The three-term series problem*

PHILIP N. JOHNSON-LAIRD

University College London

If 'Images' are better than 'Operations', and
'Deep structures' are better than 'Operations',
What is best?

Abstract

There are three apparently conflicting theories about the processes by which an individual solves three-term series problems, i.e. problems such as 'A is better than B, B is better than C, who is best?' An information-processing analysis is presented which reveals that the IMAGE model of De Soto, Handel and London (1965) can easily be reconciled with the LINGUISTIC model of Clark (1969). However, an examination of some slight inconsistencies among the empirical findings reveals that such a reconciliation may be misleading. It is suggested instead that there is likely to be a change of approach during an experimental session. The individual initially may utilize a procedure analogous to the IMAGE model, modified to incorporate one of the principles of Hunter's (1957) OPERATIONAL model. Subsequently, as the result of practice, he may develop a procedure more analogous to the LINGUISTIC model.

This paper is about a particular sort of deductive problem in which the answer depends upon the relations between items in the premises. Such *relational* inferences can involve any number of premises with a variety of different types of relation, e.g.

John stood in the last local elections in Camden.
Camden is a borough of London.
London had its annual borough elections on Tuesday.
Therefore, John stood in the elections on Tuesday.

* The author is indebted to Dr. P.C. Wason for a critical reading of an earlier draft of this paper.

However, our concern is with a special class of problems, which have only two premises, containing either the same comparative term or a comparative and its converse e.g.

Ken is taller than Bill.
Tom is shorter than Bill
Who is tallest?

These are known as 'three-term series' problems (or 'linear syllogisms').

The most important logical point about the answers to such problems is that they are not, strictly, valid deductions. They depend upon additional and unstated premises. To the ordinary intelligent individual, however, these premises are granted immediately by his knowledge of the language. For example, part of the meaning of *taller* is that it is transitive, i.e. if 'A is taller than B' and 'B is taller than C' then it follows that 'A is taller than C' regardless of what 'A', 'B', and 'C' denote. This is appreciated without conscious reflection just as it is appreciated that the relation of, say, *is the father of* is not transitive. A considerable portion of logical work is thus performed by knowledge of the language.

The fundamental problem in making a relational inference is to set up some internal representation of the premises, be it abstract or concrete, that will allow the relation between those items, not specifically linked in a premise, to be determined. A potential controversy about the process seems implicit in the earliest experimental observations. Störring, who introduced the problem into the laboratory, described a variety of methods used by his subjects (cf. Woodworth, 1938). Some individuals evidently formed a mental diagram of the premises, representing it by a visual or kinaesthetic image. Others pondered upon the meaning of the premises and solved the problem in a purely verbal way.

Although series problems continued to be studied after Störring's pioneer efforts e.g. by Burt (1919) and Piaget (1921, 1928), it was not until fairly recently that comprehensive proposals were made about the way they were solved. Broadly speaking, there are three alternative theories, which will be described in the order of their historical development. Two of them are currently claimed to make rival statements about the phenomena – one is an 'image' theory, the other a 'linguistic' theory. Our aim is to show, as is implicit in Störring's findings, that the parties to the controversy are both partly right and partly wrong.

1. The OPERATIONAL Model

The first definitive proposal about the *process* of inference in series problems was made by Ian Hunter (1957), in what we shall call the OPERATIONAL model. It takes as its starting point an idea which William James described as the fundamental principle of inference, namely, that with a linear series of the form a > b > c > d ... 'any number of intermediaries may be expunged without obliging us to alter anything in what remains written' (James, 1890, p. 646). Hunter assumes that the same principle applies to relational premises provided that they lead naturally on one from the other as in:

A is larger than B
B is larger than C

In other words, the reasoner can simply delete the two occurrences of 'B' and infer that 'A is larger than C'. However, says Hunter, where the premises are not arranged in this 'natural' order, certain cognitive OPERATIONS have to be performed in order to bring them into it. This is a psychological analogue of the traditional idea in logic of reducing an inference to a simpler form known to be valid. But, of course, in the present case the reduction is not necessarily a deliberate or conscious stratagem.

There are two main operations. First, a premise may have to be *converted.* This usually applies to the second premise because Hunter assumes that the first premise creates a 'set' for its interpretation. Thus, with a problem such as:

A is larger than B
C is smaller than B

the second premise has to be converted to 'B is larger than C' before the middle term can be expunged. Second, the premises themselves, and not their terms, may have to be *re-ordered,* when they do not naturally follow on one from the other. Such an operation is required in the case of:

B is smaller than A
C is smaller than B

in order to yield the 'natural' order: C is smaller than B, B is smaller than A. We have to imagine that the premises are cognitively re-arranged so that in the mind's eye they *do* follow on from each other. Some combinations of premises require both operations to be performed; e.g.

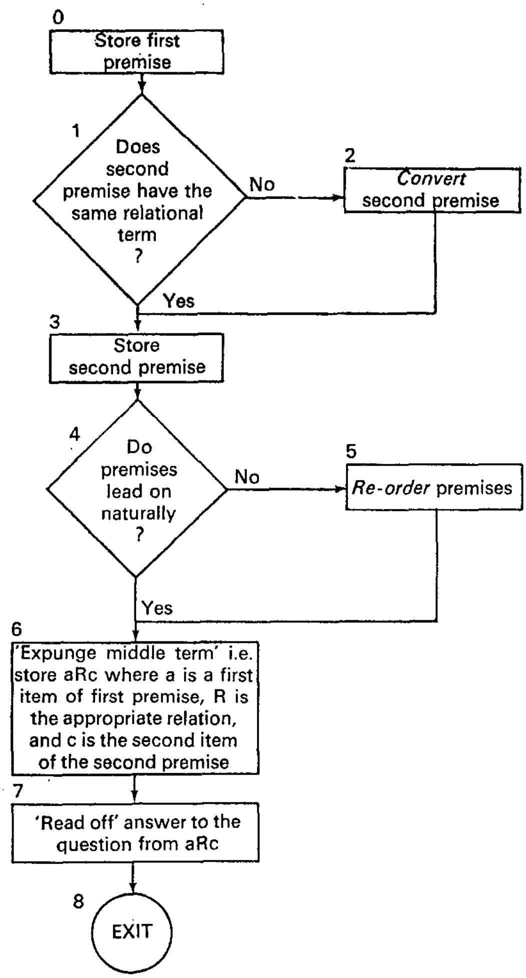
B is smaller than A
B is larger than C

Here it is necessary to convert the second premise to 'C is smaller than B', and then to re-order the two premises. (The 'natural' order could also be obtained by going back to the first premise to convert it. This, Hunter claims, would also involve two operations: *reverting* to the first premise and *converting* it to 'A is larger than B'.)

In order to facilitate comparison with the subsequent models, we have attempted to express the model in information-processing terms, summarizing it in the flow diagram of Figure 1.

According to the model, the individual sets up some internal representation of the first premise (0). He then checks to see whether the second premise involves the same rela-

Figure 1. *The OPERATIONAL model for the solution of three-term series problems (based on Hunter, 1957)*



tion (1), and, if it does not, he converts it (2). If the premises follow on 'naturally' (4), the middle term can be deleted and the two remaining terms put together in their appropriate relation (6). If the premises do not follow on 'naturally', they are re-ordered (5). Finally, the answer to whatever question is put about the premises can be 'read off' (7), and the process terminates (8).

(It should be noted that some possible combinations of premises do not yield any conclusions, either because there is no item common to them or because they are contradictory or indeterminate. It is a simple matter to add to the model procedures that ensure that such problems are successfully detected, but, for the sake of simplicity, they have been omitted from the description of this and all subsequent models.)

On the supposition that these processes follow one another in serial fashion, it is clear that problems involving *conversion* or *re-ordering* should take longer to solve than those involving neither operation, and that those involving both operations should take longest of all. Hunter tested these predictions using four sorts of problems, and these are given in Table 1 together with the hypothetical operations involved in their solution. The actual premises involved relations such as *taller* (>) and *shorter* (<),

Table 1. *The predictions of the OPERATIONAL model for the material used by Hunter (1957)*

		1 A > B B > C	2 A > B C < B	3 B < A C < B	4 B < A B > C
*Operations predicted by the model	Conversion	—	+	—	+
	Re-ordering	—	—	+	+

* A '+' indicates that the operation occurs according to the model, a '—' indicates that it does not occur.

and were followed by a question, either 'who is tallest?' or 'who is shortest?'. The subjects were 16 year-old and 11 year-old children.

The distribution of the times taken by the 16 year-olds to solve the problems supported the model: type 1 problems were easier than types 2 and 3, and type 4 were hardest of all. Since type 2 problems were easier than type 3, Hunter argues that *conversion* is a simpler operation than *re-ordering*. Somewhat surprisingly, the 11 year-olds even found type 2 problems easier than type 1 – a finding that Hunter attributes to an 'atmosphere' effect.

2. The IMAGE Model

A more recent proposal about the solution of series problems is due to De Soto, London, and Handel (1965). They suggest that the crucial step is the combination of the interpretations of premises into a unitary representation. This unitary representation consists of a visual image of a vertical or horizontal array in which the items are located in their appropriate position. The novel aspect of the IMAGE theory, however, is that the difficulty of a problem depends to some extent upon the sort of relational terms used in the premises. To illustrate the point, they recount an anecdote about a baseball fan who went to see two great players in an exhibition game. Unfortunately, both were having an 'off day'. Finally, the fan, unable to contain his anger, bellowed at one of them, 'I came to see which of you guys was better – instead, I'm seeing which is worse!'

Clearly, this is an insult, but why? – Isn't it the case that the two statements are equivalent? De Soto and his colleagues claim that they are not, because *better* refers to the 'good' end of the scale, whereas *worse* refers to the 'bad' end. The fan is implying that neither player is any good. It follows that if such comparatives are represented by a vertical array, items related by *better* will be at the top end and items related by *worse* will be at the bottom. Moreoever, the items related by *better* will be inserted into the array working downwards, whereas those related by *worse* will be inserted working upwards.

The theory postulates that evaluative comparatives, even though they are not explicitly spatial, are indeed represented in this manner. Other comparatives such as *wider* and *narrower* call for a horizontal array, or else like *lighter* and *darker* seem to call for no consistent orientation. There are two principles governing the construction of arrays, First, there is a natural preference for constructing vertical arrays starting from the top and working downwards, and for constructing horizontal arrays working from left to right. It is unclear whether this is intended to be anything more than the principle governing occidental reading habits. Second, a premise is easier to represent in the array if its first item is an 'end-anchor', i.e. an item that occurs at one end of the final array rather than the middle item.

The evidence that De Soto obtained certainly seemed to support the principle that a premise would be easier to interpret when its items can be put into the array working downwards. A problem of the form:

A is better than B
B is better than C

was consistently easier than one of the form:

C is worse than B
B is worse than A

Equivalent results were also obtained with *above – below* and *more – less* (see Handel, De Soto, and London, 1968). Such differences cannot be explained by Hunter's OPERATIONAL model because the premises follow on naturally from one another in both types of problem, but they can be explained in terms of the preferred direction of working. The principle is also claimed to apply to the order of the premises themselves as well as to the items within them. Hence, a problem such as:

A is better than B
C is worse than B

allows the array to be constructed working downwards, since it contains the 'best' item, A, in its first premise; whereas the problem with the same premises in the opposite order:

C is worse than B
A is better than B

requires the array to be constructed working upwards, since it contains the 'worst' item, C, in its first premise. This prediction, too, was confirmed experimentally. Performance was consistent with the OPERATIONAL model only with those comparatives which did not automatically suggest one end of the array of the other, e.g. *darker – lighter*. Indeed, in these cases De Soto suggests that the first item of the first premise is assigned to the topmost (or leftmost) position of the array, and, if necessary, the second premise is converted so that the third item can be added to the array.

The second principle concerning 'end-anchoring' seems somewhat arbitrary. Why should problems be easier if their premises are end-anchored? Janellen Huttenlocher (1968) suggests that it is not the fact that the end-anchor is the first item that is crucial but rather that it is the grammatical (deep structure) subject of the sentence. She and her colleagues (Huttenlocher and Strauss, 1968; Huttenlocher, Eisenberg, and Strauss, 1968) discovered that when children have to arrange objects to fit such descriptions as, 'The blue block is on top of the brown block', they find the task easier when the block to be moved corresponds to the subject of the sentence rather than to the object. If the adult's construction of imaginary arrays is similar to the child's construction of real arrays, the difficulty of a premise should depend on the grammatical status of the item that has to be fitted into the array. Hence, the principle of end-anchoring may be reformulated: it is easier to understand a premise that refers to an end item with its grammatical subject than with its grammatical object.

De Soto's principles, together with Huttenlocher's explication of end-anchoring, seem intuitively satisfactory until one attempts to put them into information-processing terms. It is then apparent that the two principles are confounded. This can best be demonstrated by considering the different ways in which a given three-term series can be expressed. Consider the following series: $A > B > C$. There are 8 basic ways of formulating premises that will yield this series: –

First premise:	$A>B$		$B<A$		$B>C$		$C<B$	
	/	\	/	\	/	\	/	\
Second premise:	$B>C$	$C<B$	$B>C$	$C<B$	$A>B$	$B<A$	$A>B$	$B<A$

Any one of these eight formulations is completely specified by describing its characteristics solely with respect to the first principle of the IMAGE theory, which concerns the direction of working. For example, if the premises are ordered so that the 'worst' item is in the first of them, and both of them involve working downwards, then the problem must be: 'B is better than C, A is better than B'. It follows that the IMAGE theory, as formulated by De Soto *et al*, incorporates two principles which are not independent, i.e. (a) the principle governing the direction of working, and (b) the principle governing end-anchoring.

Obviously, there are a number of ways in which the theory could be reformulated, sacrificing one or other principle with regard to the first or second premise. In attempting to put it into information-processing terms, a guiding cue was Huttenlocher's (1968) remark that subjects do not report making a spatial array for the second premise. Their goal is rather to determine the end-anchor and its position relative to the array representing the first premise. This version of the IMAGE theory is summarized in Figure 2.

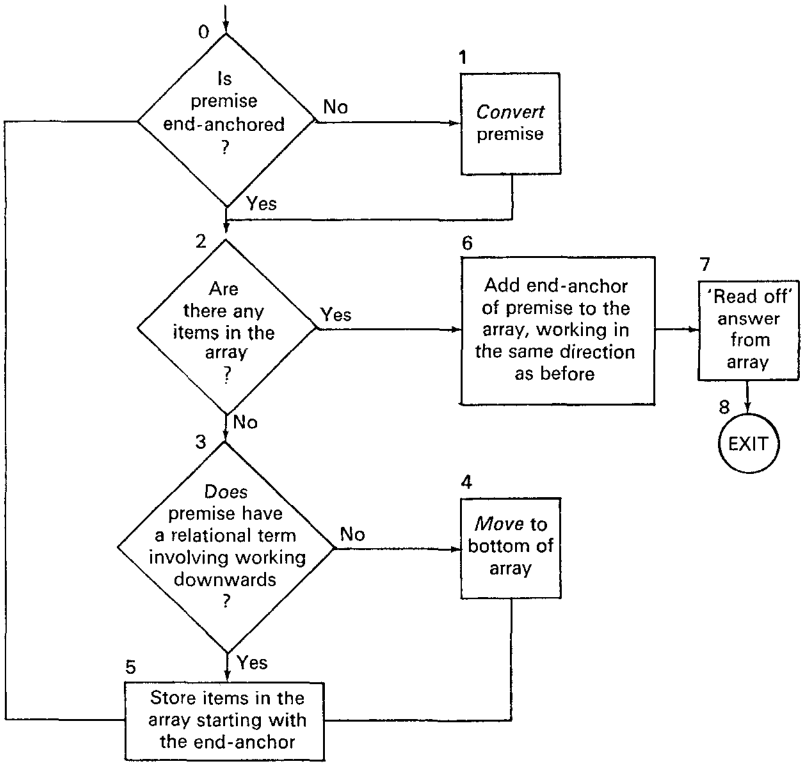
What happens according to the model is that the individual determines whether the first premise is end-anchored (0). This clearly involves considering its first item (or subject), and examining the second premise to see whether it also occurs there. For example, the first premise of the problem, 'B is better than C, A is better than B', is clearly not end-anchored since its first item, 'B', also occurs in the second premise. The model assumes that, in such cases, the premise is *converted* (1). It becomes 'C is worse than B' which clearly is end-anchored since 'C' does not occur in the second premise, 'A is better than B'.

Since no items have previously been encoded (2), the by now end-anchored premise will be represented by an array constructed from the end-anchor either working downwards from the top (3, 5), or else upwards from the bottom (3, 4, 5), whichever is appropriate. Working upwards from the bottom is assumed to be more difficult and hence requires an additional operation (4). In encoding the second premise, there is again a search for the end-anchor (0, 1); but, since there are now items in the array (2), it is only necessary to add the end-anchor (6), *working in the same direction as before*. The answer may be 'read off' from the resulting array (7, 8).

It might be argued that when a premise is not end-anchored, the individual merely encodes its second item prior to its first item. This would be theoretically distinct from the present proposal, i.e. that the premise is converted, though it is difficult initially to see quite how the two ideas differ in practice. However, one subtle consequence of conversion is that when it is applied to the first premise, in order to locate the end-anchor, the direction of constructing the array is also necessarily changed. Hence, not all premises with *better* involve working downwards: those in which the first item is not an end-anchor will be converted to *worse* and will require the array to be constructed working upwards. This can hardly be said to conflict with De Soto's theory since it is precisely at this point that his theory incorporates two opposing principles.

The simplest way to evaluate the reformulated IMAGE model is to compare its predictions with the results obtained by De Soto *et al* (1965) and Huttenlocher (1968). De Soto and his colleagues presented each combination of premises four times, once with

Figure 2. *The IMAGE model for the solution of three-term series problems (based on De Soto et al, 1965)*



each of the questions, 'Is A better than C?', 'Is C better than A?', 'Is A worse than C?', and 'Is C worse than A?', where 'A', 'B', and 'C', were in fact mens' first names. The subjects were allowed ten seconds in which to answer 'yes' or 'no', and the data consist of the percentage of correct answers over the four versions of each premise combination. Huttenlocher worked with the relations *taller* and *shorter* and used a rather different technique. After the first premise of the problem, the subject was asked both 'who is taller?' and 'who is shorter?' in a counterbalanced order. The second premise was presented, and finally the subject was asked either 'who is tallest?' or else 'who is shortest?'. Two versions of each type of problem were given to the subjects, and the data consist of the times which elapsed from the presentation of the second premise until the subject responded. Such a technique is likely to minimize the contribution of the first premise to the difficulty of the problem, and this is clearly reflected in the results. Table 2 gives the results of both these experiments together with a summary of the predictions of the reformulated IMAGE model.

Table 2. *The IMAGE model: predictions and results*

	1st premise	$A > B$		$C < B$		$B < A$		$B > C$	
	2nd premise	$C<B$	$B>C$	$A>B$	$B<A$	$C<B$	$B>C$	$A>B$	$B<A$
Operations predicted by the model	Convert 1st premise	−	−	−	−	+	+	+	+
	Work upwards	−	−	+	+	−	−	+	+
	Convert 2nd premise	−	+	−	+	−	+	−	+
	Percentage of correct answers (De Soto et al, 1965)	61.8	60.5	57.0	42.5	50.0	41.5	52.8	38.3
	Latencies (centisecs) (Huttenlocher, 1968)	141	155	142	161	142	157	135	157

The model evidently gives a good account of the findings though it appears that the effects of the hypothetical operations are not necessarily linear.

3. The LINGUISTIC Model

The third and most recent approach to the problem is the LINGUISTIC theory developed by Herbert Clark (1969a, b). Clark argues that the process of deduction is virtually identical to the process of comprehension, and that difficulties in solving three-term series problems can be accounted for by three psycholinguistic principles.

The first is the principle of *lexical marking*, which asserts that certain comparatives are easier to understand and to remember than others. According to a linguistic analysis, so-called 'unmarked' comparatives, such as *better* and *taller*, can be used in a neutral way merely to convey the relative degrees of two items on a scale, whereas the converse 'marked' terms, *worse* and *shorter*, always convey in addition something about the absolute position of the items on the scale. Thus, for example, 'A is better than B' normally informs one only of the respective merits of A and B, but 'B is worse than A' also informs one that both A and B are towards the 'bad' end of the scale. The asymmetry between such comparatives, in contradistinction to De Soto's claim that they merely refer to different ends of the absolute scale, is supported by several characteristics of unmarked comparatives. It is, for instance, quite acceptable to assert: 'John and James are both very bad but John is slightly better than Bill'. But it is somewhat deviant to assert: 'John and James are both very good but John is slightly worse than Bill'. Unmarked comparatives, as befits their neutral usage, also tend to give their name to the scale as a whole, e.g. *length, width, depth, height, heaviness, thickness, warmth.* The greater psychological simplicity of unmarked comparatives is confirmed by Donaldson and Wales (1970), who found that children tend to understand them before they can understand marked comparatives, and by Clark and Card (1969), who found that they are easier to remember than marked comparatives.

On the assumption that unmarked comparatives are taken in their neutral sense – for on occasions they can be used in an absolute way, it follows that they should be easier to work with than marked comparatives, because they lack the additional absolute information. This principle obviously accounts for the same sort of phenomena as De Soto's notion of a preferred direction for constructing the array. Both hypotheses predict that

A is better than B
B is better than C

should be easier than

C is worse than B
B is worse than A

but for different reasons. Their predictions diverge, however, for a new sort of premise introduced by Clark. These are the so-called 'negative equatives' and they have the form, 'B isn't as good as A' or 'A isn't as bad as B'. The first example involves working up the array and should therefore be more difficult than the second, according to the IMAGE theory. But the LINGUISTIC theory makes the opposite prediction because the first example contains the unmarked term 'good' whereas the second contains the marked term 'bad'. (To anticipate some results, it seems that the LINGUISTIC theory is right and the IMAGE theory is wrong.)

The second psycholinguistic principle in Clark's theory has an interesting precursor in a 'clinical' observation made by Piaget (1921). Evidently a potent source of error in children's reasoning arises in the following way. Given the statement, 'Edith is better than Suzanne', the child is likely to argue that they are both good. When this is followed by the statement, 'Edith is worse than Lily', the child argues that *they* are both bad. Hence Suzanne is good, Lily is bad, and Edith is between the two. Of course this is nonsense, but it is close to the principle of *the primacy of functional relations* which Clark proposes for adult reasoning. He suggests that with a statement such as, 'Edith is worse than Lily', an individual can comprehend that Edith and Lily are bad faster then he can comprehend their relative degrees of badness. Clark justifies this intuitively plausible assumption on linguistic grounds. The *deep structure* or underlying representation of the statement is, Clark claims, roughly of the form: (Edith is bad) more than (Lily is bad)[2], and the simple subject-predicate relations are easier to retrieve than the relation between the two clauses.

There is an addendum to this principle, and this too is related to an earlier observation. Margaret Donaldson (1963) noted that children often encode a premise such as 'Mary is taller than Martha' as 'Mary is the taller one'. (Both Burt (1919, p. 126) and Hunter (1957) suggest that on occasion young children may even encode it as 'Mary is the tallest one'.) The consequences of the abbreviated encoding depend upon the form of the second premise. If it is 'Ann is taller than Mary', it is a simple matter to infer that Ann is the tallest. But if it is 'Martha is taller than Ann', then there is no obvious series. As Clark points out, in appropriating the strategy, the subject must try to recover the whole first premise or backtrack with the information that 'Martha is the taller one' to apply it to the first premise, or revert to some other time-consuming strategy.

The principles of lexical marking and of the primacy of functional relations characterize the comprehension of premises. Let us attempt to express them in information-processing terms, before proceding to the third principle, which governs the answering of the question. Although the theory is consistent, it fails to specify the complete process of solution. The most crucial omission, surprisingly, concerns what happens after the separate interpretations of the two premises. In principle, there seem to be three major possibilities. (a) After appropriate operations, the middle item can be simply expunged leaving the remaining items in their correct relation (e.g. Hunter's theory); (b) the information in the premises can be combined to form some sort of unified representation of the three items (e.g. De Soto's theory); and (c) the three items can be separately stored with some representation of their respective possession of the attribute in question.

The distinction between these approaches may seem abstruse, and perhaps it may be clarified by an analogy with a simple everyday problem. To determine the relative weights of three objects, they could be compared two at a time using a balance, and the middle item would clearly be heavier in one comparison and lighter in the other. Or, they could be attached to three separate springbalances, and the resulting spatial

2. This analysis of the underlying structure of comparatives is due to Smith (1961). It contrasts with Lees' treatment (Lees, 1961) in which one clause is embedded within the other, e.g. (Edith is more than (Lily is bad) bad). The distinction does not affect Clark's hypothesis, though Campbell and Wales (1969) and Ross and Perlmutter (1970) have argued against both sorts of analysis. Clark points out that *good* in clauses like (A is good) is neutral in contrast to *bad* which carries absolute information. Strictly speaking, a more abstract notation is required if only to account for the fact that 'Edith is worse than Lily' logically implies 'Lily is better than Edith'. The converse deduction, if Clark is correct, is invalid unless it is established independently that Lily and Edith are both bad.

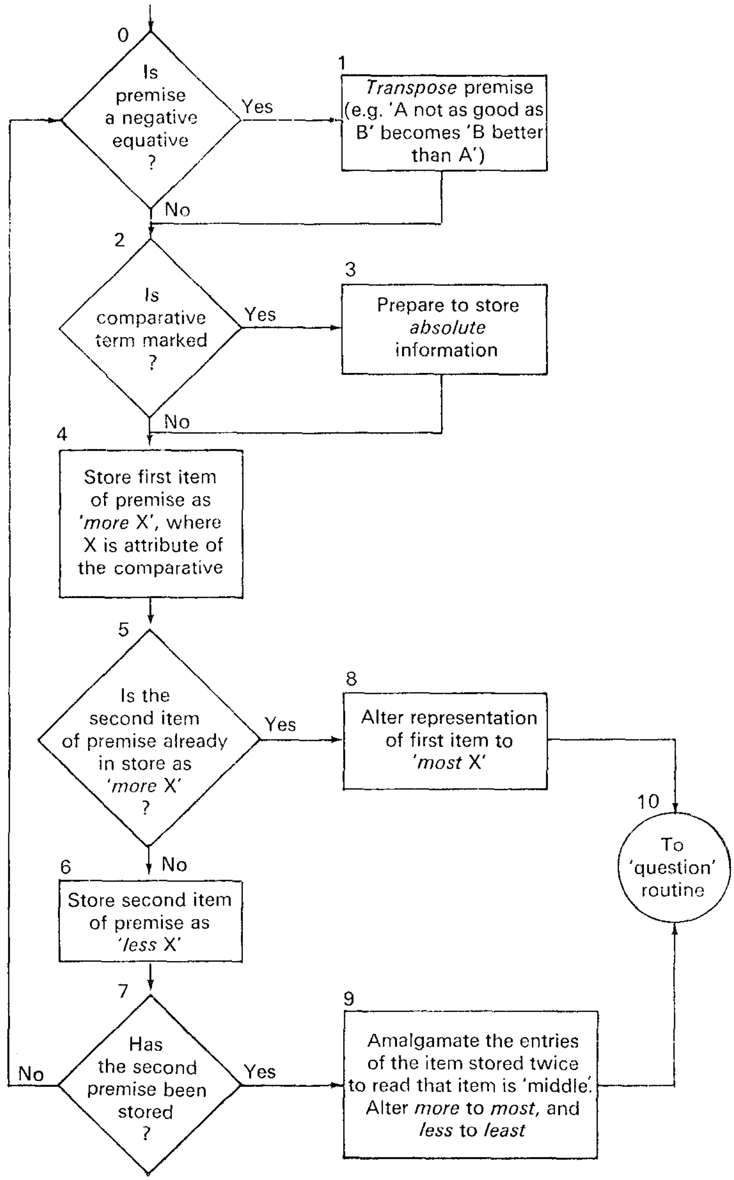
arrangement would clearly reflect their relative weights. Finally, their weights could be recorded using a conventional weighing machine. These three different methods correspond to the three possibilities for solving series problems. It seems that the spirit of Clark's theory, and in particular his principle of the primacy of functional relations, commits him to the 'weighing machine' approach. The reasoner stores each item separately according to the underlying information conveyed by the premise, and the 'weight' attached to the middle item is merely adjusted according to the information about it in the second premise. For example, the sentence 'A is better than B' would be represented by storing *A is more good* and *B is less good.* A further premise, 'C is worse than B', would be represented by *C is more bad* and *B is less bad.* This information would then be classified by amalgamating the item stored twice to read *B is middle* and by setting *more* to *most* (and *less* to *least*). Hence the final representation would be of the form: *A is most good, B is middle, C is most bad.* It should be borne in mind, of course, that the representations involving *bad,* or any marked item, will be more complex than those involving *good,* or any unmarked item, since according to the principle of lexical marking the marked terms contain an additional piece of information about absolute scale-values.

Figure 3 presents the first part of the LINGUISTIC model based on this interpretation of Clark's theory. It deals solely with the comprehension and representation of the premises; the solution of problems depends upon a further component to be described presently.

When a premise is presented, the first task is to analyze it linguistically. It is assumed that with negative equatives, the premise is transposed in order to eliminate the negation (0, 1). For example, a premise such as 'John isn't as tall as Bill' is transposed to 'Bill is taller than John'. This assumption goes beyond what Clark specifically claims, yet it is plausible since converting the premise to 'John is shorter than Bill' would introduce a marked comparative in place of the unmarked original. However, it could be claimed that negative equatives concentrate attention upon the item which comes second, and hence that the second item is merely stored before the first. Flores d'Arcais (1971) has suggested that there is this sort of re-focussing of attention in understanding such sentences as, 'Lambs are *less* ferocious than lions', which clearly resemble negative equatives. However, in this instance, it is immaterial which approach is adopted, since transposition, unlike conversion, does not change the comparative from marked to unmarked, or vice versa. But it does eliminate the negative and in this respect parallels the performance of many subjects attempting to remember negative equatives (cf. Clark and Card, 1969).

The model assumes that what the individual stores from the premises depends upon whether the comparative is marked or unmarked (2). Specifically, if the premise is of the form 'A is better than B', he stores the fact that A is *more good* (4). (This is not to

Figure 3. *The LINGUISTIC model. Part 1: the comprehension of three-term series problems (based on Clark, 1969a, b)*



be taken literally; all that is assumed is that A is symbolically represented as the 'better' item.) If the premise is 'A is worse than B', he stores the fact that A is *more bad* (3, 4). Since at this stage no premises have been previously stored (5), the other item mentioned in the premise is stored as being *less good* or *less bad*, whichever happens to apply (6).

The procedure for storing the second premise (7), is exactly the same for dealing with its first item (0, 1, 2, 3, 4). Suppose, however, the problem is of the form: 'B is better than C, A is better than B'. When A has been stored as *more good*, the model discovers that B (the second item of the second premise) has already been stored as *more good* (5), hence it is obvious that A is the best of the trio. This is represented in the model by storing A as *most good* (8). An exactly parallel process can occur with problems concerning *worse*. An obvious series has been found and the individual can now begin to deal with the specific question posed in the statement of the problem (10). But consider a problem of the form: 'A is better than B, B is better than C'. The only difference between this and the previous problem is in the order of the premises. But the difference is crucial. When B in the second premise has been stored as *more good* (4), the check (5) reveals that C (the second item in the premise) has not been previously stored as *more good*. Hence no obvious series emerges and in Clark's words it is necessary for the subject to 'backtrack' or to 'engage in some other time-consuming strategy. In fact, it is assumed that he goes on to store C as *less good* (6). There are no further premises to be interpreted (7); and the individual has stored the following items:

A is *more good*
B is *less good*
B is *more good*
C is *less good*

How should he next proceed? It is, of course, at this point that Clark gives no specific answer, and that the 'weighing machine' approach was selected. The item which has been stored twice is located and the fact that it is the *middle* item recorded (9); and, in order to simplify the business of answering the question, any remaining occurrence of *more* is altered to *most*, and any remaining occurrence of *less* is altered to *least*.

The first part of the model gives a sufficient representation of the premises for the problem to be solved. It contains three decisive choice points (i.e. (0), (2) and (5)) which will affect the psychological difficulty of a problem. First, problems involving negative equatives will be harder because they necessitate an extra operation of transposition. Second, problems involving 'marked' comparatives such as *worse* will be harder because they require more information to be stored. Third, certain problems require information to be stored about both items of both premises. They will be harder than those which require only the first item of the second premise to be stored.

Athough both Hunter (1957) and De Soto (Handel, De Soto, and London, 1968) reported that the nature of the question put to the subject about the premises exerts a significant effect upon performance, neither of them sought to integrate an explanation of the effect into their main theories. However, Hunter makes an interesting remark about the problem: 'B is shorter than A; C is shorter than B'. He writes (p. 244), 'In deriving a series from these premises, the writer was very much aware of finding that C interrupted the direction of the series and of bringing forward C as the shortest member of the series. This made it possible to answer straight away the question 'who is shortest?' whereas the question 'who is tallest?' required a further reconsidering of the two remaining terms to decide which was indeed the tallest. If this introspective evidence is generalized into a formal statement, it would be: the term which is contained only in that premise which has to be reorganized is isolated on the ground that it should come at one end or the other of the entire series.' Hunter establishes that his results, especially with the 16-year old children, tally with such an explanation.

Clark, on the other hand, assumes that the nature of the question has a profound effect on the subject's performance. His third principle is that the subject searches for information which is *congruent* with the form of the question. A premise of the form 'A is better than B' is represented by storing A is more good and B is less good. If one then asks, 'who is worse?', the question is incongruent with such a representation. It takes longer to answer even just about a single premise, as Clark (1969a) himself showed experimentally. In information-processing terms, there are two possible strategies in such a case: either the information stored with the items must be converted, or else the form of the question must be converted, with the goal being in both cases to make information and question congruent with one another. Clark makes the more parsimonious suggestion: he assumes that the question is converted from 'who is worse?' to 'who is least good?'. It is now possible for a search to be made among the items for the one that is least good. This final part of the LINGUISTIC model is represented in Figure 4.

The model assumes that the reasoner first checks to determine whether he has any information which is congruent with the question (11). If the question is 'who is worst?' but both premises contain the relation *better*, then this test will be failed, and the question will be converted (12) to 'who is least good?'; a similar conversion of 'who is best?' to 'who is least bad?' is made in the appropriate circumstances. Even if there is some information congruent in form to the question, it does not follow that the question is answerable. Consider, for example, the comprehension of the premises:

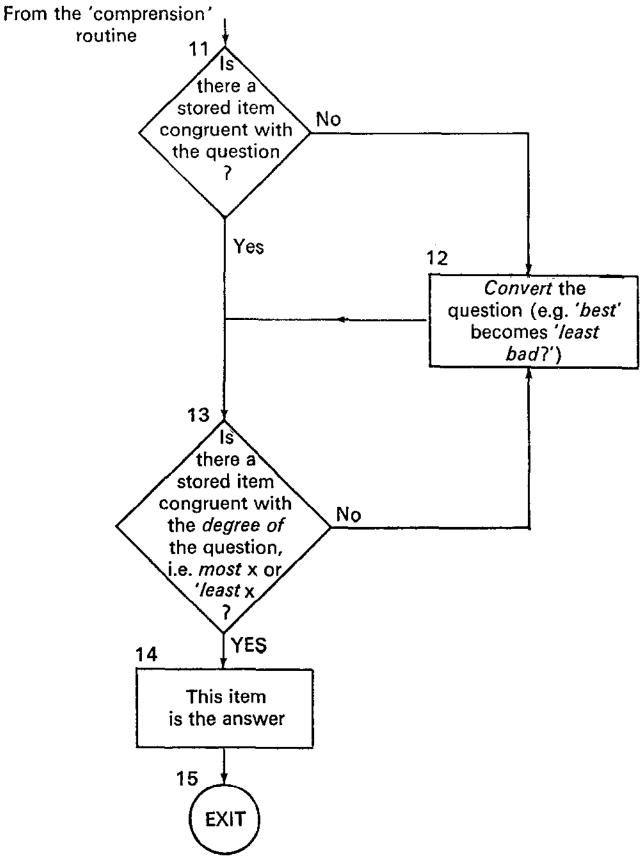
B is better than C

B is worse than A

This will lead to the representation: –

B is more good

Figure 4. *The LINGUISTIC model. Part 2: The question-answering routine (based on Clark, 1969a, b)*



C is less good
B is more bad
A is less bad

The two entries for B will be amalgamated etc. so that what is finally stored is:

C is least good
B is middle
A is least bad

Although there is an item stored with *good*, in answering the question 'who is best?' the required item is one stored as *most good*. There is no such item (13); and it is necessary to convert the question (12) to 'who is least bad?' in order to obtain the answer (14), 'A'. An analogous procedure is necessary to obtain the answer to 'who is worst?'.

This rather subtle interaction between understanding the premises and answering the question gives an alternative account of the phenomena that De Soto explained in terms of end-anchoring. The problem that has just been analyzed, for instance, is found to be more difficult to solve than one of the following form: 'C is worse than B, A is better than B'. The first problem requires both questions to be converted, but this second problem always has congruent information available because C is stored as *most bad* and A as *most good*. The difference is reflected in the results of De Soto et al (1965), Huttenlocher (1968) and Clark (1969a, b) himself.

4. The current controversy

It should now be evident that the IMAGE theory and the LINGUISTIC theory offer alternative explanations of the phenomena. For premises involving simple comparatives, the two theories compete together, making the same prediction for different reasons. Parallel to the IMAGE principle that it is easier to work down the array, there is the LINGUISTIC principle that unmarked comparative terms are simpler; and parallel to the 'end-anchoring' principle there is the principle of 'congruity'. Not all pairs of antonymous comparatives yield problems which differ empirically in their ease of solution, e.g. *lighter – darker*, *fatter – thinner*. But even here the two theories are largely in agreement. Such pairs tend not to elicit any 'directional preference' according to De Soto; and, as Clark has observed, they also tend to consist of terms which are *both* marked, and therefore involve the same amount of information. In these cases, Hunter's OPERATIONAL model also appears to characterize performance quite adequately.

The theories do not always coincide in this fashion. The IMAGE theory predicts that *to the left of* should be easier than *to the right of* because it allows a horizontal array to be built up in the natural and preferred direction, i.e. from left to right. The LINGUISTIC theory presumably makes no such prediction since both relations would seem to be marked. The evidence (De Soto et al, 1965) supports the IMAGE theory. However, it is clear that *to the left of* provides a very special mnemonic that its converse does not: the physical disposition of terms on the printed page corresponds to the spatial locations described by the statement. Subjects might quickly appreciate this fact and exploit it in their inferential strategy. The IMAGE theory also predicts that *shallower* should be easier than *deeper* because it allows the array to be constructed working downwards. The LINGUISTIC theory makes the opposite prediction because *deeper* is unmarked whereas *shallower* is marked. In this case, the results (Clark, 1969a) support the LINGUISTIC theory.

The most obvious confrontation between the two theories seems to occur with the negative equative premises. On extrapolating the IMAGE theory, problems involving the relation *not as good as* should require the array to be constructed working upwards. Hence they should be harder than those involving the relation *not as bad as*. But exactly the opposite prediction is made by the LINGUISTIC theory on the grounds of lexical marking. A similar conflict arises over the matter of end-anchoring. A problem such as

A is not as bad as B
C is not as good as B

contains premises which are both end-anchored, i.e. the first item is at one end of the array. But neither premise in:

B is not as good as A
B is not as bad as C

is end-anchored. Hence, on IMAGE theory, the first problem should be easier than the second. A little reflection should convince the reader that the LINGUISTIC theory makes the opposite prediction on the grounds of congruity between questions and premises. There is little doubt from the findings of Clark (1969a, b) and Huttenlocher et al (1970) that the predictions of the LINGUISTIC theory are confirmed on both counts. But does this mean that the IMAGE theory is eliminated?

The answer is not simple. Most investigators are agreed that subjects may construct a mental picture of the premises. But this in itself has little explanatory value. It is unlikely that all individuals can, or do, construct images, or that all relational problems can be represented in visual terms. The main function of the imaginal aspect of performance may be as an aid to memory. In blindfold chess, for example, some kind of visual representation of the pieces seems to be vital in order to keep track of the moves. But this representation no more determines which move should be made than would the use of an actual chessboard. Similarly, making an inference requires the reasoner to construct an underlying representation of the premises. The process is probably similar to the one that occurs in ordinary comprehension, which again may, or may not, involve visual imagery. But regardless of the form of the representation, the process of inference requires a number of steps in information-processing; and it is a specification of these processes, not whether the representations are abstract symbols or concrete images, which constitutes an explanation of the phenomena.

Quite apart from such explanatory considerations, several studies have found less evidence for explicitly visual imagery than was provided by De Soto's study. Sheila Jones (1970) studied her subjects' directional preferences by the ingenious expedient of giving them pencil and paper and allowing them to jot down the names mentioned in three-term series problems. Nearly three-quarters of her subjects wrote the names in systematic orders, usually preferring a vertical to a horizontal axis. The majority wrote down the names in the first premise in a preferred order (as De Soto predicts) and then added the third name to conform with the order. However, once a subject had decided upon an axis he seldom changed from it as a function of the relational terms. For example, subjects using the vertical axis did not tend to change to the horizontal to represent problems involving *lighter* and *darker*. This contrasts with De Soto's claim that certain relations require certain axes.

Yet it is a simple matter to restore the IMAGE theory, and to reconcile it with the LINGUISTIC theory. The principle that negative equatives are implicitly transposed (e.g. 'A is not as good as B' becomes 'B is better than A'), which makes good sense for the LINGUISTIC theory, makes equally good sense for the IMAGE theory. Indeed, Huttenlocher, Higgins, Milligan and Kauffman (1970) have already invoked it in arguing for their general principle that overt tasks involving the movement of objects mirror the difficulties of covert conceptual tasks of an equivalent form.[3] If this transposition is made, then the manifest differences between the two theories disappear. A premise with *not as good as* still requires the array to be constructed working downwards though the order in which the items are inserted into it is reversed. Hence the principles of preferred direction of working and of lexical marking are reconciled. Similarly, a problem of the form, 'B is not as bad as C, B is not as good as A' seems at first glance to have premises that are not end-anchored. When they are transposed, however, both are clearly end-anchored, and the principle of end-anchoring again coincides with the principle of congruity.

5. The development of strategies

Although it is easy to reconcile the two theories, this may fail to do justice both to them and to the three-term series problem. The real contrast between them does not concern the role of visual imagery or the 'primacy of functional relations' – both theories can be reformulated without them and without any loss in explanatory power. It lies, in the writer's opinion, in a more abstract distinction about the representation of premises. The IMAGE theory assumes that the two premises are combined into a unified representation of the three items; the LINGUISTIC theory assumes that information about the items is stored separately. However, it is entirely feasible that the inexper-

3. Huttenlocher's claim that a transposition of another form may also occur, e.g. from 'A is not as high as B' to 'A is lower than B', is corroborated by Jones (1970). A number of factors seem to be involved in eliciting it. It seems to be more likely with unmarked comparatives, as one would expect from the principle of lexical marking. It may also have been encouraged by Huttenlocher's use of the questions, 'which is on the top?' and 'which is on the bottom?'. They introduce a new sort of incongruity with the premises, and this may have suggested that the differences between *higher* and *lower* is less important than usual.

ienced subject represents the premises in a unified form (with or without imagery) be-because this is presumably the normal mode of interrelating different assertions on the same topic in everyday life. But the more experienced and practised subject has no need of such an elaborate procedure. It is unnecessary to form a unified representation to solve the problem, and he may well learn to store the minimum information which suffices – separate 'weights' on one or two of the items. In short, subjects seem likely to pass from an approach analogous to the IMAGE theory to one analogous to the LINGUISTIC theory.

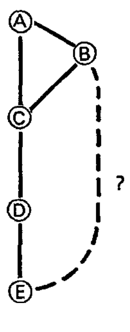
The most convincing evidence for this sort of change comes from a study by David Wood (1969). He used series problems involving up to six premises and giving rise to many different types of array. All the premises involved the comparative term *taller*; and the question was always of the form 'who is taller X or Y?'. A typical problem was:

1. D is taller than E
2. C is taller than D
3. A is taller than C
4. A is taller than B
5. B is taller than C

   Who is taller B or E?

The structure of such a problem can best be represented by what mathematicians call a 'Haas' diagram. The appropriate one is given in Figure 5. The items are represented in a vertical array according to their relative heights. The dotted line represents the question posed at the end of the premises.

Figure 5. *A Haas diagram of a five-term series problem*



Wood predicted that subjects would initially solve such problems after the fashion of De Soto's IMAGE theory: they would build up an internal representation of the items in a structure which would presumably resemble the Haas diagram (though Wood uses a different sort of diagram). However, he suspected that with experience they would develop a more sophisticated strategy. Granted the constraint that the premises always involved the term *taller*, a subject seeking to determine whether B or E was the taller could scan the left and right hand sides of the premises looking for them. In the example above, B occurs on both sides, but E occurs only on the right side. It follows that provided the problem is determinate, B is the taller. There are, of course, a variety of such strategies; and the key point about them is that the reasoner at no time builds up a unified representation or mental picture of the array of items.

This development from a representational to a more economical nonrepresentational strategy is consistent with our hypothesis about three-term series problems. But there is a snag. The reconciliation of the two theories has become an embarrassment, because if a subject passes from an approach analogous to the IMAGE theory to an approach analogous to the LINGUISTIC theory, surely one would expect their predictions to differ. Why, indeed, do they make such similar predictions? The most natural explanation rests upon one of Wood's findings.

He invented a simple but elegant test to determine the nature of his subjects' approach to the problems. After a subject had solved a certain number of conventional problems, he would be given a special test problem in which, having answered the main question, he would be asked a further unexpected question such as 'who is taller A or D?'. These supplementary questions were so framed that they could be readily answered only by those subjects who had formed a unified representation of the premises. By varying the number of conventional problems encountered before the test problem, Wood was able to confirm that subjects began by using the representational strategy but rapidly abandoned it in favour of more specialized nonrepresentational procedures. What was particularly striking was the rapidity of this development. No doubt this was aided by the uniform content of the problems. Nevertheless, it is surprising that the biggest drop in the ability to answer the supplementary question was from those subjects who had previously encountered two conventional problems to those who had previously encountered three.

This finding suggests that subjects in the more orthodox studies of *three*-term series problems are likely to have abandoned the representational strategy fairly rapidly. These studies have tested subjects repeatedly, and hence both the IMAGE theory and the LINGUISTIC theory have almost certainly been based upon the performance of fairly experienced subjects. No wonder that despite their divergent assumptions they have tended to converge upon the same empirical predictions. They are both likely to have miscalculated slightly, with the IMAGE theory failing to be sufficiently 'naive' and the LINGUISTIC theory failing to be sufficiently 'sophisticated'. Both are perhaps guilty of 'regression towards the mean', and our final task is to offer some suggestion on how they might be reformulated to correct this bias.

6. A reformation of the two theories

It is extremely difficult to say how a unified representation is formed by an inexperienced individual. What is needed are studies concentrating on the subject's initial performance and perhaps abandoning the customary procedure of asking a specific question about the premises. This constrains the individual; and a more general question merely asking what follows from the premises might be more revealing about inferential strategy. Meanwhile, it is possible to make some tentative suggestions based on a number of small but interesting discrepancies in the results of the experiments in the literature.

First, it seems likely that Hunter's notion of a 'natural' order for premises is relevant, and that inexperienced subjects find it easier to represent premises in a unified form when they do follow one another naturally. For example, judging by the number of errors they made, the subjects tested by De Soto *et al.* (1965) seem to have been less experienced than those tested by Clark (1969b). De Soto's subjects found problems of the form 'A>B, B>C', easier than those of the form, 'B>C, A>B'; whereas Clark's subjects yielded the opposite results. The hypothesis is also corroborated by the finding (Handel *et al.*, 1968) that with only twenty problems to solve, those of the form, 'A>B, B>C', were easier than those of the form, 'A>B, C<B'. This conflicts with the performance of Huttenlocher's (1968) and Clark's (1969a, b) more practised subjects. It also conflicts with the performance of the younger children tested by Hunter (1959), probably because they were prey to 'atmosphere' effects, but is consistent with the performance of the older children.

Second, according to the LINGUISTIC theory, the nature of the question should have *no* effect on a problem of the form:

B<A
B>C

since the alternative questions are both incongruous with the representation of the items. However, Hunter's subjects found that 'who is best?' was reliably easier than 'who is worst?'. There are a number of possible explanations for this finding. Hunter himself suggests that the subjects may *revert* to the first premise and *convert* it, and that the answer to the question is particularly salient when it is in a premise that has been operated upon in this way. It is also possible that congruity between question and the adjacent second premise is more important than that between question and the first premise.

Finally, it is likely that the principle of lexical marking applies to both naive and experienced subject alike. This seems more parsimonious than the notion of a preferred direction of working, since it seems to apply to language behavior in general. (For example, marked comparatives are harder to remember than unmarked comparatives.) It would be premature to present an information-processing model of naive performance; but it is clear that it would differ from the IMAGE theory principally by incorporating an aspect of Hunter's OPERATIONAL theory and some sort of question-answering procedure.

The effect of practice is probably to induce a more 'mechanized' approach to the problem, which minimizes effort and which is appropriate to the particular constraints of the material. At the same time such an approach is likely to be less flexible and may make it harder to solve an unexpectedly novel type of problem. In characterizing the effect of practice upon drawing conclusions from problems involving five premises, Hunter (1957) drew attention to an important aspect of performance. He wrote:

> 'When the student first tackles such a problem, his activity is haphazard. He may combine a couple of premises, draw a part conclusion, leave it, combine a further pair of premises, draw a second part conclusion, and try to see if this can be combined with the first part conclusion... But after solving a few more problems of this type, his performance is characteristically transformed. He largely ignores the order in which the premises are presented: he reads through the statements in search of that one of the two terms which are not repeated and which he will take as his starting point: and from this starting point, he considers each premise in such a sequence that the terms form a consecutive chain with identical terms juxtaposed.'

Likewise, the most natural modification in solving three-term series problems is to read the question *before* reading the premises. Of course, subjects may glance fleetingly at the premises to obtain a global impression of them, but it is suggested that their detailed interpretation will be guided by the nature of the question. The procedure resembles working backwards from the conclusion of an inference to its premises, and its great advantage is that it often renders it unnecessary to examine more than one premise in any detail. Where only one premise is congruent with the question, then this premise will be processed first; and once it has been interpreted and the item which is *more*-X stored (where X is the relevant attribute), there is a simple time-saving procedure. If the item does not occur in the other premise, it is the solution to the problem. Where both premises have the same comparative, there is likely to be a natural tendency to interpret them in a standard order. The same technique of establishing the item which is *more*-X, where X is congruent with the question, can be used; and it is only necessary to interpret the other premise if this item is also mentioned in it.

It is a simple matter to express this sort of model in information-processing terms. A more pertinent issue is whether there is any evidence to support it. Ironically, although in principle it entails less processing than the LINGUISTIC model, it yields exactly the same differential predictions for those problems which have been studied so far.

One of the chief difficulties with the proposal that one approach to a problem is succeeded by another is to account for the process of transition. It is necessary to invoke a higher-order conceptual skill responsible for generating new strategies out of old (cf. Miller, Galanter, and Pribram, 1960). Yet it is not too difficult to envisage how this might occur. Wood (1969), for instance, suggests that a record is kept of all the procedures used in tackling a problem, and its eventual outcome. Hence, if a certain procedure always leads ultimately to a particular outcome, other intervening procedures may be dropped as redundant. In this way, for instance, a subject might learn that it was unnecessary to interpret the second premise in any detail if the item which is *more* X in the first premise does not occur in it. However, in addition to the relatively passive monitoring of processes, a more radical and active search mechanism may have to be postulated in order to account for genuine innovative changes in strategy.

If the present analysis of series problems is correct, then we may conclude that one of the most important, though neglected, independent variables in a cognitive task is the number of problems a subject is given to solve. Subjects think, not only in solving a problem, but also about how to solve it. They are likely to move from general and flexible procedures to more economical and specialized strategies. One can no longer ask how an individual solves a three-term series problem without asking when in his intellectual development *within the experiment* it was given to him.

REFERENCES

Burt, C. (1919) The development of reasoning in school children, *J. exper. Ped.* 5, 68–77 and 121–7.

Campbell, R.N. and Wales, R.J. (1969) Comparative structures in English. *J. Ling.* 5, 215–251.

Clark, H.H. (1969a) Linguistic processes in deductive reasoning. *Psychol. R.* 76, 387–404.

Clark, H.H. (1969b) The influence of language in solving three term series problems. *J. exper. Psychol.* 82, 205–215.

Clark, H.H. and Card, S.K. (1969) The role of semantics in remembering comparative sentences. *J. exper. Psychol.* 82, 545–553.

De Soto, C.B., London, M. and Handel, S. (1965) Social reasoning and spatial paralogic. *J. Pers. soc. Psychol.* 2, 513–521.

Donaldson, M. (1963) *A study of children's thinking.* London, Tavistock.

Donaldson, M. and Wales, R.J. (1970) On the acquisition of some relational terms. In J.R. Hayes (ed.): *Cognition and the development of language.* New York, Wiley.

Flores D'Arcais, G.B. (1970) Linguistic structure and focus of comparison in processing comparative sentences. In G.B. Flores d'Arcais and W.J.M. Levelt (eds.): *Advances in Psycholinguistics.* Amsterdam, North-Holland.

Handel, S., De Soto, C.B. and London, M. (1968) Reasoning and spatial representation. *J. verb. Learn. verb. Behav.* 7, 351–357.

Hunter, I.M.L. (1957) The solving of three-term series problems. *Brit. J. Psychol.* 48, 286–298.

Huttenlocher, J. (1968) Constructing spatial images: a strategy in reasoning. *Psychol. R.* 75, 550–560.

Huttenlocher, J., Eisenberg, K. and Strauss, S. (1968) Comprehension: relation between perceived actor and logical subject. *J. verb. Learn. verb. Behav.* 7, 527–530.

Huttenlocher, J. and Strauss, S. (1968) Comprehension and a statement's relation to the situation it describes. *J. verb. Learn. verb. Behav.* 7, 300–304.

Huttenlocher, J., Higgins, E.T., Milligan, C. and Kauffman, B. (1970) The mystery of the 'negative equative' construction. *J. verb. Learn. verb. Behav.* 9, 334–341.

James, W. (1890) *The principles of psychology*, Volume II. New York, Holt.

Jones, S. (1970) Visual and verbal processes in problem-solving. *Cognitive Psychol.* 1, 201–214.

Lees, R.B. (1961) Grammatical analysis of the English comparative construction. *Word* 17, 171–185.

Miller, G.A., Galanter, E. and Pribram, K.H. (1960) *Plans and the structure of behavior.* New York, Holt, Rinehart and Winston.

Piaget, J. (1921) Une forme verbal de la comparaison chez l'enfant. *Archiv. Psychol.* 18, 141–172.

Piaget, J. (1928) *Judgment and reasoning in the child.* London, Routledge and Kegan Paul.

Ross, J.R. and Perlmutter, D.M. (1970) A nonsource for comparatives. *Ling. Inquiry*, 1, 127–128.

Smith, C.S. (1961) A class of complex modifiers in English. *Language* 37, 342–365.

Wood, D.J. (1969) *The nature and development of problem-solving strategies.* Unpublished Ph.D. Thesis, University of Nottingham.

Woodworth, R.S. (1938) *Experimental Psychology.* New York, Holt

Résumé

Il existe trois théories apparemment contradictoires traitant des processus utilisés pour résoudre les problèmes à trois termes enchaînés, comme par exemple: 'A est mieux que B, B est mieux que C, quel est le meilleur?'. Nous présentons une analyse par traitement de l'information, qui montre que le modèle de l'IMAGE de De Soto, Handel et London (1965), peut facilement s'accorder au modèle LINGUISTIQUE de Clark (1969). Néanmoins, si l'on examine en détail les données empiriques, nous trouvons de légères incompatibilités qui suggèrent qu'un tel rapprochement est peut-être trompeur. Nous proposons donc l'hypothèse suivante: le sujet doit, au cours d'une séance expérimentale, changer son type d'approche. Initialement, il utiliserait une procédure analogue à celle proposée par le modèle de l'IMAGE, tout en la modifiant afin d'incorporer un des principes du modèle OPERATIONNEL de Hunter (1957). Par la suite, et après un certain entraînement, il développerait une procédure se rapprochant d'avantage du modèle LINGUISTIQUE.