

Relational processing in higher cognition: Implications for analogy, capacity and cognitive development.

Graeme S. Halford
University of Queensland

William H. Wilson
University of New South Wales
and

Steven Phillips
Electrotechnical Laboratory

Abstract

It is proposed that models based on processing relations capture the structure sensitivity of higher cognitive processes while they can also be compared with more basic processes such as associations. Relations have the following properties that are not shared by associations: there is an explicit symbol for each relational instance, allowing it to be manipulated, higher-order relations can be formed that have lower-order relations as arguments, given any N-1 components of an n-ary relation the remaining component can be retrieved (omni-directional access), and representation of relational instances is a prerequisite to analogical mapping. A model is proposed in which each component of a relational instance is represented by a vector, and the binding is represented by computing the outer product of the vectors. This architecture has been used to model analogy and human memory. It can also be used to model structural effects on both similarity and category formation. Computational cost increases exponentially with representational rank, defined as number of components that are bound into a representation. Thus the model provides a natural explanation for processing capacity limitations in humans and higher animals. Each rank corresponds to a class of psychological processes, neural nets, and empirical criteria. The ranks and typical concepts which belong to them, are: Rank 0, elemental association; Rank 1, content-specific representations and configural associations; Rank 2, unary relations, class membership, variable-constant bindings; Rank 3, binary relations, proportional analogies; Rank 4, ternary relations, transitivity and hierarchical classification; Rank 5, quaternary relations, proportion and the balance scale. Rank 6, quinary relations. Rank 0 can be performed by 2-layered nets, rank 1 by 3-layered nets, and ranks 2-6 by tensor products of the corresponding number of vectors. All animals with nervous systems perform rank 0, vertebrates perform rank 1, other primates perform rank 2-3, but ranks 4-6 are uniquely human. Rank also increases with age. Implications of this model are developed for human reasoning and cognitive development.

Paper presented to workshop on analogy, New Bulgarian University, Sofia, Bulgaria., July 17-20, 1998.

Author's address:
Graeme Halford

**School of Psychology
University of Queensland
Brisbane
4072
Australia**

**Phone: 61 7 3365 6401 (w) 61 7 3844 0183 (h)
Facsimile: 61 7 3365 4466**

Email: gsh@psy.uq.oz.au

WWW: <http://www.psy.uq.edu.au/people/Department/gsh/index.html>.

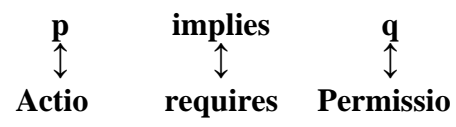
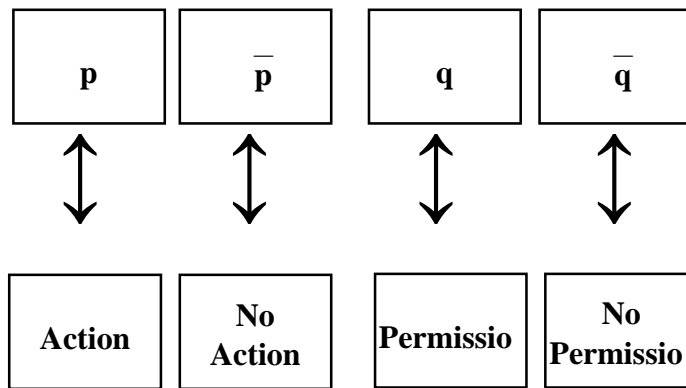
Relational processing in higher cognition: Implications for analogy, capacity and cognitive development.

In this paper we will present an outline of a theory that provides a general metric for cognitive complexity, and specifies properties of higher cognitive processes in a way that enables them to be distinguished systematically from more basic cognitive functions. The theory distinguishes the cognition of humans from other animals, distinguishes levels of cognitive development, and accounts for processing loads in cognitive tasks, within a common metric based on structural complexity. The levels of complexity are related systematically both to neural net architectures and to empirical criteria. Analogy has a central role in this theory, first because it is a core mechanism in higher cognition, and second because lower cognitive processes cannot implement analogy.

Although interest in analogy dates back to near the beginning of scientific psychology (Piaget, 1950; Spearman, 1923) understanding of human analogical reasoning accelerated dramatically in the 1980s (Gentner, 1983; Gick & Holyoak, 1983). Analogy is a natural mechanism for human reasoning, but we will suggest that its involvement in higher cognition might be even greater than previously realised. It has proven difficult to produce effective models of human reasoning based on logical inference rules. Such models do exist (Braine, 1978; Rips, 1989) but most theorists have chosen to model reasoning on the basis of alternative psychological mechanisms such as memory retrieval (Kahneman & Tversky, 1973) mental models (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) or pragmatic reasoning schemas (Cheng & Holyoak, 1985). Analogy can play a role in a human reasoning and is also entailed in some significant ways with a number of other models. We can illustrate this using pragmatic reasoning schemas.

Although it has become fashionable to interpret pragmatic reasoning schemas as being specialised for deontic reasoning, they may be more widely applicable. Consistent with this, we will use the definition of pragmatic reasoning schemas as structures of general validity that are induced from ordinary life experience. One type of pragmatic reasoning schema, permission, is known to improve performance on the Wason Selection Task (Cheng & Holyoak, 1985). In this task participants are given four cards containing p , $\sim p$, q , $\sim q$ and asked which cards must be turned over to test the rule $p \rightarrow q$. Analogy plays a central role here, because as Figure 1 shows, the elements and relations presented in the WST task can be mapped into a permission or prediction schema. This can be done by application of the principles that are incorporated in contemporary computational models of analogy (Falkenhainer, Forbus, & Gentner, 1989; Gray, Halford, Wilson, & Phillips, 1997; Hummel & Holyoak, 1997; Mitchell & Hofstadter, 1990) and no special mechanism is required.

Figure 1



A possible reason why induction of a permission schema improves performance is that, as Table 1 shows, permission is isomorphic to the conditional. Extending this argument, a possible reason for the tendency to respond in terms of the biconditional $p \leftrightarrow q$, is that participants may otherwise interpret the rule as a prediction. As Table 1 shows, prediction is isomorphic to the biconditional. This argument has been presented in more detail elsewhere (Halford, 1993). It implies that the importance of permission is not that it is deontic, but that it is isomorphic to implication. While we would not suggest that this argument does justice to the extensive literature on either the Wason Selection Task or pragmatic reasoning schemas, it does serve to illustrate that analogy can serve as the basic mechanism even in tasks such as WST that might normally be considered to entail logical reasoning.

Table 1
The Structure of the Permission Schema

<i>Permission schema</i>		<i>Action</i> \rightarrow <i>Permission</i>	<i>Permission Schema (Symbolic)</i>			<i>Conditional</i>			<i>Biconditional (Prediction)</i>		
			A	\bar{P}	$A \rightarrow P$	A	P	$A \rightarrow P$	p	q	$p \leftrightarrow q$
Action	permission	allowed	A	\bar{P}	+	1	1	1	1	1	1
Action	no permission	not allowed	A	\bar{P}	-	1	0	0	1	0	0
No action	permission	allowed	\bar{A}	\bar{P}	+	0	1	1	0	1	0
No action	no permission	allowed	\bar{A}	\bar{P}	+	0	0	1	0	0	1

Analogy, relations and higher cognitive processes

Although there are big differences between contemporary computational models of analogy, there is some degree of consensus about the core processes. In particular, it seems clear that analogy is a matter of mapping relations or relational instances between two representations. The core principles seem to be: the elements in one structure, the *base* are mapped uniquely to the elements of the other structure, the *target*, and; if a predicate P in the base is mapped to the predicate P' of the target, the arguments of P are mapped to the arguments of P'. The relational instances may be coded in the input (Gray et al., 1997; Hummel & Holyoak, 1997, Falkenhainer, 1989 #1136) or they may be constructed dynamically during the running of the model (Mitchell & Hofstadter, 1990) but a mapping between relational instances seems to constitute the essence of analogy in most models. It seems fair to say that an organism that could not represent relations or relational instances could not perform analogy. If we accept that analogy is one of the core processes in higher cognition, then ability to process relations and relational instances is also likely to be important in higher cognition. This is really an argument for the importance of structure in higher cognition, because relations are the essence of structure (a structure is a set on which one or more relations is defined).

Our next step is to consider those properties of higher cognitive processes on which there seems to be reasonable consensus. One such property is representation of structure, together with ability to operate on that structural representation. This is generally seen as the essence of higher cognition. The central role of structure in higher cognitive processes has been recognised historically (Humphrey, 1951) and by a number of writers in this century, including Gestaltists (Wertheimer, 1945), Piagetians (Piaget, 1950), information processing theorists (Anderson, 1983; Hunt, 1962; Miller, Galanter, & Pribram, 1960; Newell, 1990) and linguists (Chomsky, 1980; Fodor, 1975). One role of analogy is to form mappings between structures, so on these grounds also analogy might be considered a core process in higher cognition.

There is also reasonable consensus that higher cognitive processes entail variables, which are essential to the generality and content-independence that characterise higher cognitive processes.. An entire generation of cognitive models are based on rules, a distinguishing characteristic of which is that they relate variables. Production rules are perhaps the most common example (Anderson, 1983; Newell, 1990) and production systems normally have provision for variable binding. Smith, Langston and Nisbett (1992) make a case for logical inference rules being used in natural reasoning. These rules relate variables. For example, modus ponens is a logical inference rule of the form *if p then q, p therefore q*, where *p* and *q* are variables. Pragmatic reasoning schemas (Cheng & Holyoak, 1985) are more content-specific than abstract logical inference rules, but still relate variables. Thus the permission schema can be expressed as: to perform act *a*, you must have permission *p*.

Analogies can simulate variables by putting instances of a relation in correspondence with each other. Consider for example the following relational instances;

larger(whale,fish),
larger(horse,dog),
larger(5,3).

Each relational instance has two roles or slots, one filled by the larger entity and one by the smaller entity in a given pair. Because each role can be instantiated in a variety of ways, it effectively functions as a variable, but only if the arguments are in correspondence. It would not be true if the relational instances were cross-mapped, as in this case;

larger(whale,fish)
larger(dog,horse)

Models of analogy include mechanisms for ensuring structural correspondence. Indeed this is a core process in analogy models. Therefore they provide a mechanism that is capable of at least limited processing of variables.

Higher cognitive processes are widely regarded as incorporating symbols, even though the issue has become complicated by the debate between proponents of symbolic and connectionist models. Newell argued that symbols and a system that operates on them are necessary for intelligent action (see Newell, 1990, p. 170). Fodor and Pylyshyn (1988) argue that symbols are vital to cognition. Smolensky (1988), a connectionist modeller, does not deny the importance of symbols *per se*, but seeks to explain them at the subsymbolic level, rather than accepting them as a primary datum. With this proviso, there does seem to be widespread acceptance of the importance of symbols in higher cognitive processes.

Analogical reasoning mechanisms operate on relations that are symbolic in the sense that they include a label that specifies the link between the entities that are related. Thus in the instances considered above, the entities in the pairs (whale,fish) and (horse,dog) are linked by the relation symbol “larger”. Mathematically, an n -ary relation is a subset of the cartesian product of n sets, but the subset is typically specified by a label; for example, $\{ \dots (3,2), \dots, (5,1), \dots \}$. The existence of a label and an ordering over relational elements (i.e., $R(a,b)$ is not the same as $R(b,a)$) are important characteristics that distinguishes relations from other psychological structures such as associations, as we will argue later. We will briefly consider some further properties of higher cognition.

Compositionality has come to be accepted as a property of higher cognitive processes since the work of Fodor and Pylyshyn (1988). In essence it means that the

components of a cognitive representation retain their identity when they are composed into more complex representations, and both the components and the composites are semantically evaluable. As we will see, there are cognitive processes such as configural association, for which these properties do not hold.

Systematicity is another property that has been accepted as important in higher cognition since Fodor & Pylyshyn (1988) although it too has been the subject of some controversy (Niklasson & van Gelder, 1994; Van Gelder & Niklasson, 1994). In essence, it implies generalisation to all logically or structurally equivalent situations, although it is generally accepted that content can also influence performance, independent of structure, to some extent. Analogy clearly has the potential to be a core mechanism in achieving systematicity.

Categories are another property of higher cognition. We will not consider this complex topic here, except to say that categories must entail a label that is independent of content.

Modifiability on line is a property of higher cognitive processes that has been highlighted by the work of Clark and Karmiloff-Smith (1993). Higher cognitive processes should also be productive or generative, in the sense that they can produce or comprehend new sentences, can generate new representations, and make new inferences. This is true of both human and nonhuman primates, because apes show some inventiveness (Kohler, 1957) and ability to draw inferences (Tomasello & Call, 1997). Furthermore, though we will not pursue the question here, the approach we have adopted can model some limited forms of creativity (Halford, Wiles, Humphreys, & Wilson, 1993).

We do not include conscious awareness and language as criterial properties of higher cognition. Awareness has proven to be a difficult criterion to use, as the implicit learning literature has shown (Neal & Hesketh, 1997). As we wish to include some nonlinguistic, nonhuman species as having at least some forms of higher cognition, then language cannot be included either. We see conscious awareness and language as correlated rather than criterial properties of higher cognition.

We want to suggest that relational processing can capture the properties of higher cognition. Relations are preferable to rules, which have been used to model higher cognitive processes and to distinguish them from basic processes that have been characterised as associative (Sloman, 1996) or instance-based (Smith et al., 1992). Some cognitive representations such as loves(John,Mary) or contains(cup,drink) are not rules, but can be expressed as relations. The concept of n-ary relation is general enough to express any rule, it has the advantage of a precise mathematical definition, and effects of relational complexity on processing load are known -a{Blank or a = BBS, Which one b??}(Halford et al., in press). Relations are increasingly being utilised as the basis for models of higher cognitive processes. In addition to analogy, the importance of relational processing has been recognised in similarity (Markman & Gentner, 1993), induction (Lassaline, 1996), and

categorisation (Medin, 1989). Mental model theory, which can now account for a wide range of phenomena in human reasoning (Gentner & Stevens, 1983; Halford, 1993; Johnson-Laird, 1983; Polk & Newell, 1995), is based on representation of relations between entities. Phillips, Halford, and Wilson (1995) have argued that the associative-relational distinction can subsume the implicit-explicit distinction of Clark and Karmiloff-Smith (1993). Propositions, which are the core of some models of higher cognitive processes, can be treated as relational instances (Halford, Wilson, & Phillips, in press), section 2.2.2). For example the proposition loves(Joe,Jenny) is a relational instance.

Another big advantage of relations is that they can be compared directly with associations. The importance of this is that association has been accepted as a fundamental process in psychology virtually throughout the history of the discipline, and even many contemporary models incorporate it in one form or another. Therefore it is a disadvantage for associative and cognitive models to exist in conceptual worlds that do not communicate. We will suggest that basic processes, such as association, and higher cognitive process, which we identify with relations, can be incorporated into an overarching theory that integrates psychological processes at all levels. First however we will consider the properties of associative and relational knowledge in more detail.

Association

By contrast with higher cognition, association is not seen as inherently structural (Fodor & Pylyshyn, 1988; Humphrey, 1951). It differs from relational knowledge in a number of critical ways, one of which is that it is not symbolic. To illustrate, let us consider two commonplace relations, between cup and drink and between cup and saucer: i.e. contains(cup,drink) and placed-on(cup,saucer). The relation-symbols (or predicates) *contains* and *placed-on* specify the type of link represented, containment or superposition. Contrast this with associations; cup is associated with drink, and cup is associated with saucer. The associations *per se* do not specify the relations between cup and drink, or between cup and saucer, nor do these associations *per se* capture the fact that the relations are quite different. It is easy to overlook this because we know that a cup *contains* a drink and that a cup is *placed on* a saucer, so we tend to see this information in the associative link. The associative link is causal but does not capture the structure (Fodor & Pylyshyn, 1988). Associative links are unlabelled and all of the same kind, differing only in strength (Humphrey, 1951). The need for labelled links has been recognised however in models of higher cognitive structures such as propositional networks, in which links between nodes carry labels such as “agent”, “object”, “location”. An explicit symbol for a link therefore appears to be a property that distinguishes relational from associative processes. Our aim now is to define the properties of relational processes so that they capture the essence of higher cognition and can be compared directly with association.

Properties of relational processes

A relation that relates n entities, or n -ary relation is a subset of the cartesian product of n sets: i.e. $R(a_1, a_2, \dots, a_n)$ is a subset of $S_1 \times S_2 \times \dots \times S_n$. A relation is identified by the relation symbol, R , and the entities by argument symbols, a_1, a_2, \dots, a_n . For example the relation “larger” identifies a specific subset of a cartesian product, that subset in which the first

entity is always larger than the second; i.e. $a_1 > a_2$. There must be a binding between entities and arguments which preserves the truth of the relation; thus `contains(cup,drink)` is true but `contains(drink,cup)` is not.

Symbolisation, or an explicit label specifying the link, is a property of relations, but not of associations.

Higher-order relations have lower-order relations as arguments; e.g. in `cause(shout-at(Tom,John), hit(John,Tom))` `cause` is a higher-order relation, with `shout-at(Tom,John)` and `hit(John,Tom)` as arguments.

Systematicity means that relations imply other relations, and can be captured by higher-order relations; e.g. `>(a,b)` implies `<(b,a)`, can be written as the higher-order relation `implies(>(a,b),<(b,a))`.

Association does not share these properties. Associations can be chained, so that the output of one association is the input to another: $E_1 \rightarrow E_2 \rightarrow E_3 \dots \rightarrow E_n$, and may converge, so that E_1 and E_2 elicit E_3 , or diverge, so that E_1 elicits E_2 and E_3 . However associations are not identified by a symbol, and the associative link *per se* cannot be an argument to another association. Therefore the recursive, hierarchical structures that can be formed using higher-order relations do not appear to be possible with associations.

Compositionality means that the components of the relation, symbol and arguments, retain their identity when bound into a structure; e.g. in `larger(whale,dolphin)`, the components “larger”, “whale” and “dolphin” retain their identity when bound into the relation. This is not inherent in association, as we will see when we consider configural associations.

Modifiability by strategic processes, without information input, is possible for relations, whereas associations are modified incrementally on the basis of experience.

Omni-directional access means that, given all but one of the components of a relational instance, we can access (i.e. retrieve) the remaining component. For example, given the relational instance `mother-of(woman,child)`, and given `mother-of(woman,?)` we can access “child”, whereas given `mother-of(?,child)` we can access “woman”, and given `?(woman,child)` we can access “mother-of”. Although backward association may be possible, omni-directional access does not appear to be inherent in association.

Complexity can be defined by the “arity” or number of arguments of a relation (Halford et al., 1994; Halford et al., in press). Each argument corresponds to a source of variation or dimension, so an n -nary relation is a set of points in n -dimensional space. Dimensionality is related to processing load. Capacity is limited by the number of dimensions (or number of interacting variables) that can be processed in parallel. Data in the literature, and from our own laboratory, indicates quaternary relations (Rank 5) are the most complex that can be processed in parallel by most humans. Concepts too complex to be processed in parallel are handled by *segmentation* (decomposition into smaller segments that can be processed serially) and *conceptual chunking* (recoding representations into lower rank, but at the cost of making some relations inaccessible). For example, `velocity = distance/time`, is a ternary

relation, and is Rank 4, but can be recoded to rank 2, a binding between a variable and a constant (Halford et al., in press), Section 3.4.1). Difficulty can vary because of factors other than capacity, including declarative and procedural knowledge and amount of iteration (e.g. constructing a 5-term series from premises $a>b$, $b>c$, $c>d$, $d>e$ requires the integration process to be iterated 3 times; $a>b$, $b>c$ yields a,b,c , then this is integrated with $c>d$ to yield a,b,c,d , etc.).

In the next section we argue that each level of cognitive functioning can be assigned to an equivalence class of equal structural complexity, and that the classes can be ordered according to their complexity. They are ordered according to representational rank, defined as the number of components in cognitive representations, given that the components retain their identity when bound into more complex representations. An important feature of this idea is that the ranks correspond across the three domains of psychological process, neural net structure, and empirical observation. Each rank corresponds to a class of neural net architectures and can be identified by specific empirical criteria. It is an extension of a theory that defines processing capacity in terms of relational complexity (Halford et al., in press).

Representational rank

Representational rank corresponds to the number of components of a representation, given that the components retain their identity when bound in a more complex representation. The metric is shown in Figure 2, together with corresponding psychological processes and neural net architectures. The metric combines relational complexity with two nonstructural levels, elemental and configural association, enabling the basic properties of all levels of cognition to be defined within a single system. Rank = $n+1$ where n is the dimensionality or arity of a relation. We will now give an overview of the ranks.

Figure 2

Rank 0 corresponds to Elemental associations, which comprise links between pairs of entities: $E_1 \rightarrow E_2$

They are Rank 0 because there is no representation other than input and output, and they can be implemented by 2-layered nets. In principle Rank 0 can be assessed by any associative learning test, and because ability to perform at this level is not in question for vertebrates, or even for most invertebrates, no special assessment is intended.

Rank 1 corresponds to Configural associations, in which one cue is modified by another.

They have the form: $E_1, E_2 \rightarrow E_3$. An example is conditional discrimination, shown in Table 2. This cannot be acquired through elemental association, because of associative interference (each element, colour or shape, is equally associated with each outcome). They can be learned by fusing or "chunking" elements into a configuration such as "black/triangle". This avoids associative interference but at the cost that the components lose their

Cognitive Process	Neural net specification	Representational Rank
elemental association		0
configural association		1
unary relation		2
binary relation		3
ternary relation		4
quaternary relation		5
quinary relation		6

Figure 2. Ranks 0-6, with schematic neural nets. Input and output layers are omitted for Ranks 2-6.

identity, (e.g. “triangle” is not the same in “black/triangle” as in “white/triangle”) so the structure of the task is not represented. Thus the representation is holistic and nonstructural. Configural learning cannot be implemented by 2-layered nets (Minsky & Papert, 1969; note that conditional discrimination is isomorphic to exclusive-OR). They can be implemented with three-layered nets, by using units in the hidden layer to represent configurations of features such as "black&triangle " (Schmajuk & DiCarlo, 1992).

Ranks 2-6 are structural, and complexity increases with rank. We will consider the main properties of each rank.

Rank 2 corresponds to unary relations which are a binding between a relation symbol and an argument symbol. An example would be the proposition happy(John). Indicators of Rank 2 include symbolic representation of categories and understanding word reference.

Rank 3 corresponds to binary relations, which represent common states and actions in the world, such as larger(whale,dolphin), or loves(Joe,Jenny)..

Rank 4 corresponds to Ternary relations such as “love-triangle”, which is a relation between three people. They can be interpreted as bivariate functions, and binary operations. For example, the binary operation of arithmetic addition consists of the set of ordered triples of $+{\dots, (3,2,5), \dots, (5,3,8), \dots, \dots}$ and is a ternary relation. Many cognitive tasks that cause difficulty for young children, including transitivity and class inclusion, are ternary relations (Halford, 1993; Halford et al., in press)..

Rank 5 corresponds to quaternary relations. Proportion, $a/b = c/d$, is a quaternary relation. Comparison of moments on the balance scale (Siegler, 1981) is another example.

Rank 6 corresponds to quinary relations. Some complex reasoning tasks, such as categorical syllogisms and meta logical tasks, require Rank 6..

Table 2. Conditional discrimination, with isomorphic transfer task.

	Original task			Transfer task		
black	triangle	→	R+	green	circle	→ R+
black	square	→	R-	blue	cross	→ ?
white	triangle	→	R-	green	cross	→ ?
white	square	→	R+	blue	circle	→ ?

Neural net modeling of representational ranks

Neural nets can be rank-ordered according to the structural complexity of their internal representations (excluding input and output layers), and this rank ordering corresponds both to classes of psychological processes and to empirical criteria. Two-layered nets have no internal representation. Three-layered nets contain a representation that is computed from the input. While allowing that there are many variations, and potential for development, the representation in a typical three-layered net is “holistic” and is not structured in a way that meets the criteria for representation of relations. Three-layered nets can represent content-specific information and can form prototypes (Quinn & Johnson, 1997) but they lack compositionality and systematicity (Fodor & Pylyshyn, 1988; Phillips, 1994). They can only mediate transfer based on similar content (Marcus, submitted) and not between isomorphic structures with different contents (Phillips & Halford, 1997).

Nets that model higher cognitive processes should implement the properties of relational processes defined above. There are currently a number of competing models that can meet these criteria, discussed by (Halford et al., in press). In the model we will present

here, each relational instance is represented as a unique n-tuple, by representing bindings between relation symbol and arguments as outer products. Thus to represent loves(Joe,Jenny), each component, *loves*, *Joe* and *Jenny* is represented as a vector, and the binding is represented as the outer product of these vectors. The outer product corresponds to the binding units, shown for Rank 2 in Figure 1 but omitted for simplicity at higher ranks. Other instances of *loves* are represented in the same way, and can be summed to form a tensor product which represents the relation *loves* (Halford et al., in press, section 4.1.1.2). Thus loves(Joe,Jenny) and loves(Tom,Wendy) are represented as:

$$V_{\text{loves}} \otimes V_{\text{Joe}} \otimes V_{\text{Jenny}} + V_{\text{loves}} \otimes V_{\text{Tom}} \otimes V_{\text{Wendy}}$$

Neural net representations of relations from unary to quinary are shown schematically in the rightmost column of Figure 2. An n-ary relation is represented by the rank-n tensor, $V_R \otimes V_{a1} \otimes \dots \otimes V_{an}$. A unary relation such as happy(John) is represented by the outer product of vectors representing “happy” and “John”: $V_{\text{happy}} \otimes V_{\text{John}}$. In Figure 2 the two vectors are bound by a set of connections to a matrix of binding units. Rank 2 is the lowest structural level, but the transition from Rank 1 to Rank 2 can be envisaged by imagining the hidden layer at Rank 1 (Figure 2) being divided into two components which are then connected so as to form a matrix as shown for Rank 2. More complex relations are represented by tensor products of higher rank. A binary relation is represented by $V_R \otimes V_{a1} \otimes V_{a2}$, and so on. There is one component representing the symbol and one for each argument, so the representation of an n-ary relation has $n+1$ components. The components retain their identity, and the representations have the compositionality property. The model provides a natural explanation for empirical observations that cognitive processing load increases with relational complexity (Halford et al., in press, Section 5.). Representation of a relation of rank r with m units in each vector, requires m^r bindings units. The model implements all properties of relational knowledge (Halford et al., in press, Section 4.2) and is more efficient than models based on role-filler bindings for data bases in which relational instances are superimposed **in the sense that role-filler bindings require r units per relational instance, where symbol-argument bindings require 1 unit per instance** (Halford et al., in press, sections 2.2.1.2 and 4.1.3).

Associations, relations and analogy

It follows from this analysis that higher cognitive processes differ from associative processes in that the former entail representation and processing of structure. A task is cognitive to the extent that it entails a representation and processing of the structure of the task or situation. The representation should have the properties identified above. Representation of structure (relations) is essential to analogy, and this principle can be used to devise what is probably the most objective and straightforward test for cognitive processes.

The essential idea is that if the structure of a task is learned, it can be transferred to isomorphs using analogical mapping, and unknown items in the new task can be predicted. This principle has been applied successfully with tasks based on mathematical groups (Halford, Bain, Maybery, & Andrews, in press) but can be easily illustrated with the conditional discrimination task summarised in Table 2. Suppose someone has learned the original task. While this can be done by configural association, as noted above, configural discrimination does not lead to a representation of structure because the elements lose their

identity. However the task can also be learned by acquiring a representation of structure. The two modes of learning can be distinguished because only representation of structure enables transfer to isomorphs with prediction of new items. Notice that, in the transfer task in Table 2, once the first item is known and is mapped into the structure, the remaining three items can be easily predicted, irrespective of order of presentation.

Prediction of unknown items in an isomorphic task in this way requires analogical mapping, which in turn requires representation of structure. It is not possible if the task has been learned by configural association. Therefore transfer between isomorphs, with prediction of unseen items is a clearcut and objective measure of structural processing. It is a good way to assess higher cognitive processes. Notice too that it does not impose any extraneous task demands. The isomorphic task is assessed by the same procedure as the original task, and structure processing can be assessed by the number of correct items on the first trial of a new problem. It is not necessary to ask participants to describe the structure or to define rules, both of which impose an additional demand for articulation. We have been able to use this methodology successfully (Halford, 1980; Halford, Bain, et al., in press; Halford & Wilson, 1980) and have found that was related in a systematic way to other criteria.

Categories, structure and similarity

Although natural categories can be based on prototypes, prototypes do not represent structure (Medin, 1989). This problem can be overcome by forming prototypes based on relational instances. Relational instances such as Lives-in (chair, living room), Lives-in (vase, living room), Lives-in(couch, living room) can be represented as outer products of vectors and superimposed on a tensor product. The superimposed representation automatically averages features of the relational instances and corresponds to a prototype of living room furniture, but it also incorporates structure in the form of propositional information.

Similarity depends on more than common features, and is influenced by structure. For example grey hair is rated more similar to white hair than to black hair, whereas grey clouds are more similar to black clouds than to white clouds, because of our intuitive theories of ageing and weather respectively (Medin, 1989). Our model can handle similarity based both on elements and structure.

Element similarity can be assessed by computing the dot (inner) products of vectors representing two elements. If “desk”, “chair” and “vase” were coded by vectors representing sets of features, the dot products of vectors representing “desk” and “chair” would be higher than dot products of vectors representing “desk” and “vase”, reflecting greater similarity in the former pair.

Structural similarity can be handled by computing dot products of tensor products. The propositions feeds(soup-kitchen,woman) and feeds(woman,squirrel) have low similarity because “woman” occupies different roles. If we represent the propositions respectively as $\mathbf{v}_{\text{feeds}} \otimes \mathbf{v}_{\text{soup-kitchen}} \otimes \mathbf{v}_{\text{woman}}$ and $\mathbf{v}_{\text{feeds}} \otimes \mathbf{v}_{\text{woman}} \otimes \mathbf{v}_{\text{squirrel}}$, the dot products of these tensors will have a low value (expected value is zero with orthonormal vectors, low with sparse random vectors). This reflects the relational context, because woman is bound to soup-kitchen in one

case and squirrel in the other. However cases such as feeds(man,woman) and feeds(woman,man) are distinguished solely by the roles occupied by entities “woman” and “man”. We represent these in analogous fashion as $\mathbf{v}_{\text{feeds}} \otimes \mathbf{v}_{\text{man}} \otimes \mathbf{v}_{\text{woman}}$ **and** $\mathbf{v}_{\text{feeds}} \otimes \mathbf{v}_{\text{woman}} \otimes \mathbf{v}_{\text{man}}$. Dot products of these vectors will again be low, reflecting “man” and “woman” being in different roles. This occurs because dot products are computed so as to respect structural alignment (the elements of \mathbf{v}_{man} are multiplied by the elements of $\mathbf{v}_{\text{woman}}$, and vice-versa, giving the dot product an expected value of zero with orthonormal vectors, or a low value with sparse random vectors). This illustrates the sensitivity of the model to structural alignment.

Relational Context Similarity

The similarity of two items can be based on the degree to which they are used in the same relational context. For example, in the relational domain constructed around the items chair, desk and vase detailed above, chair and desk would achieve a high similarity as they both occur frequently in the same relational context (ie. Made_of(chair, wood) and Made_of(desk, wood), Stands_on(chair, floor) and Stands_on(desk, floor)). Chair and vase, however would achieve a lower similarity as they occur less frequently in the same relational context. Furthermore “woman” in feeds(soup-kitchen,woman) is dissimilar to “woman” in feeds(woman,squirrel) because the relational contexts are different, “soup-kitchen” in one case and “squirrel” in the other.

The relational context similarity of two items, a and b is computed as a normalised dot product of the rank 2 tensors retrieved from computing the dot product of each item’s vector against an appropriate dimension of the rank 3 tensor storing the relations¹. This can be applied to the hair-colour and cloud-colour examples above. We will represent a naive theory of ageing by propositions such as old-people-have(hair,gre), old-people-have(hair,white), young-people-have(hair,black), young-people-have(hair,brown) etc. These propositions can be superimposed on a tensor product representation. If we query this representation with “grey” we retrieve “old-people-have(hair,_)”. If we query it with “white” we retrieve “old-people-have(hair,_)”. The dot products of these tensors will be high, reflecting high similarity. However if we query the representation with “black” we retrieve “young-people-have(hair,_)” and the dot product of this with “old-people-have(hair,_)” is low.

By contrast, our knowledge of weather is represented by propositions threatening(clouds,gre), threatening(clouds,black), nonthreatening(clouds,white) etc. Querying with “grey” and “black” yields threatening(clouds,_) in both cases, with high dot products representing high similarity. Querying with “white” yields nonthreatening(clouds,_) which is dissimilar to threatening(clouds, _). Thus the model represents naive theories as sets of propositions coded in a tensor product. Relational context, as defined above, accounts for the effect of naive theories on similarity.

¹ This is illustrated with the rank 3 tensor as only binary relations are being stored. A higher rank tensor would allow ternary, quaternary, etc. relations to be stored.

Representational ranks are really points on a continuum, and limits on processing capacity are soft, so performance declines gracefully as the rank demanded by a task increases. It is proposed to model performances of intermediate rank, using the graceful degradation and graceful saturation properties of tensor products (Wilson & Halford, 1994).

Empirical indicators of ranks

Each rank has a unique set of empirical indicators. We will consider the main indicators for each rank.

Rank 0 is indicated by elemental association. Since this is evidently universal to all animals with nervous systems, no special predictions are made.

Rank 1 is best assessed by conditional discrimination. It is indicated in general by tasks that require content-specific representations. Representation of vanished objects and prototype formation both entail this requirement, and are performed by infants 3-6 months (Baillargeon, 1987). Consequently the theory predicts that with suitable testing and training techniques, infants of this age can acquire conditional discrimination. The significance of this can be seen from the fact that in the past children under five years have had great difficulty with this task (Rudy, 1991). Two further predictions follow. The first is that transfer to isomorphs of conditional discrimination will not be possible until a median age of five years. The second is that formation acquisition of conditional discrimination will be related to representation of vanished objects as assessed by Baillargeon (1987) and to prototype formation.

Rank 2 entails a relation-symbol that is independent of the entity to which it is bound, and is the simplest symbolic representation. Tasks that require this level of structure include:

Explicit category membership, such as *dog(Rover)*, where the category label *dog* is represented independently of the entity to which it is bound, *Rover*. As with all relations, the argument slot functions as a variable, and can be instantiated in a variety of ways such as *dog(Fido)*, *dog(Penny)* etc. Representation of explicit categories, in which there is a binding between a category symbol and instances of the category, seems to occur at approximately one year (Gershkoff-Stowe, Thal, Smith, & Namy, 1997; Sugarman, 1982).

Inferences about numerosity based on category membership Xu and Carey (1996).

Word comprehension, or understanding that words function as symbols for their referents.

Representing the binding between an object and its location, as assessed in the A-not B task (Halford, 1993, pp. 51-56; Wellman, Cross, & Bartsch, 1986).

Match-to-sample requires choosing an object that matches the sample (e.g. if shown an apple as sample, required to choose between an apple and a hammer). This task has been analysed by Premack (1983) and Halford et al. (in press) and is an analogy based on a unary relation. Transfer to an isomorphic task (e.g. the sample is a hammer, and the choices are a

banana and a hammer) demonstrates the principle is recognised independently of specific content.

This theory appears to be unique in predicting a correspondence between all five tasks.

Rank 3 entails symbolic processes based on binary relations, which develop at a median age of two years (Halford, 1993). Tasks that can be used to test this level of performance include:

Binary relational match-to-sample requires choice of a pair of objects that has the same relation as the sample (e.g. if the sample is XX, they should choose AA rather than BC. If the sample is XY, they should choose BC rather than AA). This implies a form of analogical reasoning based on binary relations, a Rank 3 representation (Gentner & Stevens, 1983; Halford et al., in press; Holyoak & Thagard, 1995).

Sorting into two categories can be assessed using the technique of Gershkoff-Stowe et al. (1997). Balance scale - weight and distance rules requires children to decide whether a beam should balance, or which side will go down, based on either weight or distance, with the other factor held constant [Halford, 1995 #2927]. This requires binary relations (Halford et al., in press, Section 6.3.1).

Rank 4 entails ternary relations. This level of structure is required for transitive inference, hierarchical classification, class inclusion, hypothesis testing, cardinality and comprehension of sentences (Andrews, 1996; Halford, 1993; Halford et al., in press). Other tests that require this level of structure include:

Transfer between isomorphs of conditional discrimination tasks with prediction of unseen items. Conditional discrimination has a well defined structure that can be assessed by transfer to isomorphs. As pointed out above, if the relations in the original task in Table 1 are learned, and given any one item of the isomorphic transfer task, the remaining three items can be predicted, irrespective of order of presentation. This is a case of analogical reasoning (Gentner & Stevens, 1983; Holyoak & Thagard, 1995) in which the structure of the original task (the base or source) is mapped into the transfer task (target). The structure of conditional discrimination is basically a ternary relation, in that it consists of ordered 3-tuples (e.g. colour,shape,response). Therefore, while original learning can be used to infer nothing more complex than configural association (*Rank 1*), prediction of unseen items of a new isomorphic transfer task reflects processing ternary relations (*Rank 4*). The same paradigm can be used to assess two different levels of cognitive process, with procedure held constant and without additional demands such as articulation. Infants should be able to learn the original discrimination but should not be able to predict unseen items on the isomorphic transfer task. Five year olds should be able to do both. These predictions are more optimistic than previous findings that conditional discrimination is not learned before age 5 (Gollin, 1966; Rudy, 1991).

The tendency to prefer reversal over nonreversal shifts (Kendler, 1995). Ability to make efficient reversal shifts in multidimensional discrimination problems was first analysed in detail by Kendler and Kendler (1962) and there is a long history of research (see review by

Kendler, 1995, and commentary by Halford, 1997). Reversal shifts depend on representation of the relevant dimension, which requires processing a ternary relation, because a dimension is a set on which an asymmetric, transitive relation is defined. Representation of a dimension requires induction of a relational schema (Halford, Bain, et al., in press). Consequently this longstanding enigma can be explained as a form of relational processing. Many predictions follow from this, but the one on which we will focus here is that preference of reversal shifts should correspond to other ternary relations tasks.

Ranks 5 and 6 entail quaternary and quinary relations respectively. Rank 5 is typically understood at age 11 (Halford, 1993), but there is virtually no useable data on Rank 6, though it is believed to occur only in a minority of adults. However we will consider two tasks that appear to require this level of processing, but have not been analysed in this way before.

Relational processes in reasoning

In this section we will consider how relational processes could be involved in two reasoning tasks, knights and knaves and categorical syllogisms.

Knights and knaves problems are based on the following scenario. Suppose there is an island where there are just two sorts of inhabitants - *knights* who always tell the truth and *knaves* who always lie. An example problem is: *A says "I am a knave and B is a knave". B says, "A is a knave". What is the status of A and B: Knight, knave, or impossible to tell?* (Rips, 1989, pp. 85-86). The solution entails two or more steps, but we focus on the step that requires the highest relational complexity: If we assume *A* is a knight, then *A*'s statement that *A* and *B* are knaves must be true, but *A* says *A* is a knave, which is a contradiction. Therefore *A* must be a knave. Symbolically:

$$kt(\underline{A}) \text{ and says}(\underline{A},(kv(\underline{A}) \text{ and } kv(\underline{B}))) \rightarrow kv(\underline{A}).$$

Using the type of analysis developed by Halford et al. (in press-a) there are five variables in this expression, corresponding to the five underlined arguments. Therefore this inference is quinary. The second step is to reason that if it is false that *A* and *B* are knaves, and that *A* is a knave, then *B* must be a knight: $false(kv(\underline{A}) \text{ and } kv(\underline{B})) \text{ and } kv(\underline{A}) \rightarrow kt(\underline{B})$. This step is quaternary, so task complexity, defined by the most complex step, is quinary.

Categorical syllogism tasks have been more extensively investigated, but we will focus on the following example tasks:

All *A* are *B*, all *B* are *C*. This would be represented by Johnson-Laird & Byrne (1991, Table 6.1) as the mental model: $[[a]b]c$. This mental model can be expressed as a relation between the following classes of entities (where $\neg A$ means "not *A*"): ABC , $\neg ABC$, $\neg A\neg BC$. We can think of this as follows: There is one class of entities with properties *A*, *B* and *C*, another class with properties not-*A*, *B* and *C*, and another class with properties not -*A*, not -*B* and *C*. The mental model that relates these three classes has the complexity of a ternary relation. Now consider the syllogism:

Some A are B , No B are C . The premises express a relation between the following classes of entities: $A \rightarrow BC$, $A \rightarrow B \rightarrow C$, $AB \rightarrow C$, $\neg A \rightarrow BC$, $\neg AB \rightarrow C$ (c.f. J-L&B, 1991, Table 6.1). The problem relates 5 classes of entities, so it has the complexity of a quinary relation. J-L&B define complexity in terms of the number of mental models required for a problem. The first problem above requires one model and is easy (88% correct) while the second requires 3 models and is difficult (38% correct). However more difficult problems tend to entail more complex relations. Of the 27 syllogisms with valid conclusions, there are 7 with ternary relations that entail 1 mental model, and 17 with relations more complex than ternary that entail more than 1 mental model (contingency coefficient $C = .61$). Therefore the relational complexity metric has potential to provide an alternative explanation to number of mental models for difficulty of categorical syllogisms.

Conclusion

We wish to propose that the representation and processing of structure, including analogical mapping, are core processes in higher cognition. They can be used as criteria for distinguishing tasks that demand higher cognitive processes from those that can be performed by more basic processes. Ability to form analogies can also be used as criterion for neural net models of higher cognitive processes. The relational complexity metric permits levels of structure to be distinguished.

Cognitive tasks can be grouped into equivalence classes of equal structural complexity, and the classes can be ordered according to representational rank. Ranks 0 and 1 are associative, do not entail explicit representation of structure, and do not enable analogical mappings to be made. Ranks 2-6 entail explicit representation of relations, from unary to quinary. In general they have the properties normally attributed to higher cognitive processes. There is a correspondence between three domains: level of structural complexity, neural net architecture, and observable properties of performance.

Reference List

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Andrews, G. (1996, August 1996). Assessment of relational reasoning in children aged 4 to 8 years. Paper presented at the Conference of the International Society for Study of Behavioral Development, Quebec City.

Baillargeon, R. (1987). Young infants' reasoning about the physical and spatial properties of a hidden object. *Cognitive Development*, 2, 179-200.

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, 85, 1-21.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391-416.

Chomsky, N. (1980). Rules and representations. The Behavioral and Brain Sciences, 3, 1-61.

Clark, A., & Karmiloff-Smith, A. (1993). The cognizer's innards: A psychological and philosophical perspective on the development of thought. Mind and Language, 8(4), 487-519.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. Artificial Intelligence, 41, 1-63.

Fodor, J. A. (1975). The language of thought. New York: Thomas Y. Crowell Company.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. Cognition, 28, 3-71.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. Cognitive Science, 7, 155-170.

Gentner, D., & Stevens, A. L. (1983). Mental models. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gershkoff-Stowe, L., Thal, D. J., Smith, L. B., & Namy, L. L. (1997). Categorization and its developmental relation to early language. Child Development, 68, 843-859.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Gollin, E. S. (1966). Solution of conditional discrimination problems by young children. Journal of Comparative and Physiological Psychology, 62, 454-456.

Gray, B., Halford, G. S., Wilson, W. H., & Phillips, S. (1997, September 26-28). A Neural Net Model for Mapping Hierarchically Structured Analogs. Proceedings of the Fourth conference of the Australasian Cognitive Science Society, University of Newcastle.

Halford, G. S. (1980). A learning set approach to multiple classification: Evidence for a theory of cognitive levels. International Journal of Behavioral Development, 3, 409-422.

Halford, G. S. (1993). Children's understanding: the development of mental models. Hillsdale, NJ: Erlbaum.

Halford, G. S. (1997). Review of Levels of Cognitive Development. Tracey S. Kendler. Mahwah, NJ: Erlbaum, 1995. Pp vii + 187. Merrill-Palmer Quarterly, 43(4), 694-699.

Halford, G. S., Bain, J. D., Maybery, M. T., & Andrews, G. (in press). Induction of relational schemas: Common processes in reasoning and learning set acquisition. Cognitive Psychology.

Halford, G. S., Wiles, J., Humphreys, M. S., & Wilson, W. H. (Eds.). (1993). Parallel distributed processing approaches to creative reasoning: Tensor models of memory and analogy. Menlo Park, CA: AAAI Press.

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

Halford, G. S., Wilson, W. H., Guo, J., Gayler, R. W., Wiles, J., & Stewart, J. E. M. (1994). Connectionist implications for processing capacity limitations in analogies. In K. J. Holyoak & J. Barnden (Eds.), Advances in connectionist and neural computation theory, Vol. 2: Analogical connections (pp. 363-415). Norwood, NJ: Ablex.

Halford, G. S., Wilson, W. H., & Phillips, S. (in press). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. Behavioral and Brain Sciences.

Holyoak, K. J., & Thagard, P. (1995). Mental leaps. Cambridge, MA: MIT Press.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. Psychological Review, 104, 427-466.

Humphrey, G. (1951). Thinking: An introduction to its experimental psychology. London: Methuen.

Hunt, E. B. (1962). Concept learning: An information processing problem. New York: Wiley.

Johnson-Laird, P.N. (1983). Mental models. Cambridge: Cambridge University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). Deduction. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. Psychological Review, 80(4), 237-251.

Kendler, H. H., & Kendler, T. S. (1962). Vertical and horizontal processes in problem solving. Psychological Review, 69, 1-16.

Kendler, T. S. (1995). Levels of cognitive development. Mahwah, NJ: Erlbaum.

Kohler. (1957). The Mentality of Apes (2nd rev. ed. Ella Winter, Trans). Harmondsworth: Penguin.

Lassaline, M. (1996). Structural alignment in induction and similarity. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(3), 754-770.

Marcus, G. F. (submitted). Rethinking eliminative connectionism.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology*, 25(4), 431-467.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, 44(12), 1469-1481.

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

Minsky, M., & Papert, S. (1969). *Perceptrons: An introduction to computational geometry*. Cambridge, MA.: MIT Press.

Mitchell, M., & Hofstadter, D. R. (1990). The emergence of understanding in a computer model of concepts and analogy-making. *Physica D*, 42(1-3), 322-334.

Neal, A., & Hesketh, B. (1997). Future directions for implicit learning: Toward a clarification of issues associated with knowledge representation and consciousness. *Psychonomic Bulletin and Review*, 4(1), 73-78.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Niklasson, L., & van Gelder, T. (1994,). Can connectionist models exhibit non-classical structure sensitivity? *Sixteenth Annual Conference of the Cognitive Science Society, Atlanta, Georgia*, 664-669.

Phillips, S. (1994). Strong systematicity within connectionism: The tensor-recurrent network. *Sixteenth Annual Conference of the Cognitive Science Society*, 723-727.

Phillips, S., & Halford, G. S. (1997,). Systematicity: Psychological evidence with connectionist implications. *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society, Stanford University*, 614-619.

Phillips, S., Halford, G. S., & Wilson, W. H. (1995, July). The processing of associations versus the processing of relations and symbols: A systematic comparison. *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society, Pittsburgh, PA*, 688-691.

Piaget, J. (1950). *The psychology of intelligence*. (M. Piercy & D. E. Berlyne, Trans.) London: Routledge & Kegan Paul, (Original work published 1947).

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, 102, 533-566.

Premack, D. (1983). The codes of man and beasts. *The Behavioral and Brain Sciences*, 6, 125-167.

Quinn, P. C., & Johnson, M. H. (1997). The emergence of perceptual category representations in young infants: A connectionist analysis. *Journal of Experimental Child Psychology*, 66, 236-263.

- Rips, L. J. (1989). The psychology of knights and knaves. Cognition, 31, 85-116.
- Rudy, J. W. (1991). Elemental and configural associations, the hippocampus and development. Developmental Psychobiology, 24(4), 221-236.
- Schmajuk, N. A., & DiCarlo, J. J. (1992). Stimulus configuration, classical conditioning, and hippocampal function. Psychological Review, 99, 268-305.
- Siegler, R. S. (1981). Developmental sequences within and between concepts. Monographs of the Society for Research in Child Development, 46, 1-84.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. Psychological Bulletin, 119, 3-22.
- Smith, E. E., Langston, C., & Nisbett, R. E. (1992). The case for rules in reasoning. Cognitive Science, 16, 1-40.
- Smolensky, P. (1988). On the proper treatment of connectionism. Behavioral and Brain Sciences, 11(1), 1-74.
- Spearman, C. E. (1923). The nature of intelligence and the principles of cognition. London: MacMillan.
- Sugarman, S. (1982). Developmental change in early representational intelligence: Evidence from spatial classification strategies and related verbal expressions. Cognitive Psychology, 14, 410-449.
- Tomasello, M., & Call, J. (1997). Primate cognition. New York: Oxford University Press.
- Van Gelder, T., & Niklasson, L. (1994,). Classicalism and cognitive architecture. Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society, Atlanta, Georgia, 905-909.
- Wellman, H. M., Cross, D., & Bartsch, K. (1986). Infant search and object permanence: A meta-analysis of the A-not-B error. Monographs of the Society for Research in Child Development, 51.
- Wertheimer, M. (1945). Productive thinking. New York: Harper.
- Wilson, W. H., & Halford, G. S. (1994,). Robustness of tensor product networks using distributed representations. Proceedings of the Fifth Australian Conference on Neural Networks, Brisbane, 258-261.
- Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. Cognitive Psychology, 30, 111-153.