lack of any differences for both of the experiments mentioned above, it seems unlikely that an inability to recognise the redundancy of mid item premise information was affecting performance. In view of this, an alternative reason was investigated.

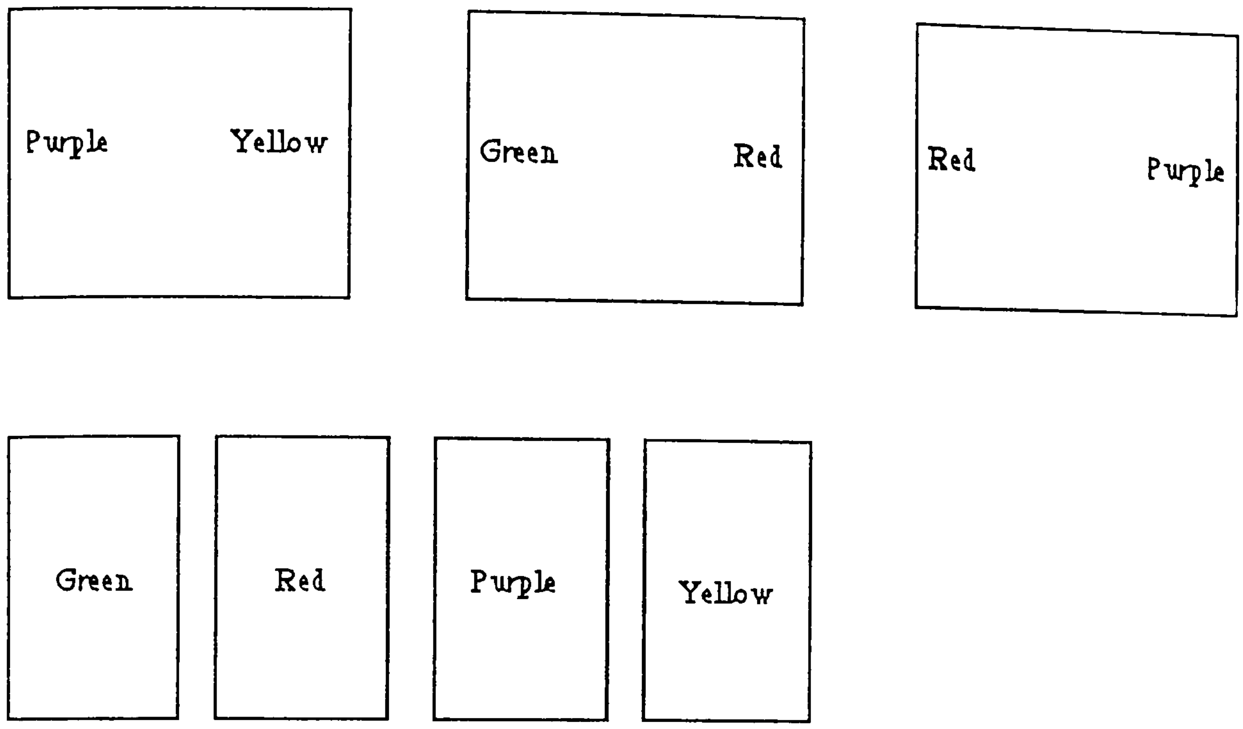
2. The psychological significance of the gap between relational information

This was initially suggested as a result of Experiment 1, where it seemed that some arrangements of premise orderings were easier than others. This was followed up in Experiment 3, which found that those orderings which required an item to be placed at the front of a partially ordered array posed particular problems. An examination of the order in which the subjects placed individual items also showed that they dealt with information in the order in which it was presented to them. This meant that a representation would be built up by initially considering the first presented premise. In Fig. 10.1, this gives the relation between the 'red' and 'green' items. The second premise concerns the relation between the 'green' and 'yellow' items. The children appeared to have no problems in realising that only a 'yellow' token (because it is the only novel item) needs to be added to the array. At this stage, they do not need to appreciate the significance of the gap between the two premises in order to correctly place the novel item. This is because they can reproduce the simple left to right ordering in the premise information and place the 'yellow' item to the right of the 'green' one. This, however, is not the case when they consider the last premise, where the 'red' item is novel. The following of the left to right ordering pattern in this instance results in an incorrect placement, as the 'red' item should be placed to the left of (at the front of) the partially ordered array. However, the fact that the children do not realise that the gap between the second and third premises is significant for their goal, means that they fail to recognise that the familiar left to right ordering pattern is inappropriate.

It is interesting to note that this limitation to a successful relational task representation is quite robust. The relational information used in these experiments consisted of parts of the complete photograph and drawing of the array. Thus, each premise and each piece of relational information was one item of information (usually one drawing) with two different task objects (usually people) depicted on it. An example of this is given in Appendix C. The three premises were placed one at a time on a desk in front of the child. It therefore seems that the gaps between premises were made relatively obvious. Items immediately adjacent to each other were shown on one single premise drawing, with no desk space between them. On the other hand, items which were not necessarily immediately adjacent to each other in the integrated representation were in separate drawings, with typically 70 to 100 centimetres of desk space between them. However, it does seem that this was not sufficient to make the significance of the gap salient to the children. This is because those orderings which required items to be placed to the left of a partially ordered array resulted in inhibition of performance. This pattern of results was replicated in Experiment 4, but using familiar toys in an everyday context as the stimulus material.

Experiment 6 investigated the existence of a similar inhibition in performance for 9 year old children. The results of this study showed a slightly different pattern, however, in that the children's performance was only affected for one of a possible six orderings. An example of this ordering, together with the appropriate integrated array, is shown in Fig. 10.2 below.

Fig. 10 2: Example of CD AB BC ordering and the appropriate integrated array

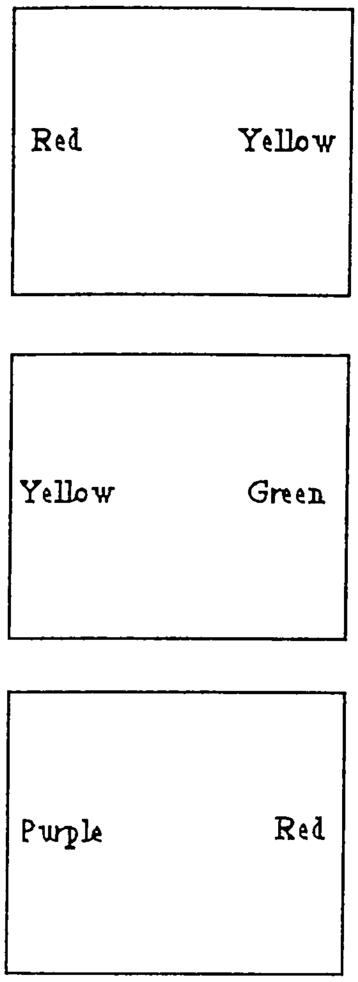


For all the other orderings which required novel items to be placed to the left of the partially ordered array, performance was not affected. This is because the 9 year olds were able to spontaneously self-correct partially ordered arrays. Typically, an incorrect item placed to the right of a partially ordered array was noticed when the subject checked their answer by scanning all three pieces of relational information. They were usually able to correct their error at this stage. However, the children working with the ordering above were unable to engage in successful self-correction. Typically they would re-order the B item (the 'red' item in the example above) after scanning the third premise. However, they would not move the 'A' item (the 'yellow' item) with this, and so the ordering would still be incorrect. It is unclear, however, whether the children did not notice that their answer was incorrect or whether they were unable to correct their error. Nonetheless, this has been interpreted as a failure to make use of the special status of the relational information given as part of the task. Although 9 year olds are able to overcome the propensity to order in a left to right manner, they are still unable to correctly integrate the relational information as their self-correction means that they become misled. Of course, this may well be a problem for the 7 year olds too. However, their performance was inhibited by the incorrect left to right ordering pattern, such that errors due to the status of relational information could not be isolated. It seems therefore, that the studies reported in Chapters 5 and 6 were demonstrating a tendency for both age groups to order the array in a left to right pattern, even when this resulted in incorrect placements. The 9 year old children could correct this, but were unable to recognise that they needed to preserve relations between items which were given as task information.

This could be explained as an over generalisation of the ordering schema hypothesised by Halford, 1992. The familiar left to right ordering pattern is firmly in place, but the children have not recognised the importance of integrating and maintaining relational information in order to arrive at the correct left to right ordering. If this were the case, then the situation for the use of analogies in structured tasks would appear rather bleak. It would seem that children are unable to derive the correct structural relations for a domain, if this means that they have to deviate from well used ordering patterns. Thus, we would have to be very cautious in situations where children are asked to structure complex relational information in order to perform analogical mappings.

Chapter 8, however, investigated the factor which was underlying the incorrect left to right ordering patterns. Rather than a simple and undifferentiated application of a left to right pattern, we have evidence that this error was due to a lack of appreciation of the importance of the gap between premise information. This was shown by manipulating the congruence of the dimension of the task relationships (vertical or horizontal) and the dimension in which the relational information was presented. Incongruent conditions (where the task and information presentation relationships were different) resulted in significantly improved performance. These conditions highlighted the gap between the separate pieces of relation information. Fig. 10. 3 below shows an example of an incongruent condition.

Fig. 10.3: Example of incongruent task and information presentation relation



In the task represented above, the child cannot simply order the array in the way it was originally presented, because the relations are not congruent. Thus, in order to construct the horizontal array, the pieces of relational information must necessarily be placed in a different dimensional relation to the one in which they are presented. This highlights the gap between each piece of relational information, as the subjects are forced to consider each set of paired items as a discrete and separate piece of data. This pattern of results was replicated with familiar toys in an everyday context, and also with 5 year old children ordering 3-item arrays. Thus, the very poor performance demonstrated by the 5 year old children might be due to the same factor. However, we would need to carry out further studies with this age group, using a larger number of items and a variety of different orderings before any firm conclusions could be reached.

The experiments carried out in Chapter 8 also ruled out the possibility that the incorrect ordering pattern was due to the cultural effects of learning to read and write. Similar patterns of inhibition for congruent task and information presentation relations were found when subjects were required to order a vertical array from the bottom upwards (Experiment 8) and from the top downwards (Experiment 9).

To summarise, therefore, children are constrained by factors inherent in the task when they are performing relational mappings. These factors are the ordering of the relational information and the congruence of task and information presentation relations. Unless the gap between pieces of relational information is made salient, children are unable to identify and preserve such information when they are constructing structural task representations. This results in incorrect analogical mappings.

10. 3 STRUCTURAL REPRESENTATIONS AND 'FAR' ANALOGIES

The experiments described above looked only at 'near' analogies. These are those tasks which share the same spatial relation with the generalised ordering schema. The final three studies used the insights gained to investigate the construction of structural task representations when reasoning with 'far' analogies; those which share only the relational structure with the generalised schema.

Experiment 11 (the first in Chapter 9) was designed to ascertain whether the facilitation effect gained for incongruent task and information presentation relations would still be present when the task was a 'far' analogy when compared to the generalised schema. Again we achieved a similar pattern of results. This evidence adds further weight to the literature reviewed in Chapter 3, maintaining that children solve abstract series problems by integrating the premises into a single spatially ordered array. If we accept Halford's claim that this is done by mapping from a generalised ordering schema, then it appears that for both spatial ('near') and abstract ('far') analogies, the generalised schema is vulnerable to features inherent in the presentation of the task. These features can inhibit the correct integration of relational information by causing the child to join the task items in the way they are presented, that is, by following a simple directional pattern (left to right, top to bottom or bottom to top). It is not that the child has an explicit misunderstanding that this is the correct action, otherwise we would not be able to affect performance by manipulating task and information presentation relations. Rather, it seems that their ordering schema is not robust, such that it can be influenced by external features.

Experiment 13 provided further evidence concerning the vulnerability of the subject's ordering schema. It was shown that the implausibility of the situation used to externally demonstrate the ordering schema adversely affected performance. Thus, when the children were reasoning with an abstract evaluate relation, they were unable to successfully integrate the separate pieces of information if the task items (cars) were ordered in a vertical relation with no visible means of support. As we had already established that the actual dimension used for the ordering resulted in successful performance for the particular relation used, it seems likely that the fact that the cars appeared to be suspended in mid air meant that the children could not successfully order the array. Doing the same task but with the cars in a multi-storey car park resulted in significantly better performance.

It seems therefore that the schema is not retrieved intact from long term memory, as Halford suggested. The general ordering pattern is in place, but the child is reconstructing the ordering schema piecemeal, and is dependent on constraints in each specific task.

Experiment 12 investigated the claims made by Handel and colleagues that certain evaluative relations are tied to a vertical relation. The results showed that whilst some evaluative relations appear to be linked to a vertical generalised schema, others are linked to an horizontal one and some appear to be linked to neither. This finding provides us with some evidence that there is not a prior preference for vertical spatial orderings, as suggested by Handel *et al.*

It also adds more weight to our assertion that children's generalised ordering schemas are still quite closely tied to everyday, real-world experiences. This is because the children were sometimes unable to utilise the 'given' schema for the evaluative relations. This implies that they have two generalised ordering schemas, one which orders vertically and the other horizontally. For the tasks which they were performing, they could not use a schema which was not tied to the relation. This also calls into question Halford's implicit assumption that the child's generalised schema is a vertical one, using the relation 'above'. However, it is difficult to give much weight to this, as Halford himself, whilst using the vertical relation 'above' to verbally describe the schema, depicts it diagrammatically in the horizontal dimension!

To summarise, the research in the thesis has shown that:

1. Children's structural representations are adversely affected by features inherent in the task.

2. The features prevent the children from recognising and using the gaps between individual pieces of relational information.

3. The resultant incorrect structural representations adversely affect analogical mapping.

4. Children's generalised schemas (i.e. their structural representations) which are used in analogical mapping, are not retrieved intact from long term memory, but appear to be reconstructed using information from the target domain.

10.4 LIMITATIONS

The above conclusions have been drawn by making two major assumptions, based on evidence which has been reported previously in the literature. These will be discussed briefly below.

1. Are series problems solved by analogy?

The primary aim of this thesis was to investigate the construction of structural task representations for the purposes of analogical mapping. Previously, reasoning using relational similarity (i.e. analogical reasoning) has been investigated by using either classical or problem analogies. However, they were both unsuitable for the studies reported here. Classical analogies have an obvious structure inherent in their presentation. Problem analogies, on the other hand, have their relational structure embedded in a narrative. In order to discover how the child is representing this, it was argued that we would inadvertently begin to structure the task ourselves, by asking the relevant questions. Because of this, a decision was made to follow Halford's theory that series problems are solved by analogy to generalised ordering schemas. This seemed intuitively appealing and allowed us to look in detail at the actual integration of separate pieces of information into a single structural representation. Thus we have gained several insights into how children go about representing knowledge at a structural level, and particularly how this can be affected by external localised features. This then is important in its own right. However, the assumption that series problems are solved by analogy has made the explicit addressing of structural representations for the purposes of analogical mapping possible. We can now use the findings from this thesis and see if they generalise to other forms of analogical reasoning.

2. The nature of the base schema

This research followed Halford, 1992, in asserting that these types of analogies are solved by the mapping of relations from a generalised ordering schema, abstracted from everyday life into the target domain. We made claims about the status of the subject's mental representation by using evidence from how they actually built up an external array of concrete task items. This was based on Riley's (1975) work which showed a strong similarity in patterns of results from children who either did or did not use external ordering aids when solving series problems. However, we can never be completely sure that these external arrays are a true and complete example of mental representations.

We have assumed that the external array is a representation of the **generalised schema** used as a base domain. Again, we can never be completely sure of this. However, for both concrete and abstract analogies, the children were happy to use spatially ordered arrays. Because abstract ('far') analogies share only relational structure with the spatially ordered array, then subjects have to be mapping relational structure in order to use the spatial array as a base domain. Thus it seems that the children did have a generalised spatially based ordering schema which they could bring to bear on problems.

Without the two assumptions above, namely, that children solve series problems by analogy and that the external array is a representation of their generalised ordering schema, it would have been impossible to address the issues raised during the course of this research. It is therefore argued that the assumptions were necessary and adequately justified by previous research.

10.5 A RECONSIDERATION OF THE TYPES OF KNOWLEDGE REQUIRED FOR ANALOGICAL REASONING

Chapter 2 reviewed the literature concerning analogical reasoning in both children and adults. Research with adults has emphasised both the importance of the representation of task structure and the induction of generalised schemas from the use of common relational structures (e.g. Holyoak, 1985). Recent work by Keane (1994) has called for a much broader view of the domain, with more research being carried out into the way pragmatic task constraints such as task instructions or the presentation of material can affect analogical mapping. This view is consistent with the argument which has been presented in this thesis, which has shown that young children's abilities are affected by factors inherent in the task presentation.

Over the last decade or so, work with young children has shown that they are capable of reasoning by analogy providing that they have the relevant domain knowledge (e.g. Goswami, 1992). Thus it appears that Piagetian claims that children below the age of 11 years were unable to reason using relational similarity were confounded as his subjects had not been pretested for possession of the necessary knowledge about the subject in which they were required to reason.

However, these studies have all used situations where the task structure has been made very explicit. Classical analogies of the form A : B :: C : ? are necessarily presented with an integral task structure. There is only one relationship (that between A and B) which can be mapped over onto the C and D terms and the necessity to do this is made explicit by the stated task goal. Research using simple problem analogies with young children have also made the goal of the task explicit and have used a narrative in which only relationships which require mapping have been emphasised (e.g. Brown, Kane and Long, 1989). Further studies have trained children to use an analogical problem solving strategy (Brown and Kane, 1988) and have also resulted in improved performance. However, manipulations such as these which have increased children's meta-knowledge of analogy have still not resulted in total success. Thus there must be factors other than domain and meta knowledge which are affecting performance. This thesis considered the role of structural knowledge in the development of the ability to reason analogically.

Both classical and problem analogies are unable to directly address the issue of **how** the structural representation is constructed. This is because classical analogies have their structure explicitly represented as part of the task. On the other hand, problem analogies are inherently ill defined and as such the crucial relations are embedded within a narrative which could contain some irrelevant details. It would therefore be difficult to explicitly ask children to represent the relational structure of the problem, as they may not be aware of exactly what we are requiring them to do. Asking them questions about relational structure as we see it does mean that we can highlight its importance, but provides no insight into how children would represent relationships if they had not been cued to do so by the experimenter.

The class of problems used in the research reported in this thesis has shown that the construction of an appropriate task representation can provide problems for young children. Thus we can argue that there is another level of knowledge which children need to be able to use for a full use of the relational similarity constraint, in addition to those already documented. These are:

1. Domain level i.e. knowledge of the actual subject area.

2. 'Meta-level' knowledge i.e. knowledge that reasoning by analogy is an appropriate way to solve the task.

The third level of knowledge investigated in this thesis concerns the ability to correctly represent relational structure. Evidence has been provided that children have to be able to represent tasks at the appropriate structural level before they can perform the necessary relational mappings between base and target domains.

10.6 IMPLICATIONS

There are two main implications which can be drawn from the research reported in this thesis. They are as follows:

1. The importance of structural representations in analogical reasoning.

This research has shown that children are affected by external task constraints when they perform analogical mappings. In other words, how the task is presented can affect the construction of a representation which depicts objects and their relations to each other. This structural representation in turn affects how children make analogical mappings between base and target domains. It therefore follows that if we think carefully about how we present tasks to children, we may well be able to affect what aspects of the base domain are mapped into the target domain. Presentation of a task in an appropriate manner will facilitate successful analogical transfer. The appropriateness of task presentation will, of course, differ according to individual tasks. However, it seems likely that we must structure children's learning such that the important features in both base and target domains are highlighted and mapped, rather than mapping occurring solely between those features which are most obvious. The representation of a task at the structural level is neither trivial nor obvious for many young children. If we can support this ability by highlighting appropriate mappings, then we should be able to improve children's use of analogy as a knowledge acquisition device.

2. How to structure the task.

The way in which performance was facilitated was by making the children structure the task in an opposite dimension to the way in which it had been presented. Thus the significance of the relational information was highlighted. This has some implications for the realm of scaffolded learning (Wood, Bruner and Ross, 1976). This term is used to describe the process by which difficult tasks are broken down into a series of smaller tasks which children can manage. Eventually these small tasks are arranged so that the children can eventually carry the whole cognitive task themselves. Wood (1988) emphasises that "built well, such scaffolds help children to learn how to achieve heights that they cannot scale alone" (p.80).

The research reported in this thesis highlights the care with which learning environments need to be structured. When young children solve series problems, their performance is inhibited by structuring the task such that they need to expend less cognitive effort. Instead, we were required to present the task such that the subjects were made to explicitly focus on the gap between relational information by presenting task information in an opposite domain to the task relation. In other words, rather than simply breaking the task down into components and carrying out some of these components for the subjects, we actually structured the task so that the children's attention was drawn to the important task components. This is an alternative means of scaffolding the task, by highlighting the features which children need to be aware of.

10.7 FURTHER RESEARCH

There are three main areas in which further studies based on the research reported in this thesis could be carried out.

1. Using the domain of series problems

As previously discussed, these are useful types of analogical problems in which to investigate children's construction of structural task representations. This is because the task structure, whilst not being presented to the child as part of the task, as in classical analogies, is a necessary and salient part of the problem solving process, so that the children are required to make the structural representation explicit. There has been little research which directly addresses children's ability to construct such representations. These types of tasks could therefore be usefully employed in looking at younger children's abilities in this area. Experiment 10 in this thesis showed that 5 year old children's performance with 3 item arrays can be facilitated by varying the way in which the task is presented. The children's performance was then similar to that of the 7 year olds, except that fewer task items were used. This meant that the 5 year olds were able to correctly integrate separate pieces of relational information when the gaps between these pieces of information was made salient.

We now need to carry out the same type of study, but with an increase in the number of task items which the 5 year old children are required to reason with. If this manipulation results in equivalent performance to that of the 7 year olds with the same number of task items, we will have some evidence that incorrect structural task representations, caused by task constraints, are affecting performance, rather than working memory limitations *per se.* This result would also explain the successful performance reported by Pears and Bryant (1990) when 5 year old children were ordering 5 or 6 bricks in a single tower. This is because the relational information was presented to the children in a horizontal dimension whereas they were required to order vertically. Thus the task presentation made the gaps between the relational information salient.

This line of research could also be extended to look at 5 year olds performance using abstract evaluative relations. One of the original aims of this thesis was to investigate any differences between 'near' and 'far' analogies (Gentner, 1989). For the purposes of the research described here, a 'near' analogy has been defined as a spatial series problem, whilst a 'far' analogy has been defined as an abstract series problem. Because our first experiment showed that 5 year olds were not successful with the types of problems we were studying, their performance with abstract series problems was not investigated further. However, now that we have some insight into the factors which are affecting spatial series problems with 5 year olds, and also abstract series problems with 7 year olds, we can begin to investigate whether manipulation of these factors will also facilitate 5 year olds performance when working with abstract series problems.

The studies reported in this thesis have investigated the presence of the 'congruence constraint' when working with left to right, top to bottom and bottom to top orderings. Similar patterns of results have been obtained for all these types of problems. We have not however investigated any effects when working with right to left orderings. It could be that these will be less strong, as there may be a strong preference for left to right ordering patterns due to the direction in which the English language is written. Nonetheless, we have also obtained an effect due to the 'congruence constraint' for both directions of vertical orderings. It would seem therefore that the 'congruence constraint' is not solely due to the cultural effects of literacy, and should be present for all types of orderings. The investigation of right to left orderings will test this hypothesis.

Most of the research reported here has looked at children's performance when working with spatial series problems. The last three reported studies, however, began to explore the construction of structural task representations when children are reasoning using abstract evaluative relations. It seems from these experiments that performance is affected by the dimension of the schema with which the children reason, and also the degree of plausibility of that schema. Handel and colleagues, working in the 1960's, claimed that some evaluative relationships were tied to vertical orderings, whilst others were 'neutral'. However, the research reported here has provided us with some evidence that evaluative relationships can also be tied to horizontal orderings. This work needs to be extended, so that other types of evaluative relationships can be investigated. If we can find evidence that some of these are tied to one dimension of ordering, then we will have an important piece of information which could help us in deciding how scaffold the structural representation of analogical tasks in educational contexts.

The degree of plausibility of the ordering schema was also shown to have an effect in the last reported experiment. Again, this could be an important result when we are thinking about task presentation, but first of all it needs to be replicated and extended, using task items other than cars and relationships other than 'better than'. In the study reported here, the plausible schema was supported with both a verbal cue (in the instructions) and with a visual cue (the surface which the children used to order on). Varying just one of these cues would give us some insights into the extent of support needed to facilitate performance.

2 Further research with problem analogies

Problem analogies were rejected as a suitable task for the research reported in this thesis as they were not sufficiently well-specified to directly investigate the hypothesis that the way in which children represent relational structure will significantly affect their performance. However, now that we have evidence that the role structural knowledge is indeed important, we can return to the domain of problem analogies and begin to investigate what factors impinge on the correct structural representation of the problem. Some candidates could be irrelevant details given as part of the base problem, the order in which information is presented and the use of familiar relationships or objects in new or unfamiliar contexts. These could encourage inappropriate mappings from the child's current knowledge to the target domain.. Further investigation of these will add to our knowledge of the extent to which young children can apply the relational similarity constraint.

3. Using insights from this thesis to look at other types of analogical reasoning.

One of the original aims of this research was to look at the development of structural representations for the purposes of analogical reasoning in educational contexts. Because of the results obtained from the initial experiments, further studies concentrated on the domain of series problems and so there has been no opportunity to look at other areas of analogical reasoning. However, now that some factors have been isolated which affect children's performance, the next phase of research can go on to determine their effects on the use of analogy for problem solving in educational settings.

We have shown that children are unable to correctly construct structural task representations when the relational information is not made sufficiently salient, and that the resultant incorrect task representations adversely affect analogical mappings. It has also been claimed that children are more successful when reasoning with some evaluative relations if they use an ordering schema to which the relationship is spatially tied and that generalised mapping schemas might well be reconstructed using information from specific tasks. These findings could explain some of the lack of success reported in the literature when analogies are used for teaching purposes.

Concrete analogies are widely used in the teaching of fractions They are commonly compared either to pieces of a cake or pie ('sub-area of a whole' analogy) or to separate items in a collection, for example 4 pieces of an 8-piece bar of chocolate ('a subset-set' analogy). Hart (1980) found that 30% to 40% of 11 year old children were unable to answer simple problems based on an understanding of fractions. A typical question would be "I have eaten 3 pieces of a 9-piece bar of chocolate. What fraction of the chocolate have I eaten?" Goswami, 1992, has emphasised the importance of beginning an early age, by using familiar concepts, when we teach fractions by using concrete analogies. This should ensure that the analogy is transparent to the child and that they learn procedures about the concrete materials at the same time as they learn about written representations. This should prevent the embedding of parallel but separate strands of knowledge in either the concrete materials or in the written representations.

However, an interesting extension to the research reported here would be to make the necessary relationships between the task items in the two domains salient, thus supporting the construction of structural task representations. This could be done by requiring the children to explicitly represent the same fraction in different ways, either diagrammatically or in a written form, and by presenting several different types of task materials in parallel, i.e. pieces of cake, groups of objects and groups of people. The children could then be asked to pick out the common relationships in the three domains, so that they could represent all three domains in a common structure, using either diagrammatic or written representations.

It has previously been suggested (Gelman, 1991) that children incorrectly focus on the number of pieces of cake when they are thinking about the 'sub-area of a whole' analogy. This means that they reach the conclusion that 1/4 is bigger than 1/3, because 4 is bigger than 3. Goswami suggests that this is because they are thinking about the cake being shared out between a number of recipients, rather than considering the size of the piece of cake which each one will receive. This seems intuitively plausible. However, a different but connected reason could be that the children are not sufficiently aware of what the denominator of the fraction represents. If this so, then perhaps they are importing an incorrect simple counting pattern over from their prior arithmetical experience, just as the children studied in this thesis imported an incorrect simple ordering strategy. It could be that we need to make the children aware of the inappropriateness of this by scaffolding the task in an appropriate way, just as we did for the 'incongruent' series problems. This could be done by explicitly focusing on what happens to the written representation as areas or sets are divided into smaller or larger fractions, and the reason for this relationship.

Another area of primary mathematics where young children commonly experience difficulty is arithmetic word problems (Halpern, 1992). These are formal problems which are based on situations found in the real world, and which can be solved by the application of an arithmetic algorithm. This difficulty is not confined to young children, and indeed extends up to undergraduate level and beyond (Steen, 1987). Most of the research in this area has been carried out with secondary school children and college students, however. For example, Schoenfeld (1985), taught college students five strategies to use in mathematical problem solving. These were:

1. Draw a diagram

2. Consider a similar problem

3. Try to establish sub-goals.

4. Look for patterns

5. Use contradiction.

It was found that students who were taught these strategies were performing significantly better after five training sessions than those who did not. It would be interesting to ascertain whether this facilitation in performance could be replicated in primary aged children. However, if we consider the five strategies, at least three of them (numbers 1, 2 and 4) are concerned with the representation of tasks at the structural level and the mapping of similar structures from base to target problems.

We now have evidence from this thesis that these processes are often difficult for young children, but that they can be successful, given appropriate support. It seems that if we encourage children to focus on the common structural patterns which exist between problem isomorphs, they will be able to perform the appropriate analogical mappings. This could be facilitated by presenting several isomorphic word problems in parallel, and requiring the children to complete structural representations depicting the common relationships, by using the appropriate task items. This would highlight the fact that different surface features can still share common deep (i.e. structural) features, and would also scaffold the children's learning by requiring them to consider the structural relationships inherent in the task. Once the importance of these had been learnt, we could present more problems but without cueing the construction of a structural task representation, to ensure that the children were able to build one without needing any support. This type of experiment would also enable us to directly test the claim that the representation of a task at the structural level facilitates analogical mapping. The performance of different groups of children on the same set of isomorphic problems, but with or without the provision of an appropriate structural representation could be compared.

10.8 GENERAL SUMMARY

This thesis has demonstrated the importance of structural task representations for the purposes of analogical mapping. In doing so, it has taken advantage of a large body of literature which exists concerning series problems, but has used these types of problems for a very different purpose. We have also shown that features external to the individual problem-solver, that is, those inherent in the task, can have a significant effect on performance, perhaps by affecting the reconstruction of generalised schemas. It is suggested that these findings could affect the use of analogy in educational settings.

REFERENCES

Adams, M. J. (1978). Logical competence and transitive inference in young children. Journal of Experimental Child Psychology, 25, 477-489.

Anderson, A., Tolmie, A., Howe, G., Hayes, J. T. C., & MacKenzie, M. (1992). Mental models of motion. In Rodgers Y., Rutherford A , & Bibby P.(Eds.), Models in the Mind London: Academic Press.

Bibby, P. A., & Payne, S. J. (1993). Internalization and the use specificity of device knowledge. Human Computer Interation, 8, 25-56.

Breslow, L. (1981). Re-evaluation of the literature on the development of transitive inferences. Psychology Bulletin, 89, 325-351.

Brown, A. L. (1989). Analogical learning and transfer: What develops? In Vosniadou S. & Ortony A. (Eds.), Similarity and Analogical Reasoning Cambridge, UK: Cambridge University Press.

Brown, A. L., & Kane, M. J. (1988). Pre-School children can learn to transfer: Learning to learn and learning from people. Cognitive Psychology, 20, 493-523.

Brown, A. L., Kane, M. J., & Echols, C. H. (1992). Young children's mental models determine analogical transfer across problems with a common goal structure. Cognitive Development, 1, 103-121.

Brown, A. L., Kane, M. J., & Long, C. (1989). Analogical transfer in young children: Analogies as tools for communication and exposition. Applied Cognitive Psychology, 3, 275-293.

Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1985). Mathematics in the streets and in schools. British Journal of Developmental Psychology, 3, 21-29.

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

Catherall, E. (1984). Investigating Fractions. Hove, U.K.: Wayland Ltd.

Catrambone, R., & Holyoak, K. J. (1985). The function of schemas in analogical problem-solving. Poster presented at the meeting of the American Psychological Association, Los Angeles, California.

Chen, Z., & Daehler, M. W. (1989). Positive and negative transfer in analogical problem solving. Cognitive Development, 4, 327-344.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. Cognitive Psychology, 17, 391-416.

Chi, M. T. H., Feltovich, P. J., & Glaser, P. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Collins, A., & Burstein, M. (1989). A framework for a theory of comparison and mapping. In Vosniadou S. and Ortony A. (Eds.), Similarity and Analogical Reasoning Cambridge, UK: Cambridge University Press.

Crisafi, M. A., & Brown, A. L. (1986). Analogical Transfer in very young children: Combining two separately learned solutions to reach a goal. <u>Child Development, 57,</u> 953-968.

DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. <u>Child Deveopment, 62,</u> 736-752.

DeSoto, C. B., London, M., & Handel, S. (1965). Social reasoning and spatial paralogic. <u>Journal of Personality and Social Psychology, 2,</u> 13-521.

Donaldson, M. (1978). <u>Childrens Minds</u>. London: Fontana.

Gallagher, J. M., & Wright, R. J. (1977). Children's solution of verbal analogies: Extension of Piaget's concept of reflexive abstraction. Paper presented to the Society for Research in Child Development, New Orleans.

Gallagher, J. M., & Wright, R. J. (1979). Piaget and the study of analogy: structural analysis of items. In Megary J. (Ed.), <u>Piaget and the helping professions</u> <u>Vol 8</u> (pp. 114-119). Los Angeles, USA: University of Southern California.

Gelmen, R. (1981). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In Carey, S. & Gelman, R. (Eds.), Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Gentner, D. (1983). Structure mapping : a theoretical framework for analogy. <u>Cognitive Science, 7,</u> 155-170.

Gentner, D. (1988). Metaphor as structure mapping - the relational shift. <u>Child Development, 59,</u> 47-59.

Gentner, D. (1989). The mechanisms of analogical learning. In Vosniadou S. & Ortony A. (Eds.), <u>Similarity and Analogical Reasoning</u> Cambridge, U.K.: Cambridge University Press.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: mental models of electricity. In Gentner D. and Stevens A. L. (Eds.), <u>Mental Models</u> Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Gentner, D., & Toupin, C. (1986). Systemacity and surface similarity in the development of analogy. <u>Cognitive Science, 10,</u> 277-300.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. <u>Cognitive Psychology, 12,</u> 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. <u>Cognitive Psychology, 15,</u> 1-38.

Goswami, U. (1989). Relational complexity and the development of analogical reasoning. <u>Cognitive Development, 4,</u> 251-268.

Goswami, U. (1991). Analogical Reasoning: What develops? A review of research and theory. <u>Child Development, 62,</u> 1-22.

Goswami, U. (1992a). <u>Analogical Reasoning in Children</u>. Hove, U.K.: Lawrence Erlbaum Associates.

Goswami, U. (1992b). Commmentary: Analogical reasoning and conceptual complexity in cognitive development. <u>Human development, 35,</u> 193-217.

Goswami, U., & Brown, A. L. (1989). Melting chocolate and melting snowmen: analogical reasoning and causal relations. Cognition, 35, 60-95.

Goswami, U., & Brown, A. L. (1990). Higher order structure and relational reasoning: Contrasting analogical and thematic relations. Cognition, 36, 207-226.

Halasz, K. G., & Moran, T. P. (1983). Mental Models and problem solving using a calculator. In Proceedings of CHI '83:Human Factors in Computing systems New York : ACM

Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? Cognitive Psychology, 16, 65-93.

Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. Human Development, 35, 193-217.

Halford, G. S., & Kelly, M. E. (1984). On the basis of early transitivity judgements. Journal of Experimental Child Psychology, 38, 42-63.

Halpern, D. F. (Ed.) (1992) Enhancing thinking skills in the sciences and mathematics Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Handel, S., DeSoto, C. B., & London, M. (1968). Reasoning and spatial representations. Journal of Verbal Learning and Verbal Behaviour, 7, 351-357.

Hart, K. M. (1980). Secondary school children's understanding of mathematics. Research Monograph. Chelsea College, University of London

Hayes, J. N., & Simon, H. A. (1974). Understanding written problem instructions. In Gregg L. W. (Ed.), Knowledge & Cognition Hillsdale, New Jersey: Erlbaum.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). Induction - processes of inference, learning and discovery. Cambridge, Mass.: MIT Press.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. Psychology of Learning and Motivation Vol. 19 New York: Academic Press.

Holyoak, K. J., Junn, E. N.& Billman D. O.(1984). Development of analogical problem-solving skill. Child Development, 55, 2042-2055.

Hunter, I. (1957). The solving of three-term series problems. British Journal of Psychology, 48, 286-298.

Inhelder, B., & Piaget, J. (1958). The growth of logical thinking from childhood to adolescence (Parsons, A. & Pilgram, S., Trans.). New York: Basic Books.

Inhelder, B., & Piaget, J. (1964). The early growth of logic in the child (Larger, E. & Papert, D., Trans.). London: Routledge & Kegan-Paul.

Johnson, M. (1987). The Body in the Mind: The Bodily Basis of Meaning, Imagination and Reason. Chicago: University of Chicago Press.

Jones, S. (1970). Visual and verbal processes in problem solving. Cognitive Psychology, 1, 201-214.

Kallio, K. D. (1982). Developmental change of a five-term transitive inference. Journal of Experimental Child Psychology, 33, 142-164.

Keane, M. T., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: a comparison of the three models. Cognitive Science, 18, 387-438.

Keane, M. T. (1990). Incremental analogizing: theory and model. In Gilhooly K. H., Keane M. T., Logie, R. and Erdos, G. (Eds.), Lines of thinking: Reflections on the psychology of thought New York.: Wiley.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. Cognitive Science, 8, 255-273.

Levinson, P. J., & Carpenter, R. L. (1974). An analysis of analogical reasoning in children. Child Development, 45, 857-861.

Lunzer, E. A. (1965). Problems of formal reasoning in test situations. In Mussen P. H. (Ed.), Monographs of the Society for Research in Child Development. Vol. 30 (pp. 19-46). Society for Research in Child Development.

Markham, E., & Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning : Taxonomic versus thematic relations. Cognitive Psychology, 16, 1-27.

Marr, D. (1982). Vision. New York: Freeman.

Novick, L. R. (1988). Analogical transfer, problem similarity and expertise. Journal of Experimental Psychology: Learning, memory and cognition, 14, 398-415.

O'Connor, J., Beilin, H., & Kose, G. (1981). Children's belief in photographic fidelity. Developmental Psychology, 17, 859-865.

Ortony, A. (1979). Beyond literal similarity. Psychological Review, 87, 161-180.

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

Pears, R., & Bryant, P. (1990). Transitive Inferences by young children about spatial position. British Journal of Psychology, 81, 497-510.

Piaget, J. (1921). Une forme verbale de la comparison chez l'enfant. Archives de Psychologie, 18, 141-172.

Piaget, J. (1970). Genetic Epistemology (Duckworth, E., Trans.). New York: Columbia University Press.

Piaget, J. (1972). Essai de logique opératoire. Paris: Dunod.

Piaget, J., Montangero, J., & Billeter, J. (1977). Les correlats. In Piaget J. (Ed.), L'Abstraction Reflechissante Paris: Presses.

Polya, G. (1957). How to solve it. Garden City, New York: Doubleday Anchor.

Riley, C. A. (1976). The Representation of comparitive relations and the transitive inference task. Journal of Experimental Child Psychology, 22, 1-22.

Rogoff, B. (1990). Apprenticeship in Thinking : Cognitive Development in Social Context. New York: Oxford University Press.

Samburski, S. (1973). Physics of the stoics. Westport, Connecticut: Greenwood Press.

Schoenfield, A. H. (1987). What's all the fuss about meta-cognition? In Schoenfield, A. H. (Ed.) Cognitive science and mathematics education. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Smedslund, J. (1963). Development of concrete transitivity of length in children. Child Development, 34, 389-405.

Spearman, C. (1923). The Nature of 'Intelligence' and the Principles of Cognition. London, UK: Macmillan.

Steen, L. A. (1987) Mathematics education: a predictor of scientific competence. Science, 237, 251-252

Sternberg, R. J. (1977). Intelligence, Information Processing and Analogical Reasoning : The Componential Analysis of Human Abilities. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Sternberg, R. J., & Nigro, G. (1980). Developmental patterns in the solution of verbal analogies. Child Development, 51, 353-378.

Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. Journal of Experimental Child Psychology, 27, 195-232.

Trabasso, T. (1975). Representation, memory and reasoning: How do we make transitive inferences? In Pick A. D.(Ed.), Minnesota Symposia on Child Psychology, Vol. 9 . Minneapolis: University of Minnesota Press.

Trabasso, T., Riley, C., & Wilson, E. (1975). The representation of linear & spatial strategies in reasoning: A developmental study. In Falmagne R. J. (Ed.), Reasoning: Representation & Process in Children and Adults Hillsdale, New Jersey: Erlbaum.

Trabasso, T. C., & Riley, C. (1975). On the construction and use of representations involving linear order. In Solso R. L. (Ed.), Information Processing and Cognition: The Loyola Symposium Hillsdale, New Jersey: Erlbaum.

Vosniadou, S. (1989). Analogical reasoning as a mechanism in knowledge acquisition : A developmental perspective. In Vosniadou S.& Ortony A.(Eds.), Similarity and Analogical Reasoning Cambridge, U.K.: Cambridge University Press.

Vosniadou, S., & Ortony, A. (1989). Similarity & analogical reasoning : a synthesis. In Vosniadou S.& Ortony A.(Eds.), Similarity & Analogical Reasoning Cambridge, U.K.: Cambridge University Press.

Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). 'Metaphorical' mapping in human infants. Child Development, 52, 728-731.

Walkerdine, V., & Sinha, C. (1975). Object properties, context rules and cognitive structure: the three terms of the three term series. Paper presented at the British Psychological Society Developmental Section Conference, Cambridge, UK.

Wechsler, D. (1976). Wechsler Intelligence Scale for Children - Revised. Windsor, U.K.: Nelson.

Willatts, P. (1991). Infant planning: What develops early, and what develops late? Paper presented at the British Psychological Society Developmental Section Conference, Cambridge, UK.

Winston P. H. (1980). Learning and reasoning by analogy. Communications of the ACM, 23, 689-703.

Wood, D. J. (1969). The Nature and Development of Problem Solving Strategies. Unpublished Ph.D. Thesis, University of Nottingham.

Wood, D. J. (1988). How Children Think and Learn. Oxford, U.K.: Blackwells.

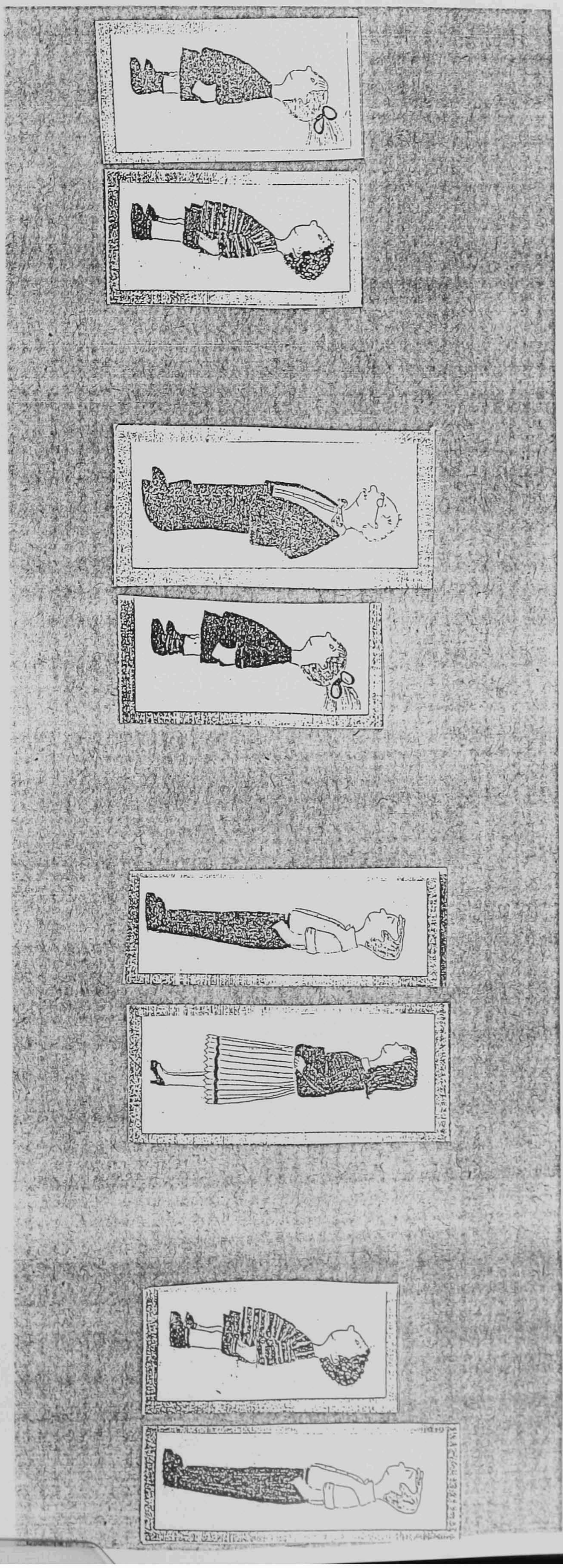
Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17, 89-100.

LIST OF APPENDICES

APPENDIX A Task materials : 5 person queue

Used in Experiment 1

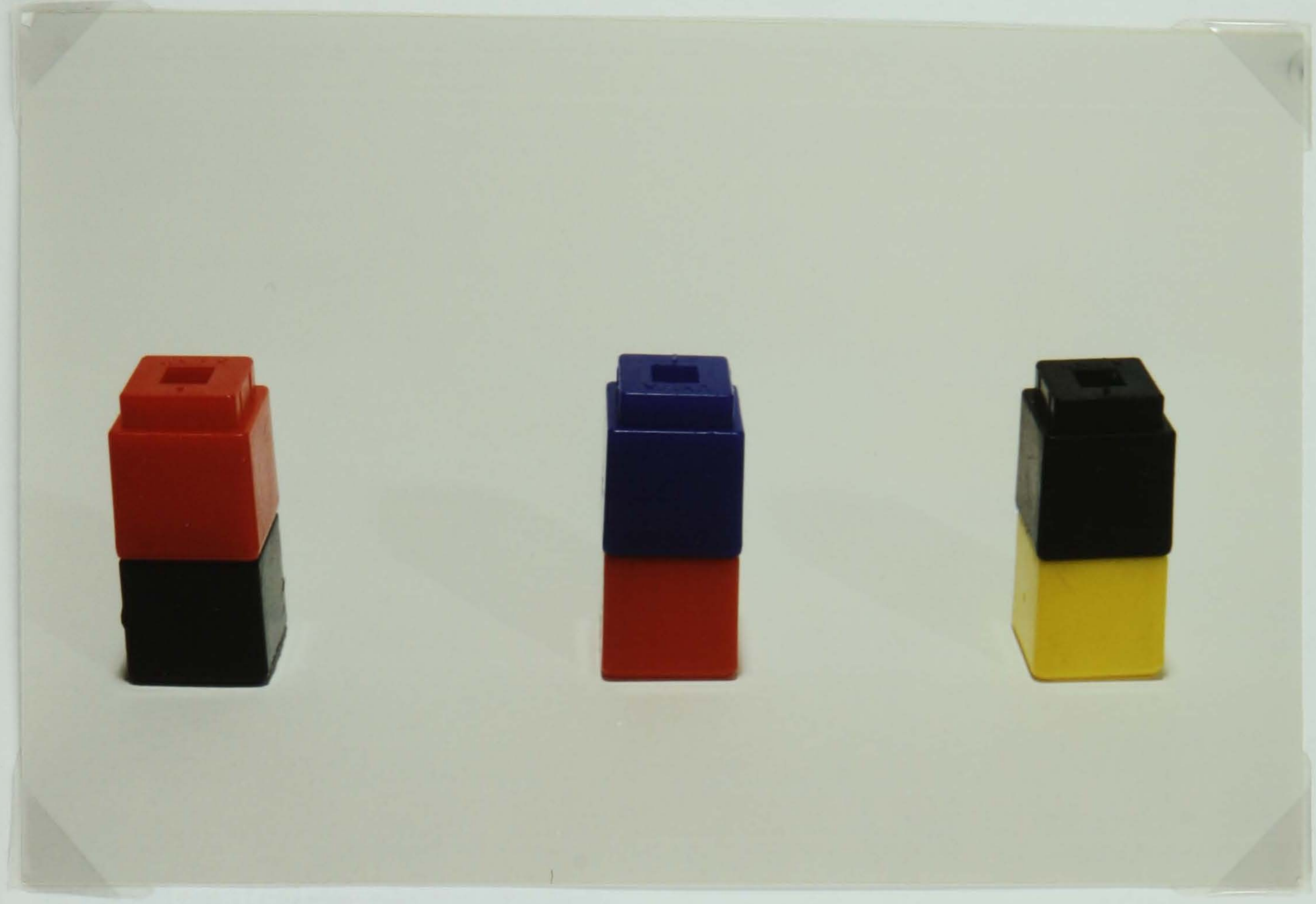
Example premise presentation (original materials, which were in colour, have been photocopied and reduced by 76%)



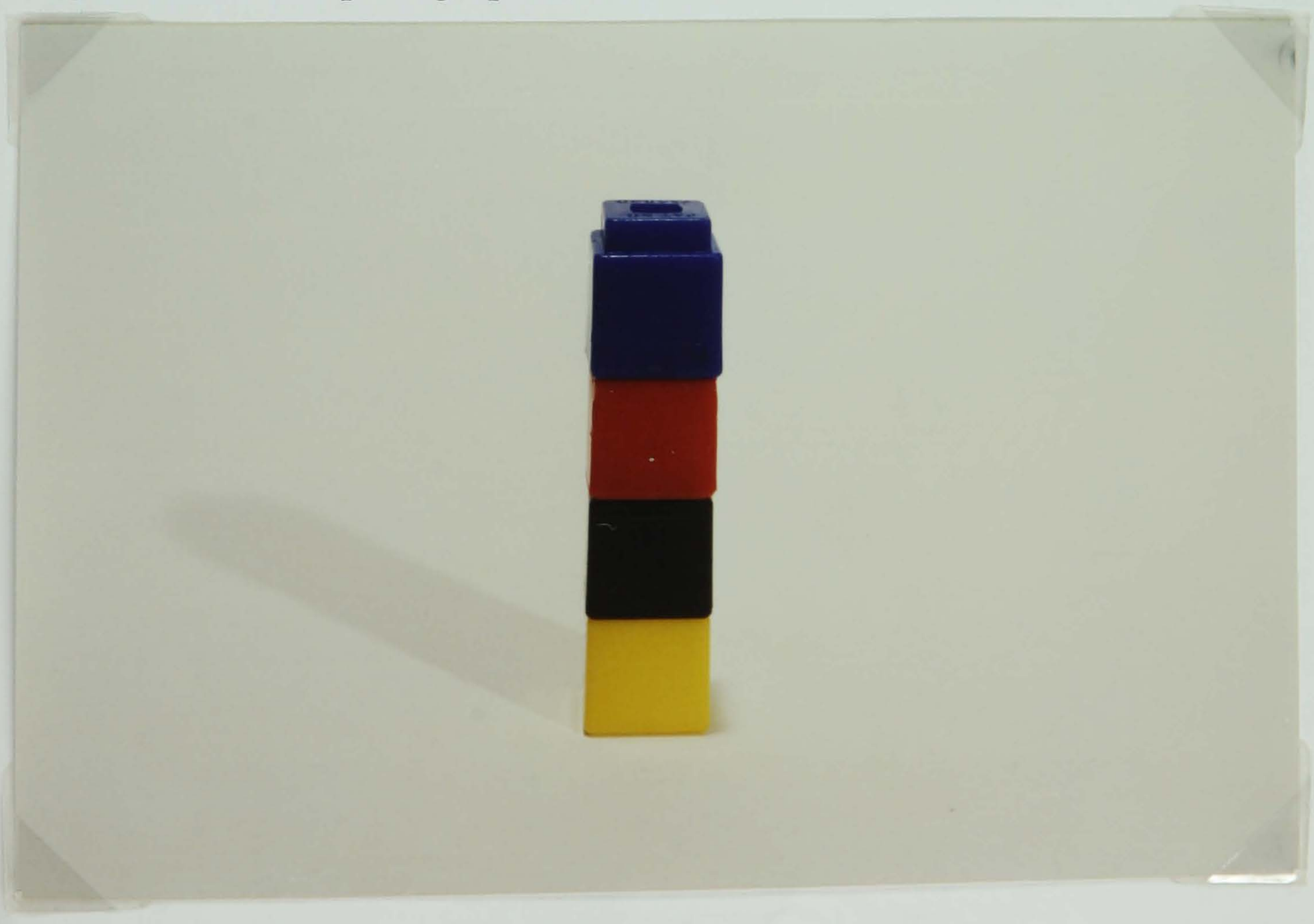
APPENDIX B Task materials : Towers of bricks

Used in Experiment 2

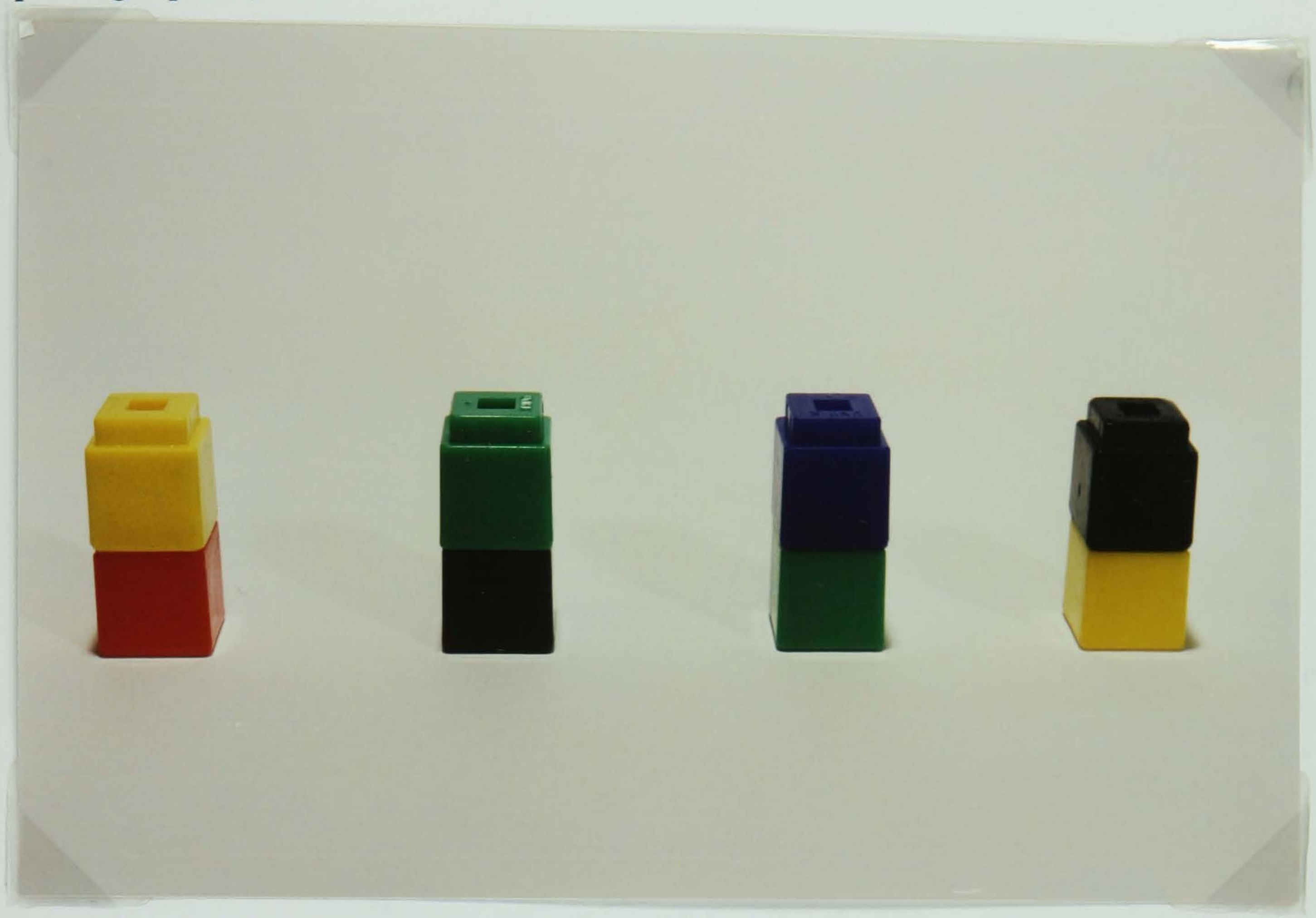
Tower of 4 : Example premise presentation (original materials photographed)



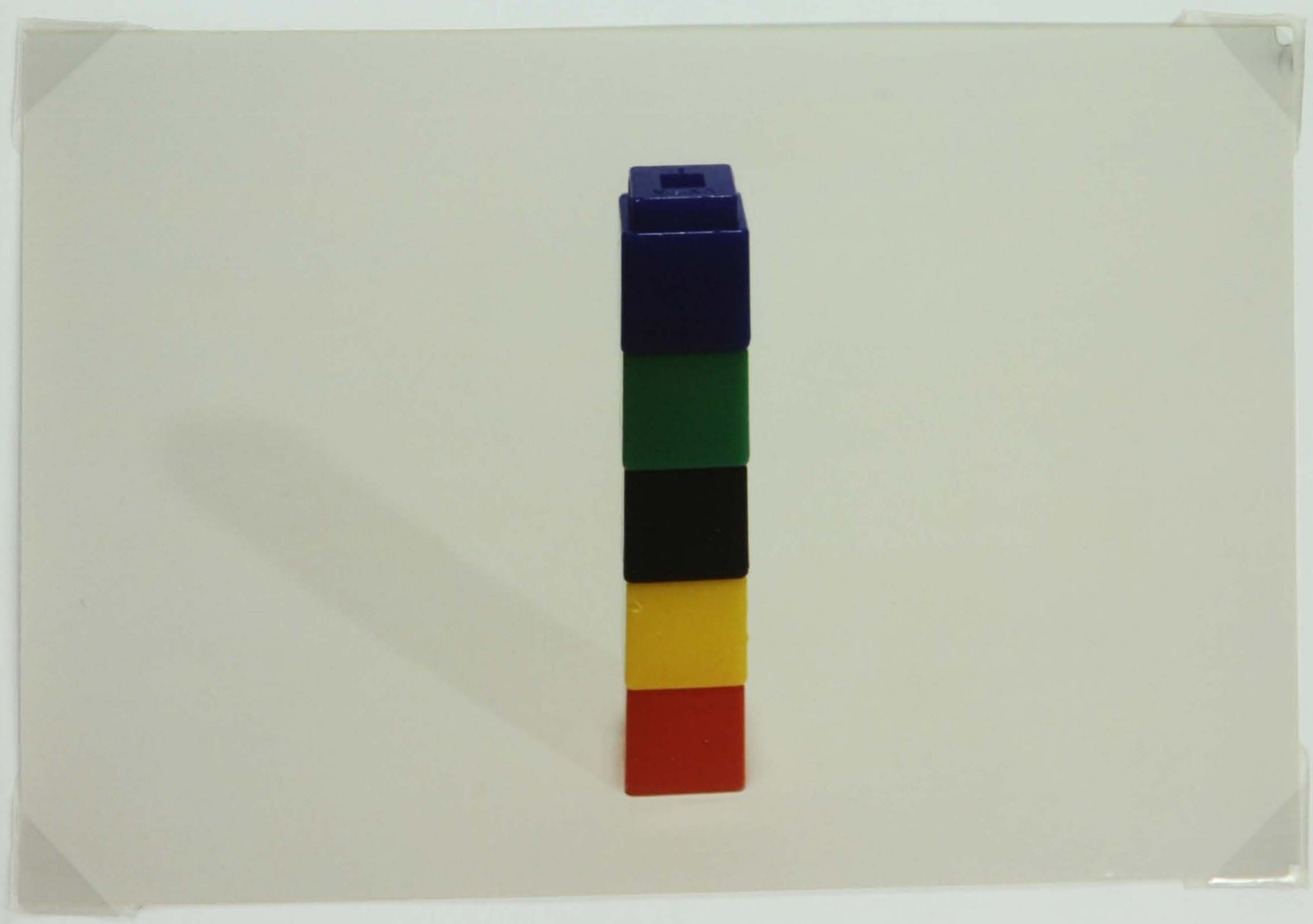
Tower of 4 : Completed array using premise information as above (original materials photographed)



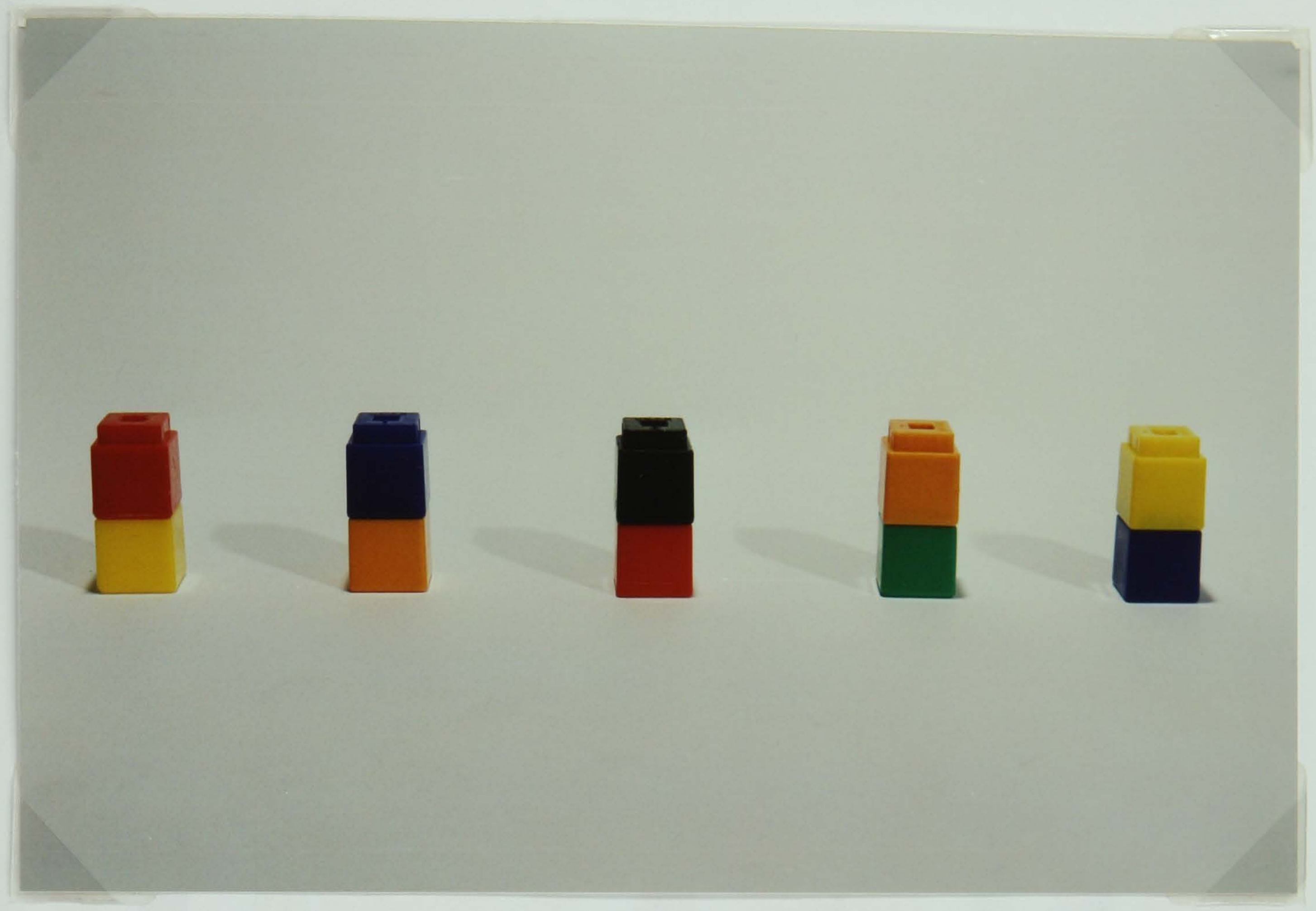
Tower of 5 : Example premise presentation (original materials photographed)



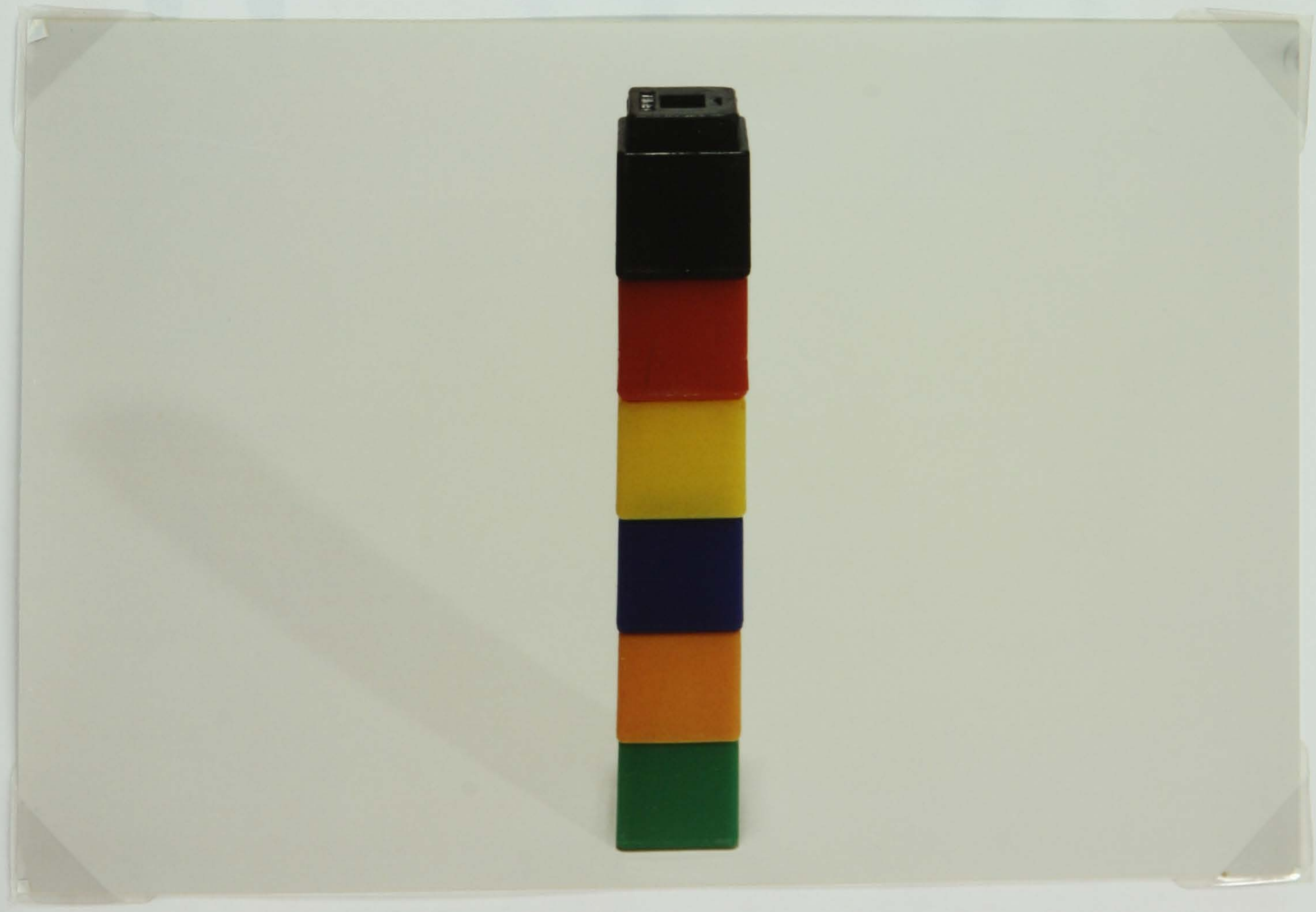
Tower of 5 : Completed array using premise information as above (original materials photographed)



Tower of 6 : Example premise presentation (original materials photographed)



Tower of 6 : Completed array using premise information as above (original materials photographed)



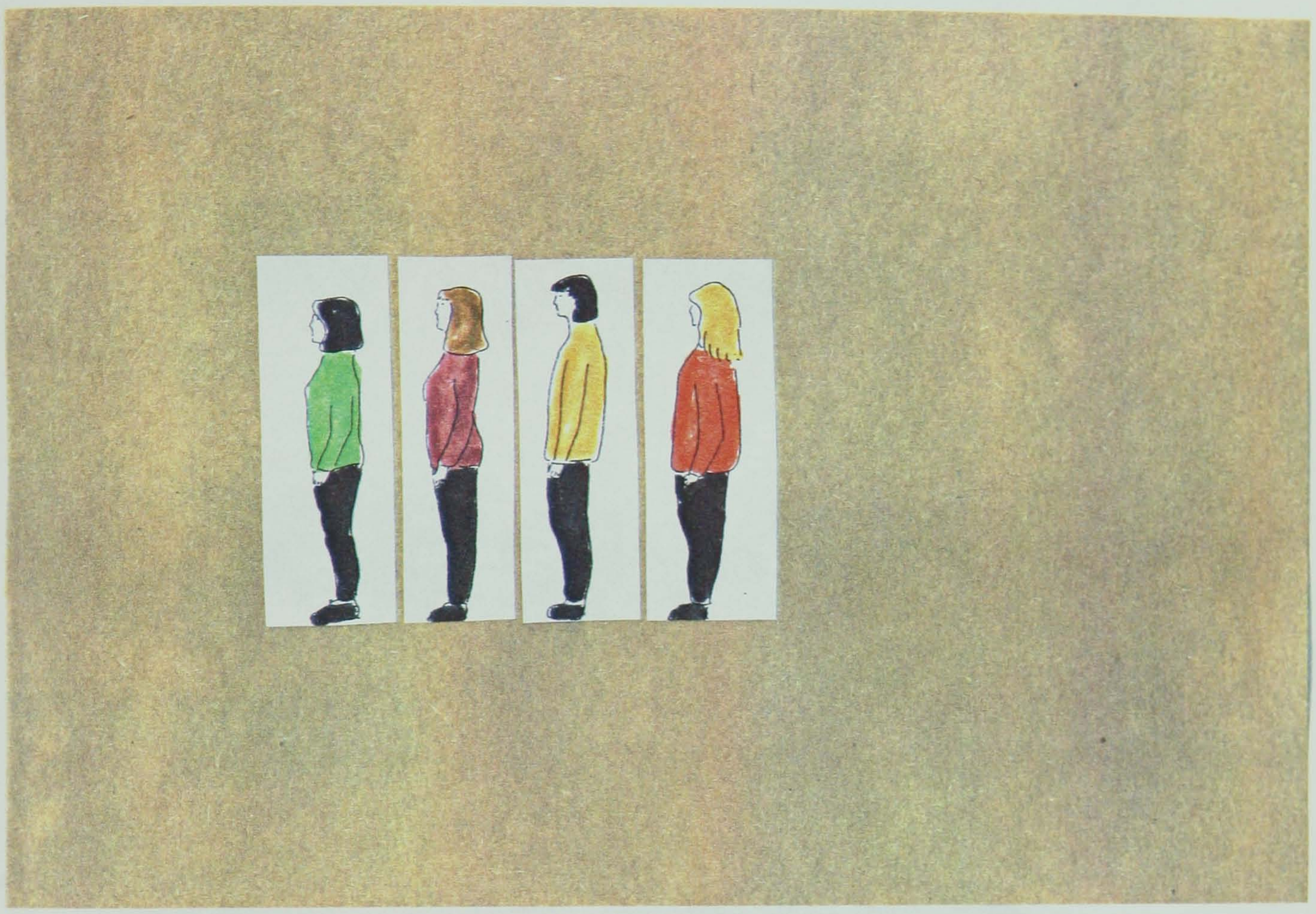
APPENDIX C Task materials : 4 person queue (drawings)

Used with horizontal premise presentation in Experiments 3, 4, 7, 8, 10 (with only 3 items) and 11.

Used with vertical premise presentation in Experiments 8, 10 (with only 3 items), 11 and 12.

Example horizontal premise presentation (original materials photocopied and reduced by 50%)

Completed array using premise information as above (original materials photocopied and reduced by 50%)



APPENDIX D   Task materials : 4 person queue (photographs)
Used in Experiments 3, 5 and 6.
Example horizontal premise presentation (original materials)
<u>Note</u> The page should be rotated through 90°. The premises are placed closer together than they were in the experiments. The actual gap was approximately 6 cms

Completed array using premise information as above (original materials)

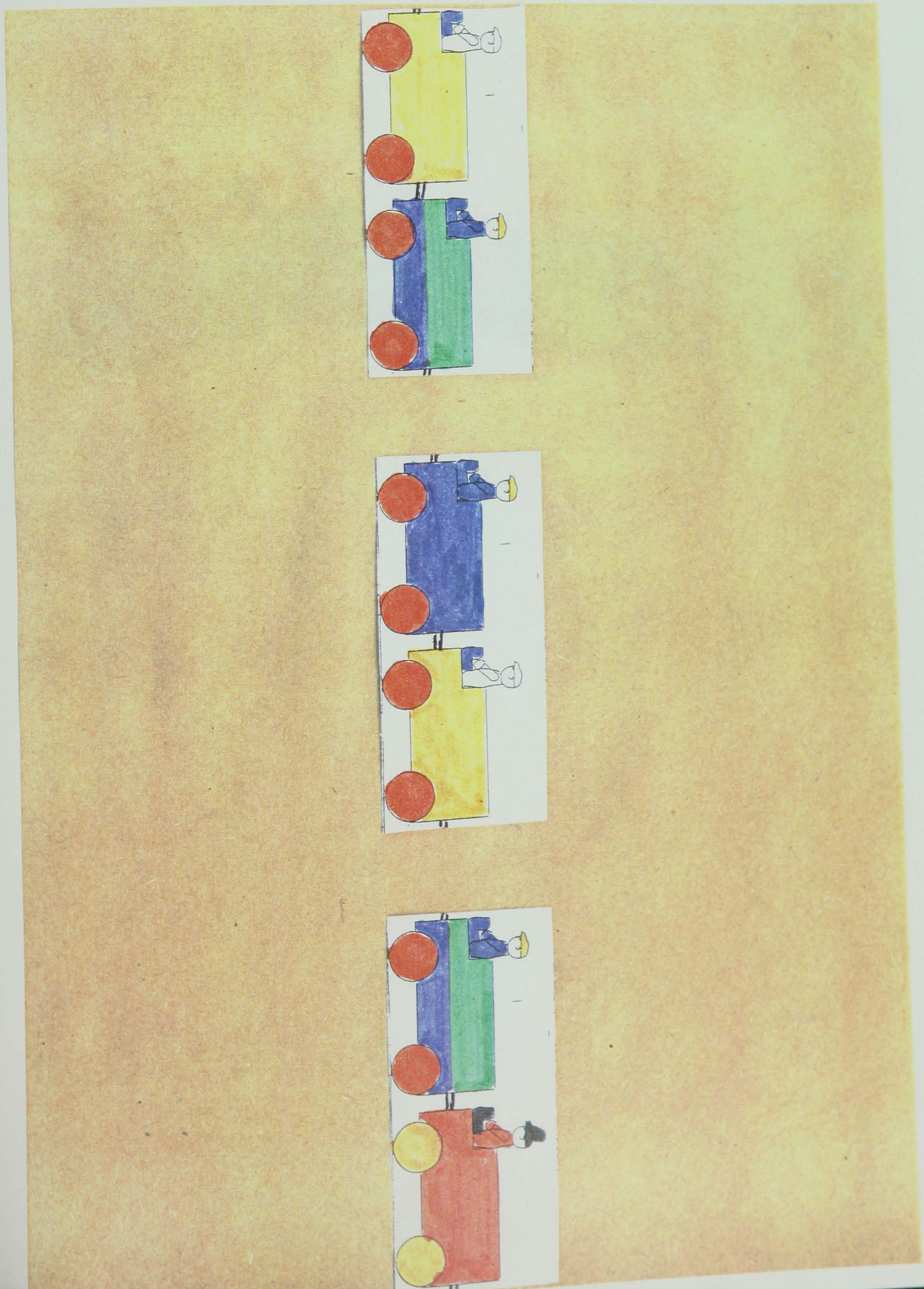
APPENDIX E Task materials : 4 carriage train

Used with horizontal premise presentation in Experiments 4 and 9.

Used with vertical premise presentation in Experiment 9.

Example horizontal premise presentation (original materials photocopied and reduced by 65%)

Note The page should be rotated through 90°.



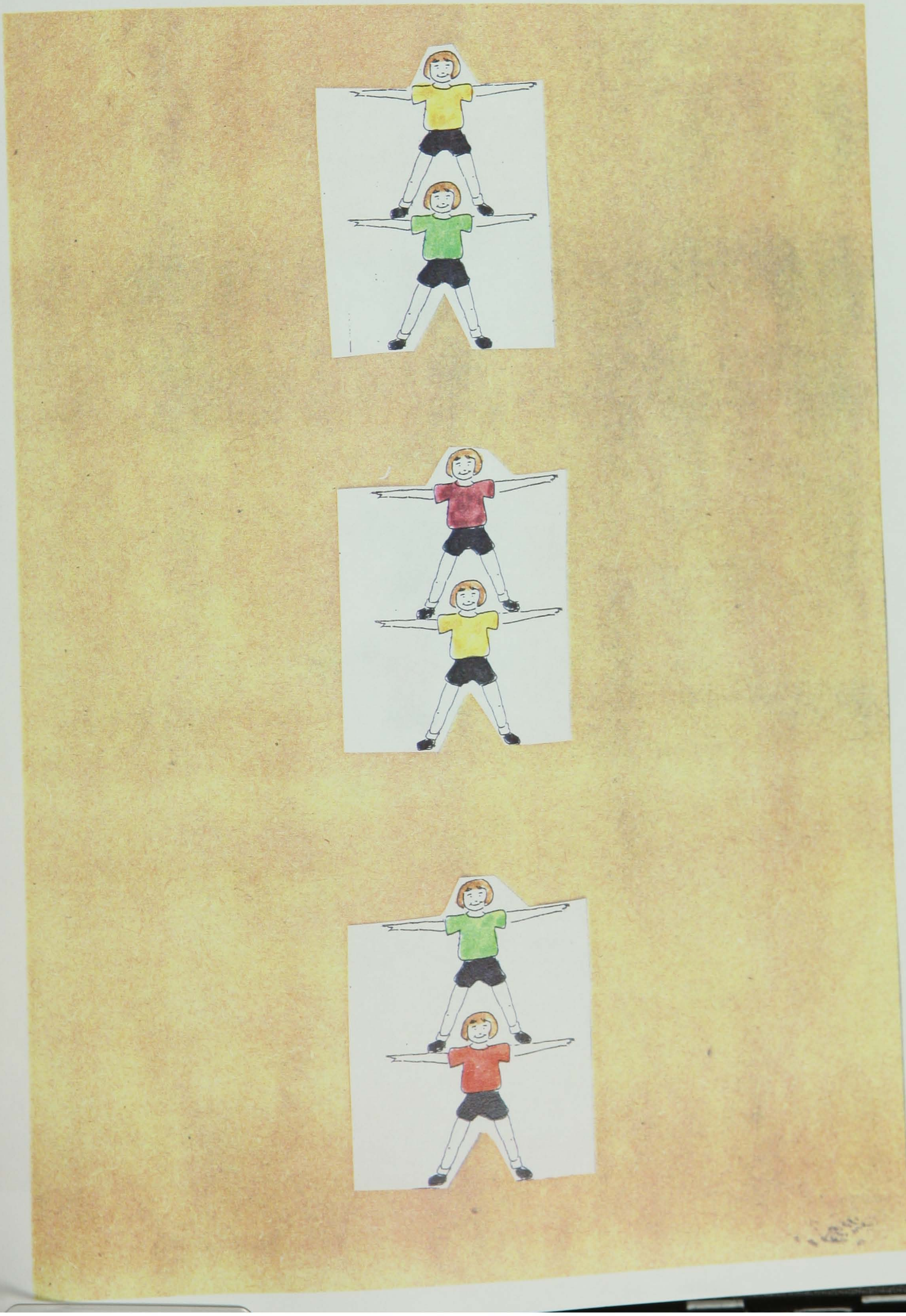
Completed array using premise information as above (original materials photographed)

APPENDIX F Task materials : 4 person human tower

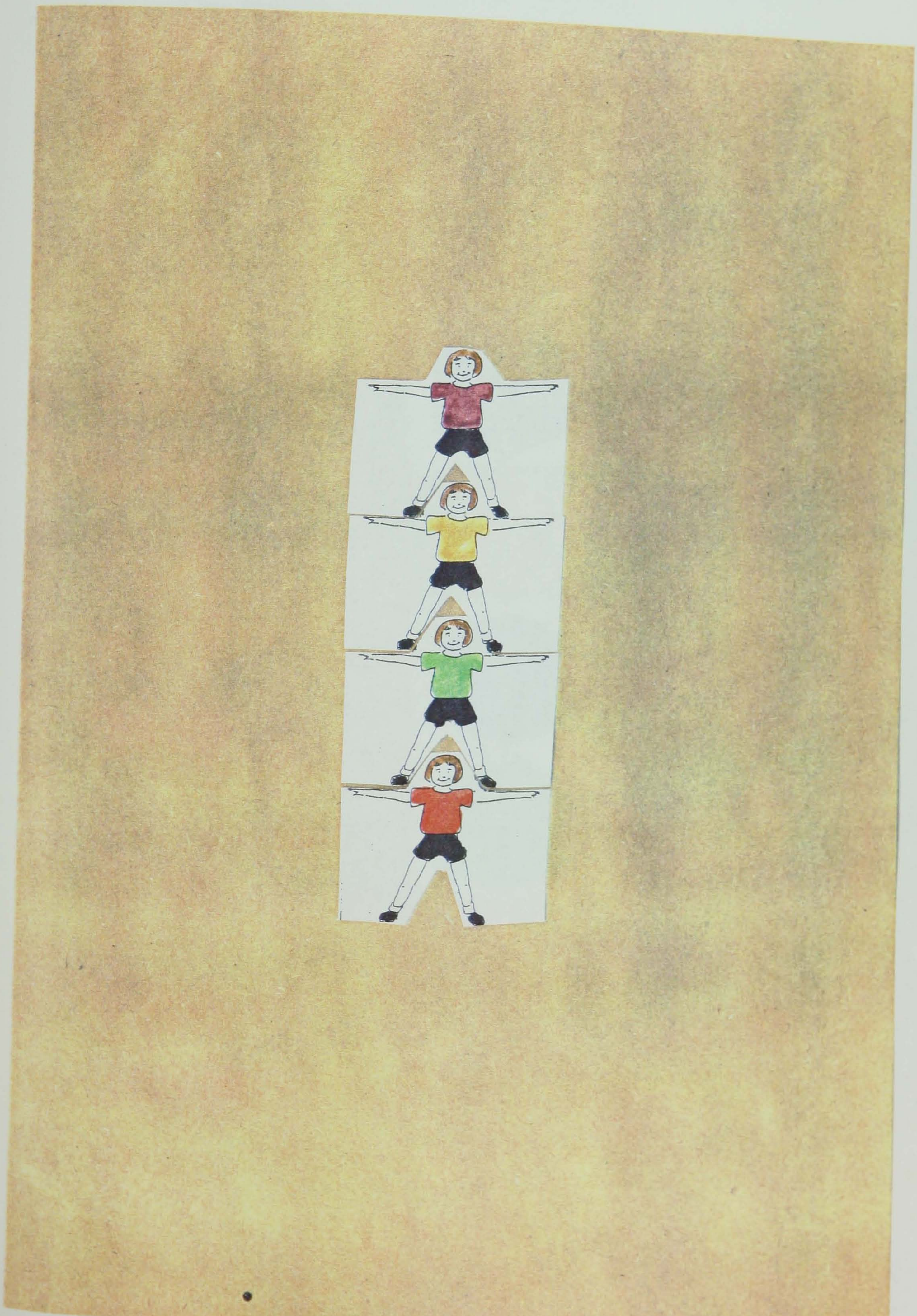
Used with horizontal premise presentation in Experiments 8, 10 (with only 3 items) and 11.

Used with vertical premise presentation in Experiments 8, 10 (with only 3 items), 11 and 12.

Example vertical premise presentation (original materials photocopied and reduced by 65%)



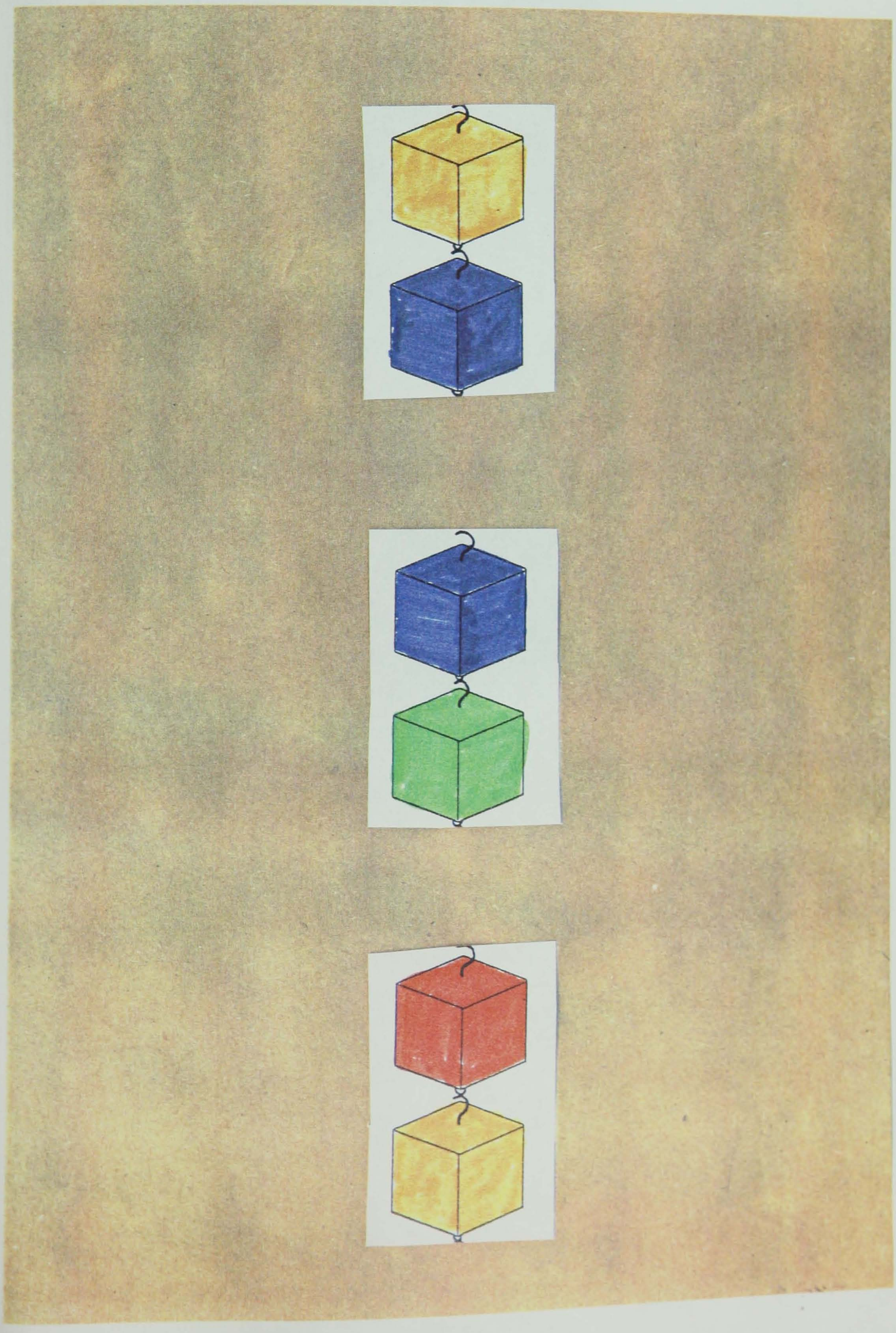
Completed array using premise information as above (original materials photocopied and reduced by 65%)



APPENDIX G Task materials : 4 crate crane

Used with both horizontal and vertical premise presentations in Experiment 9.

Example vertical premise presentation (original materials photocopied and reduced by 50%)

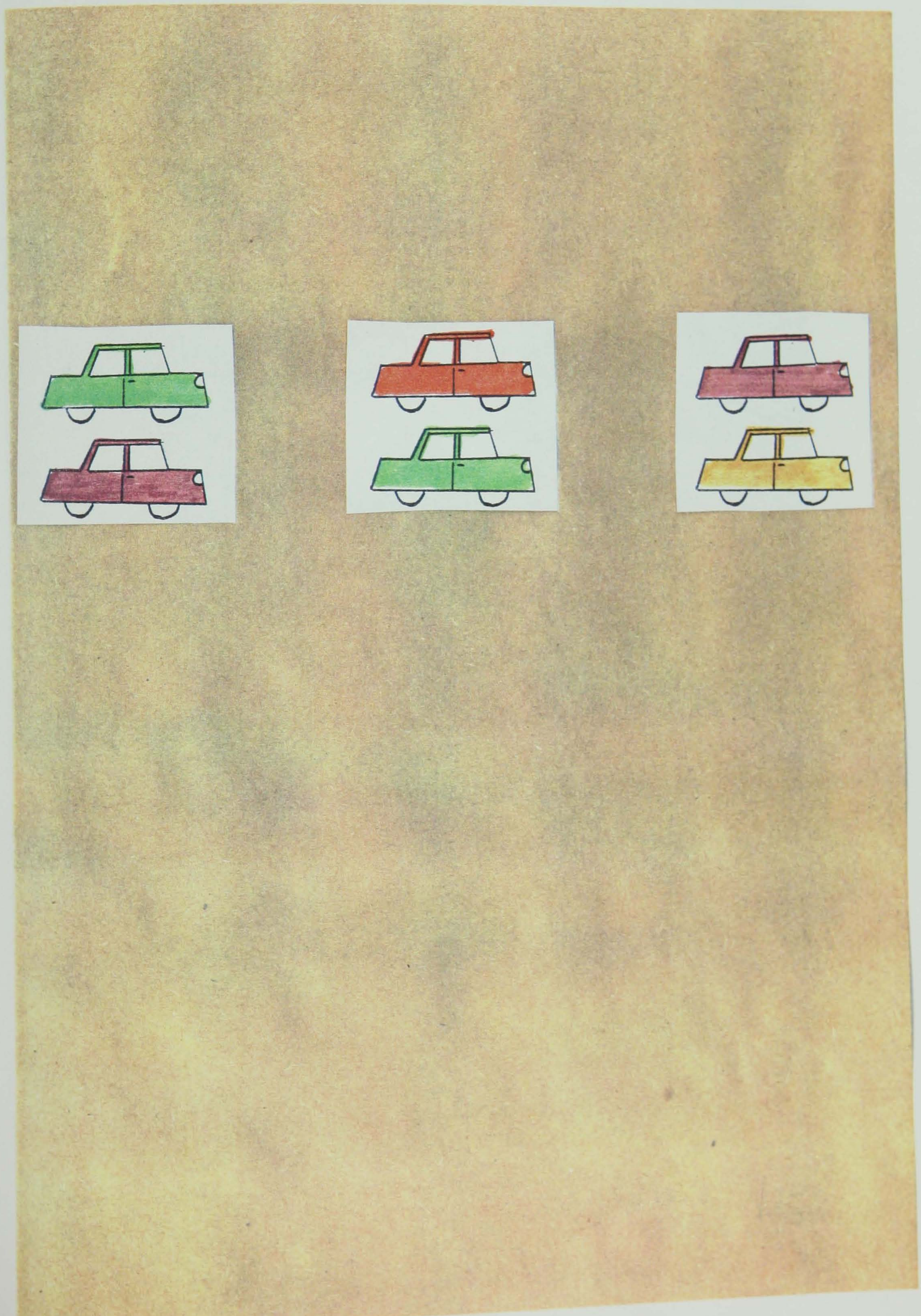


Completed array using premise information as above (original materials photographed)

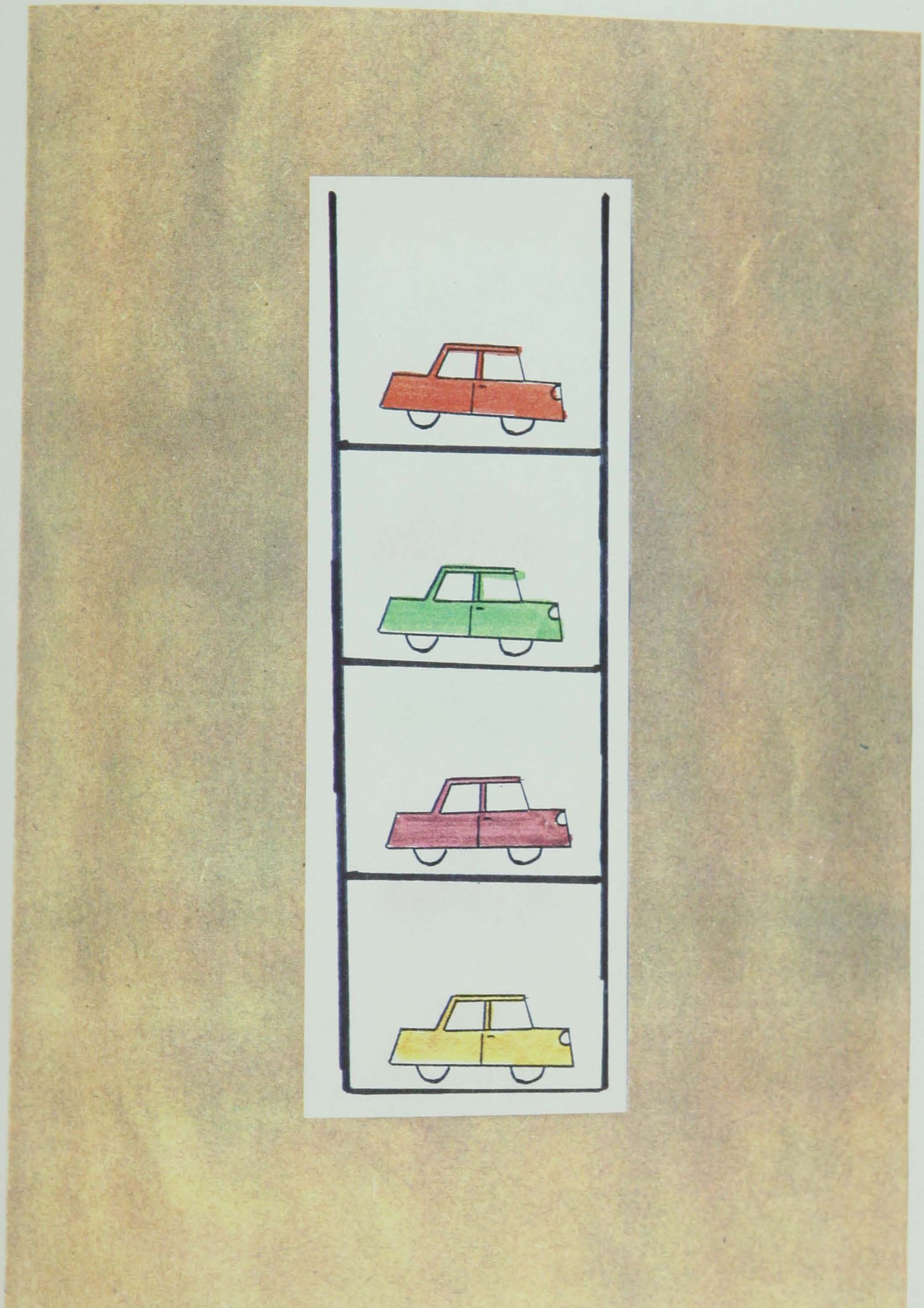
APPENDIX H Task materials : 4 car 'stack'

Used in Experiment 13

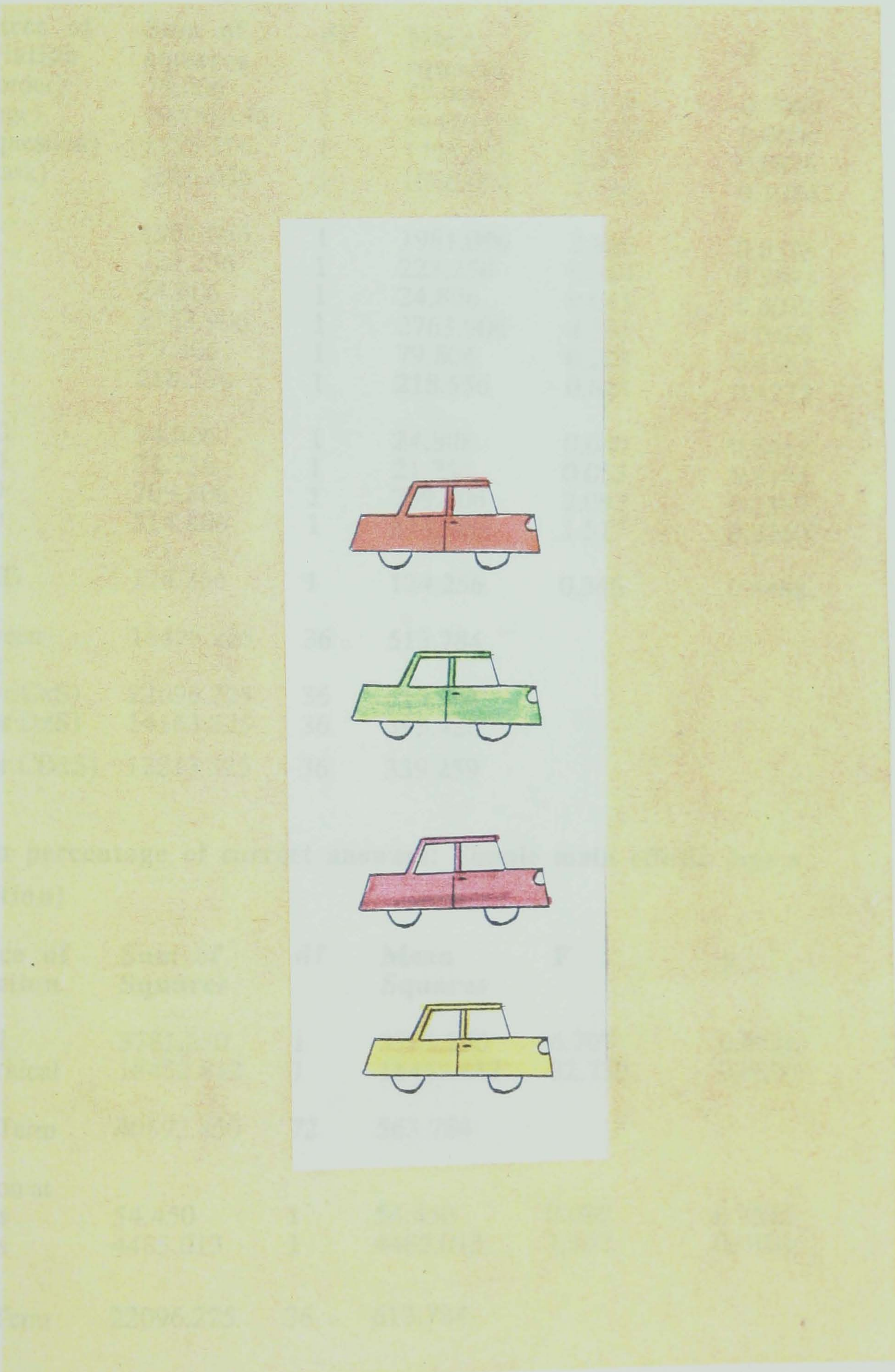
Example premise presentation (original materials photocopied and reduced by 50%)



Completed 'plausible' array using premise information as above (original materials photocopied and reduced by 50%)



Completed 'implausible' array using premise information as above (original materials photocopied and reduced by 50%)



APPENDIX I F tables (Chapter 4)

Experiment 1

Mean percentage of correct answers: 4 factor ANOVA

Source of variation	Sum of squares	df	Mean squares	F	p
A (order)	77.006	1	77.006	0.150	0.7009
B (age)	19470.156	1	19470.156	37.896	0.0000
C (question)	1775.556	1	1775.556	2.893	0.0976
D (task)	1066.056	1	1066.056	2.710	0.1084
AB	1981.056	1	1981.056	3.856	0.0573
AC	223.256	1	223.256	0.364	0.5502
AD	24.806	1	24.806	0.063	0.8032
BC	2763.906	1	2763.906	4.503	0.0408
BD	79.806	1	79.806	0.203	0.6551
CD	218.556	1	218.556	0.644	0.4275
ABC	24.806	1	24.806	0.040	0.8418
ABD	21.756	1	21.756	0.055	0.8154
ACD	709.806	1	709.806	2.092	0.1567
BCD	514.806	1	514.806	1.517	0.2260
ABCD	124.256	1	124.256	0.366	0.5488
Between Error	18496.225	36	513.784		
(Error CxS)	22096.225	36	613.784		
(Error DxS)	14163.325	36	393.426		
(Error CDxS)	12213.325	36	339.259		

Mean percentage of correct answers: Simple main effects (age x question)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
age at					
critical	3781.250	1	3781.250	6.707	0.0116
non-critical	18452.812	1	18452.812	32.730	0.0000
Error Term	40592.450	72	563.784		
question at					
7 years	54.450	1	54.450	0.089	0.7675
9 years	4485.013	1	4485.013	7.307	0.0104
Error Term	22096.225	36	613.784		

Mean number of correctly ordered arrays: 3 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (order)	0.800	1	0.800	1.025	0.3181
B (age)	8.450	1	8.450	10.826	0.0022
C (task)	1.250	1	1.250	3.659	0.0638
AB	0.200	1	0.200	0.256	0.6158
AC	16.200	1	16.200	47.415	0.0000
BC	0.450	1	0.450	1.317	0.2587
ABC	1.800	1	1.800	5.268	0.0277
Between Error	28.100	36	0.781		
(Error CxS)	12.300	36	0.342		

Mean number of correctly ordered arrays: simple simple main effects (order x age x task)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
order at					
7 yrs spatial	6.050	1	6.050	7.751	0.0085
7 yrs abstract	8.450	1	8.450	10.826	0.0022
9 yrs spatial	0.450	1	0.450	0.577	0.4526
9 yrs abstract	4.050	1	4.050	5.189	0.0288
Error Term	28.100	36	0.781		
age at					
spat f spatial	7.200	1	7.200	9.224	0.0044
spat f abstract	0.450	1	0.450	0.577	0.4526
abst f spatial	0.800	1	0.800	1.025	0.3181
abst f abstract	2.450	1	2.450	3.139	0.0849
Error Term	28.100	36	0.781		
task at					
spat f 7 yrs	6.050	1	6.050	17.707	0.0002
spat f 9 yrs	0.200	1	0.200	0.585	0.4492
abst f 7 yrs	8.450	1	8.450	24.732	0.0000
abst f 9 yrs	5.000	1	5.000	14.634	0.0005
Error Term	12.300	36	0.342		

Building the tower with serial re-ordering: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (order)	640.000	1	64.000	1.555	0.2284
B (task)	44.100	1	44.100	0.115	0.7380
AB	16.900	1	16.900	0.044	0.8358
Between Error	7409.000	18	411.611		
(Error BxS)	6876.000	18	382.000		

Experiment 2

Mean number of successful tower completions: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
Subjects	21.644	14	1.546		
A (no. of bricks)	16.178	2	8.089	20.303	0.0000
(Error AxS)	11.156	28	0.398		

APPENDIX J F tables (Chapter 5)

Experiment 3

Mean number of correct answers: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (stimulus)	0.017	1	0.017	0.017	0.8953
B (ordering)	63.333	2	31.667	33.204	0.0000
AB	0.133	2	0.067	0.070	0.9326
Error	51.500	54	0.954		

Data collapsed across type of stimulus:1 factor ANOVA

Source of Variance	Sum of Squares	df	Mean Squares	F	p
A (ordering)	63.333	2	31.667	34.947	0.0000
Error	51.650	57	0.906		

Mean time taken to complete trials: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (stimulus)	50.380	1	50.380	3.591	0.0635
B (ordering)	3679.875	2	1839.938	131.135	0.0000
AB	2.624	2	1.312	0.093	0.9109
Error	757.667	54	14.031		

Data collapsed across type of stimulus: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	3679.875	2	1839.938	129.370	0.0000
Error	810.671	57	14.222		

Experiment 4

Mean number of correct answers: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (stimulus)	0.200	1	0.200	0.138	0.714
Error	26.000	18	1.444		

Mean time taken to complete trials: Comparison with Experiment 3 (insert ordering): 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (stimulus)	5.639	1	5.639	3.163	0.092
Error	32.092	18	1.783		

Mean time taken to complete trials: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	18.164	1	18.164	1.711	0.207
Error	191.037	18	10.613		

APPENDIX K F tables (Chapter 6)

Experiment 5

Mean number of correct answers: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	5.400	2	2.700	9.346	0.0008
Error	7.800	27	0.289		

Mean time taken to complete trials: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	124.480	2	62.240	7.372	0.0028
Error	227.968	27	8.443		

Comparison of Experiments 5 and 3

Mean number of correct answers: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (age)	8.067	1	8.067	13.444	0.0006
B (ordering)	31.033	2	15.517	25.861	0.0000
AB	6.433	2	3.217	5.361	0.0075
Error	32.400	54	0.600		

Mean number of correct answers: Simple main effects (age x ordering)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
age at					
insert	0.050	1	0.050	0.083	0.7739
left to right	3.200	1	3.200	5.333	0.0248
mixed	11.250	1	11.250	18.750	0.0001
Error Term	32.400	54	0.600		
ordering at					
7 years	32.067	2	16.033	26.722	0.0000
9 years	5.400	2	2.700	4.500	0.0156
Error Term	32.400	54	0.600		

Mean time taken to complete trials: 2 factor ANOVA (age x ordering)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (age)	558.821	1	558.821	44.359	0.0000
B (ordering)	1293.351	2	646.675	51.333	0.0000
AB	574.132	2	287.066	22.787	0.0000
Error	680.274	54	12.598		

Mean time taken to complete trials: Simple main effects (age x ordering)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
age at					
insert	20.503	1	20.503	1.628	0.2075
left to right	271.216	1	271.216	21.529	0.0000
mixed	841.234	1	841.234	66.777	0.0000
Error Term	680.274	54	12.598		
ordering at					
7 years	1743.003	2	871.501	69.180	0.0000
9 years	124.480	2	62.240	4.941	0.0107
Error Term	680.274	54	12.598		

Experiment 6

Mean number of correct answers: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	20.000	1	20.000	75.000	0.0000
Error	4.800	18	0.267		

Mean time taken to complete trials: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	613.389	1	613.389	173.487	0.0000
Error	63.642	18	3.536		

Comparison of Experiments 6 and 5

Mean number of correct answers: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	28.275	3	9.425	32.314	0.0000
Error	10.500	36	0.292		

Mean time taken to complete trials: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	757.707	3	252.569	38.189	0.0000
Error	238.089	36	6.614		

APPENDIX L F tables (Chapter 7)

Experiment 7

Mean number of correct answers: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (number of items)	0.2000	1	0.200	0.277	0.605
Error	13.000	18	0.722		

Mean time taken to complete trials: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (number of items)	11.011	1	11.011	0.746	0.399
Error	265.682	18	14.760		

APPENDIX M F tables (Chapter 8)

Experiment 8

Mean number of correct answers: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (task rel.)	0.100	1	0.100	0.133	0.7171
B (prem. rel.)	0.400	1	0.400	0.533	0.4699
AB	8.100	1	8.100	10.800	0.0023
Error	27.000	36	0.750		

Mean number of correct answers: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
task rel. at					
horizontal	3.200	1	3.200	4.267	0.0461
vertical	5.000	1	5.000	6.667	0.0140
Error Term	27.000	36	0.750		
prem. rel. at					
horizontal	2.450	1	2.450	3.267	0.0791
vertical	6.050	1	6.050	8.067	0.0074
Error Term	27.000	36	0.750		

Mean time taken to complete trials: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (task rel.)	0.967	1	0.967	0.171	0.6815
B (prem. rel.)	8.575	1	8.575	1.518	0.2259
AB	527.221	1	527.221	93.322	0.0000
Error	203.382	36	5.650		

Mean time taken to complete trials: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
task rel. at					
horizontal	241.513	1	241.513	42.749	0.0000
vertical	286.676	1	286.676	50.744	0.0000
Error Term	203.382	36	5.650		
prem. rel. at					
horizontal	200.661	1	200.661	35.518	0.0000
vertical	335.135	1	335.135	59.321	0.0000
Error Term	203.382	36	5.650		

Comparison of Experiments 8 and 3

Mean number of correct answers: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	0.867	2	0.433	2.017	0.153
Error	5.800	27	0.215		

Mean time taken to complete trials: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (ordering)	30.419	2	15.209	6.137	0.0064
Error	66.916	27	2.478		

Experiment 9

Mean number of correct answers: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (task rel.)	0.025	1	0.025	0.046	0.8311
B (prem. rel.)	0.625	1	0.625	1.154	0.2899
AB	18.225	1	18.225	33.646	0.0000
Error	19.500	36	0.542		

Mean number of correct answers: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
task rel. at					
horizontal	9.800	1	9.800	18.092	0.0001
vertical	8.450	1	8.450	15.600	0.0003
Error Term	19.500	36	0.542		
prem. rel. at					
horizontal	6.050	1	6.050	11.169	0.0019
vertical	12.800	1	12.800	23.631	0.0000
Error Term	19.500	36	0.542		

Mean time taken to complete trials: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (task rel.)	74.474	1	74.474	19.757	0.0001
B (prem. rel.)	7.797	1	7.797	2.068	0.1590
AB	293.547	1	293.547	77.875	0.0000
Error	135.701	36	3.769		

Mean time taken to complete trials: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
task rel. at					
horizontal	36.154	1	36.154	9.591	0.0038
vertical	331.868	1	331.868	88.041	0.0000
Error Term	135.701	36	3.769		
prem. rel. at					
horizontal	102.831	1	102.831	27.280	0.0000
vertical	198.513	1	198.513	52.663	0.0000
Error Term	135.701	36	3.769		

Experiment 10

Mean number of correct answers: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (prem. rel.)	0.025	1	0.025	0.134	0.7161
B (task rel.)	0.025	1	0.025	0.134	0.7161
AB	3.025	1	3.025	16.254	0.0003
Error	6.700	36	0.186		

Mean number of correct answers: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
prem. rel. at					
horizontal	1.250	1	1.250	6.716	0.0137
vertical	1.800	1	1.800	9.672	0.0036
Error Term	6.700	36	0.186		
task rel. at					
horizontal	1.250	1	1.250	6.716	0.0137
vertical	1.800	1	1.800	9.672	0.0036
Error Term	6.700	36	0.186		

Mean time taken to complete trials: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (task rel.)	1.243	1	1.243	0.542	0.4662
B (prem. rel.)	0.371	1	0.371	0.162	0.6899
AB	286.921	1	286.921	125.253	0.0000
Error	82.467	36	2.291		

Mean time taken to complete trials: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
prem. rel. at					
horizontal	162.963	1	162.963	71.140	0.0000
vertical	125.200	1	125.200	54.655	0.0000
Error Term	82.467	36	2.291		
task rel. at					
horizontal	133.334	1	133.334	58.206	0.0000
vertical	153.957	1	153.957	67.208	0.0000
Error Term	82.467	36	2.291		

APPENDIX N F tables (Chapter 9)

Experiment 11

Mean number of correct answers: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (base)	6.400	1	6.400	10.569	0.0025
B(prem. rel.)	4.900	1	4.900	8.092	0.0073
AB	2.500	1	2.500	4.128	0.0496
Error	21.800	36	0.606		

Mean number of correct answers: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
base at					
vertical	0.450	1	0.450	0.743	0.3944
horizontal	8.450	1	8.450	13.954	0.0006
Error Term	21.800	36	0.606		
prem. rel. at					
vertical	7.200	1	7.200	11.890	0.0015
horizontal	0.200	1	0.200	0.330	0.5691
Error Term	21.800	36	0.606		

Mean time taken to complete trials: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (base)	65.715	1	65.715	35.834	0.0000
B (prem. rel.)	38.475	1	38.475	20.980	0.0001
AB	51.597	1	51.597	28.135	0.0000
Error	66.020	36	1.834		

Mean time taken to complete trials: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
base at					
vertical	0.426	1	0.426	0.232	0.6326
horizontal	116.886	1	116.886	63.737	0.0000
Error Term	66.020	36	1.834		
prem. rel. at					
vertical	89.591	1	89.591	48.853	0.0000
horizontal	0.480	1	0.480	0.262	0.6119
Error Term	66.020	36	1.834		

Comparisons of Experiments 11 and 8

Mean number of correct answers: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (task rel.)	14.400	1	14.400	36.507	0.0000
B (base rel.)	6.400	1	6.400	16.225	0.0003
AB	4.900	1	4.900	12.423	0.0012
Error	14.200	36	0.394		

Mean number of correct answers: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
task rel. at					
vertical	1.250	1	1.250	3.169	0.0835
horizontal	18.050	1	18.050	45.761	0.0000
Error Term	14.200	36	0.394		
base rel. at					
abstract	11.250	1	11.250	28.521	0.0000
concrete	0.050	1	0.050	0.127	0.7239
Error Term	14.200	36	0.394		

Mean time taken to complete trials: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (task rel.)	71.182	1	71.182	36.314	0.0000
B (base rel.)	63.706	1	63.706	32.499	0.0000
AB	40.040	1	40.040	20.426	0.0001
Error	70.568	36	1.960		

Mean time taken to complete trials: Simple main effects (task x premise presentation relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
task rel. at					
vertical	2.224	1	2.224	1.135	0.2938
horizontal	108.998	1	108.998	55.605	0.0000
Error Term	70.568	36	1.960		
base rel. at					
abstract	102.378	1	102.378	52.228	0.0000
concrete	1.368	1	1.368	0.698	0.4091
Error Term	70.568	36	1.960		

Experiment 12

Mean number of correct answers: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (base rel.)	0.000	1	0.000	0.000	1.0000
B (task rel.)	0.433	2	0.217	0.262	0.7702
AB	13.300	2	6.650	8.052	0.0009
Error	44.600	54	0.826		

Mean number of correct answers: Simple main effects (task relation x base relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
base rel. at					
best	6.050	1	6.050	7.325	0.0091
fastest	7.200	1	7.200	8.717	0.0047
oldest	0.050	1	0.050	0.061	0.8066
Error Term	44.600	54	0.826		
task rel. at					
horizontal	7.267	2	3.633	4.399	0.0170
vertical	6.467	2	3.233	3.915	0.0258
Error Term	44.600	54	0.826		

Mean time taken to complete trials: 2 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (base rel.)	2.440	1	2.440	0.922	0.3412
B (task rel.)	44.766	2	22.383	8.460	0.0006
AB	287.174	2	143.587	54.269	0.0000
Error	142.876	54	2.646		

Mean time taken to complete trials: Simple main effects (task relation x base relation)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
base rel. at					
best	115.440	1	115.440	43.631	0.0000
fastest	174.109	1	174.109	65.805	0.0000
oldest	0.065	1	0.065	0.025	0.8761
Error Term	142.876	54	2.646		
task rel. at					
horizontal	195.176	2	97.588	36.883	0.0000
vertical	136.764	2	68.382	25.845	0.0000
Error Term	142.876	54	2.646		

Experiment 13

Mean number of correct answers: 1 factor ANOVA

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (base)	5.000	1	5.000	8.824	0.0082
Error	10.200	18	0.567		

Mean time taken to complete trials: 1 factor ANOVA

Source of Variation	**Sum of Squares**	**df**	**Mean Squares**	**F**	**p**
A (base)	8.218	1	8.218	11.855	0.0029
Error	12.477	18	0.693		