Cognitive Development 27 (2012) 481–499

Contents lists available at SciVerse ScienceDirect

Cognitive Development

# Computational models of relational processes in cognitive development

Graeme S. Halford[a,*], Glenda Andrews[b], William H. Wilson[c], Steven Phillips[d]

[a] School of Applied Psychology, Griffith University, Nathan, 4111, Australia
[b] School of Applied Psychology, Griffith University, Gold Coast, 4222 Australia
[c] School of Computer Science & Engineering, The University of New South Wales, Sydney, 2052, Australia
[d] Human Technology Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba Central 2, Tsukuba Science City, Ibaraki, 305-8568, Japan

ARTICLE INFO

*Keywords:*
Computational models
Connectionist models
Self-modifying production system
Strategy acquisition
Relational knowledge

ABSTRACT

Acquisition of relational knowledge is a core process in cognitive development. Relational knowledge is dynamic and flexible, entails structure-consistent mappings between representations, has properties of compositionality and systematicity, and depends on binding in working memory. We review three types of computational models relevant to relational knowledge. The first are formal models of structural commonalities among concepts, including some that differ in surface characteristics. The second is a self-modifying production system model of the role of relational knowledge in strategy acquisition. The third comprises symbolic connectionist models that implement key properties of relational cognition. These models are complemented by the semantic cognition model that shows how some developmentally important concept acquisition mechanisms can emerge from learning input–output functions. We conclude that no one type of model fully suffices as an account of cognitive development but there is potential for future development, including hybrid models that could meet most or all of the criteria.



* Corresponding author.
*E-mail addresses:* g.halford@griffith.edu.au (G.S. Halford), g.andrews@griffith.edu.au (G. Andrews), billw@cse.unsw.edu.au (W.H. Wilson), steve@ni.aist.go.jp (S. Phillips).
*URLs:* http://www.psy.uq.edu.au/ gsh (G.S. Halford), http://www.griffith.edu.au/health/school-psychology/staff/dr-glenda-andrews (G. Andrews), http://www.cse.unsw.edu.au/ billw (W.H. Wilson), http://staff.aist.go.jp/steven.phillips (S. Phillips).

0885-2014/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cogdev.2012.08.003

The computational models outlined in this article are designed to implement a theory of development of higher cognitive processes that are characterised as symbolic or analytic and are distinguished by requiring "... access to a single, capacity-limited central working memory resource, ..." (Evans, 2008, p. 270). We propose that relational knowledge provides a conceptual basis for higher cognition (Gentner, 2010; Halford, Wilson, & Phillips, 2010), so our models are designed to implement relational processing. Although our theory is wide ranging, we use transitive inference as our reference task because it is basic to all inference (James, 1890) and there exists a large, high-quality database on it. We distinguish implicit and explicit transitive inference (Goel, 2007); only the latter entails relational representations formed in working memory (Halford, Maybery, & Bain, 1986; Maybery, Bain, & Halford, 1986).

Symbolic processes require a structured operating system to give them meaning (Newell, 1980), just as the natural numbers 1, 2, ... 5, ... 9, ... are given meaning initially by counting and later by other operations such as addition and multiplication. The requirement for structure is consistent with a number of theoretical positions, including gestalt (Wertheimer, 1945), Piagetian developmental (Piaget, 1950), and psychometric (Spearman, 1923). Relations are the essence of structure because, mathematically, a structure is a set of elements on which one or more relations is defined. The theory of relational knowledge has now been applied to analogy (Gentner, 2010; Gick & Holyoak, 1980; Kokinov, Holyoak, & Gentner, 2009), reasoning (see especially mental models theory; Goodwin & Johnson-Laird, 2005; Johnson-Laird, 2005), categorisation (Zielinski, Goodwin, & Halford, 2010) and cognitive complexity (Halford, Wilson, & Phillips, 1998). Relational knowledge processes also play a role in language acquisition (Gentner, 2010; Golinkoff & Hirsh-Pasek, 2008; Naigles, Hoff, & Vear, 2009; Tomasello & Brandt, 2009). For example, the representation of verbs includes an argument structure with inherently relational slots such as *agent*, *patient*, *instrument*.

Relational knowledge has partly replaced logic as a basis for human reasoning. Logic was once considered as constituting the laws of thought (Boole, 1854). Piaget (1950) proposed development through progressively more elaborate psycho-logics as the basis of cognitive development. However, the difficulty of accounting for thought on the basis of logic gave rise to alternatives including information processing models (Anderson, 1991; Andrews & Halford, 2011), heuristics (Kahneman, Slovic, & Tversky, 1982), mental models (Goodwin & Johnson-Laird, 2005; Johnson-Laird, 2005), Bayesian rationality (Oaksford & Chater, 2007), and analogy (Gentner, 2010; Halford, 1993). Relational knowledge has been proposed as an alternative to logic in accounting for cognitive development (Andrews & Halford, 2011; Halford & Andrews, 2004). Relational knowledge can support probabilistic inferences (Halford et al., 2010) and is therefore consistent with Bayesian approaches (Oaksford & Chater, 2007), but the criterion for relational knowledge is representation of relations with the properties outlined later.

Cognitive complexity can be accounted for by the number of entities related in a single representation (Halford, Bain, Maybery, & Andrews, 1998; Halford, Wilson, et al., 1998). The relational complexity (RC) metric has been applied to cognitive development (Halford, 1993), logical inference (Zielinski et al., 2010), mathematics education (English & Halford, 1995), human factors (Boag, Neal, Loft, & Halford, 2006), language (Andrews, Birney, & Halford, 2006), and cognitive neuroscience (Christoff & Owen, 2006; Kroger et al., 2002). The RC metric has also made it possible to predict previously unrecognised cognitive capacities, as in the balance scale (Halford, Andrews, Dalton, Boag, & Zielinski, 2002). Halford, Cowan, and Andrews (2007) present a method for analysis of relational complexity that provides details of assessment criteria.

A correspondence across tasks of equal structural complexity exists, even when they belong to different domains. Andrews and Halford (2002) assessed transitivity, hierarchical classification, class inclusion, cardinality and sentence comprehension in children aged 3–8 years. In each case, tasks of higher relational complexity were more difficult and attained at a later age than closely matched tasks of lower relational complexity. Relational complexity accounted for 80% of age-related variance in fluid intelligence. Tasks of equal relational complexity constituted an equivalence class that transcended task variables and content domains. The structural similarity of transitivity and class inclusion tasks is portrayed in Fig. 1. Thus, relational complexity defines a deep structural property of cognitive processes. By contrast, Piagetian tasks have generally been found not to be equivalent across

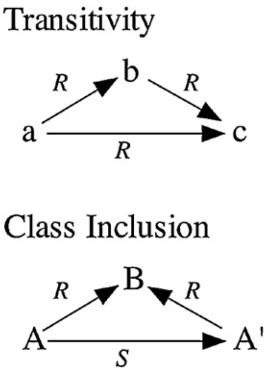


**Fig. 1.** Basic structure of transitivity and class inclusion.

domains, even though Piaget's specific observations were replicated (Halford & Andrews, 2006). This deep structural property is one of the phenomena that we model.

The concept of relational knowledge has been elaborated to the point where it arguably provides the foundation for higher cognition (Halford et al., 2010). We also have proposed that the associative – relational distinction corresponds to the implicit–explicit distinction (Karmiloff-Smith, 1992) that has been of considerable importance in cognitive development (Phillips, Halford, & Wilson, 1995). Therefore we propose that the nature of relational knowledge, and the processes by which it is acquired, are at the core of cognitive development. In this article we review our attempts to model these processes.

## 1. Core properties of relational knowledge

We first outline the properties of relational knowledge that we have attempted to model.

### 1.1. Structure-consistent mapping between representations

This property is a fundamental one (Halford, 1993; Halford & Andrews, 2004). It is the basis of analogy (Gentner, 2010; Holyoak & Thagard, 1989) and has been identified as the factor that best distinguishes human cognition from that of other animals (Penn, Holyoak, & Povinelli, 2008). Quinian bootstrapping, proposed by Carey (2009) as a major transition process, arguably entails structure-consistent mapping in some form.

Structure-consistent mapping between representations has long been known to be influenced by the capacity to represent complex relations (Halford & Wilson, 1980). Children aged 4–6 years were taught structures consisting of movements between the corners of a square array. They were then required to map the structure into an isomorphic transfer task, and the amount of information required was manipulated. Only the 5–6-year-olds could make mappings based on more complex information.

A formal model of structured knowledge was developed based on category theory, a branch of mathematics that defines structures in terms of transformations rather than objects. We consider it in more detail later with respect to subsequent work by Phillips, Wilson, and Halford (2009) but, as applied to the task used by Halford and Wilson (1980), it means that the structure is composed of movements between elements, rather than the elements themselves. It therefore corresponds to acquisition of an abstract concept from specific examples, a phenomenon investigated in subsequent studies (Halford, Bain, et al., 1998; Halford & Busby, 2007).

### 1.2. Binding between a relation symbol and arguments

A relational representation is a binding between a relation symbol and a set of ordered tuples of elements, corresponding to the arguments of the relation. Thus the relation *larger_than* can be represented as a binding between the symbol *larger_than* and ordered pairs representing instances where one entity is larger than another, such as "whale larger_than fish", "ship larger_than canoe." These instances of this relation are written as "larger_than{. . . (whale, fish), . . ., (ship, canoe), . . .}." The symbol *larger_than* distinguishes the type of link between the entities (e.g., whale and fish are related by size-difference in this context, but might share other relationships, such as same phylum). Therefore, relational instances are more than associative links (Halford et al., 2010). The symbol might not be a conventional element such as a word, and in some cases it might be an internal stimulus that has become a symbol by binding to relational instances. For example, after noting that certain entities are larger than others, an infant might adopt a gesture such as raising the hand to indicate what is different about pairs of elements of varying size. Propositions correspond to relational instances, so "larger_than(whale, fish)" is a proposition and is an instance of the *larger than* relation.

### 1.3. Dynamic and flexible representations

This property of relational knowledge is illustrated by explicit transitivity (Goel, 2007; Halford, 1993). For example, premises *Tom is taller_than Peter* and *Bob is taller_than Tom* can be integrated into the ordered triple *Bob taller_than Tom taller_than Peter* from which the inference *Bob taller_than Peter* can be read off (Andrews & Halford, 2002). This integration can be performed by mapping to an ordering schema such as *top-above-middle-above-bottom*, which functions as an analog of *tallest-middle-shortest* (Halford, 1993; Halford et al., 2010, Fig. 2A). There is a structural correspondence between top-above-middle-above-bottom and Bob-taller-than-Tom-taller-than-Peter (e.g., *above* consistently corresponds to *taller_than*) and the validity of the mapping is determined by relations between elements, rather than by semantic cues. Explicit transitive inference is based on dynamic mapping to a schema, a type of coordinate system, in working memory. The representations may be fuzzy (Brainerd & Reyna, 2001) but the relational properties still apply. For example, if the premises provide no semantic basis for mapping, structural correspondence must be processed in some form.

The mapping of elements Bob, Tom, Peter to positions above, middle, bottom in the ordering schema is dynamic and flexible. The mappings serve the current task of making an inference; they are not long-term acquisitions and are not acquired incrementally. By contrast, implicit transitivity (Goel, 2007) is based on incremental knowledge acquisitions in long-term memory. For example, McGonigle and Chalmers (1977) trained squirrel monkeys to choose one member of each pair in a series (A+B−, B+C−, C+D−, D+E−), where [+] indicates a rewarded choice and [−] a nonrewarded choice). Monkeys showed a 90% preference for B over D, though this pair was untrained. This inference is based on incrementally acquired knowledge (Couvillon & Bitterman, 1992; Markovits & Dumas, 1992; Wynne, 1995). It is a very different type of process from explicit transitive inference and it cannot be assumed that the tasks represent equivalent attainments.

### 1.4. Compositionality and systematicity

These have been proposed as fundamental properties of cognition (Fodor & Pylyshyn, 1988; but see also Van Gelder & Niklasson, 1994). Compositionality indicates that the entities in a compound representation retain their identity and are accessible, e.g., given "larger_than(whale, fish)" we can access elements in this representation by asking, "What is larger than a fish?" (whale), "What is a whale larger than?" (fish), and "What is the relation between whale and fish?" (larger_than).

Systematicity was defined by Fodor and Pylyshyn (1988). Halford et al. (2010) commented on systematicity in terms of a relational framework. In essence it means that certain cognitive capacities are intrinsically connected through their structural properties. Thus a proposition like "larger_than(___, ___)" has an intrinsic structure comprising a slot for a larger entity and a slot for a smaller entity; both slots can be filled in a variety of ways. This property means that if we understand that a whale is larger than a fish, we can understand other instances of a relation because the structure, consisting of a binding between relation symbol-bound slots, is available.

### 1.5. Higher-order representation

An important property of relations is that they can relate other representations (Gentner, 2010). For example, "Mary believes that Ken resents that Jim gave Mary a ticket," entails a series of relational instances (propositions) embedded in higher-order relations. Higher-order relations play a major role in cognitive complexity and control theory (Zelazo, Müller, Frye, & Marcovitch, 2003).

### 1.6. Working memory (WM)

The use of relational knowledge in a dynamic and flexible way, as required in higher cognition, means that it must be possible to form representations in WM based on both structural correspondence and semantic information (Blanchette & Dunbar, 2000; Halford et al., 2010). Earlier conceptions of WM (Daneman & Carpenter, 1980) were based on a combination of processing and storage, but a more recent conception (Oberauer, 2009, chap. 2) is based on dynamic binding to a coordinate system. Dynamic formation of relational representations also entails mapping to a coordinate system, as illustrated earlier with explicit transitive inference. There is therefore a close correspondence between our theory of relational knowledge and Oberauer's conception of WM.

The relational complexity metric measures the number of entities that are related in a cognitive representation (Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998; Halford, Bunch, & McCredden, 2007; Halford, Cowan, et al., 2007), as illustrated by the explicit transitive inference task in which premises *Tom is taller than Peter* and *Bob is taller than Tom* are integrated into a mental model in which the elements *Bob*, *Tom*, *Peter* are ordered according to the relation specified in the premises. This is equivalent to the ternary relation, *monotonically_taller*(*Bob*, *Tom*, *Peter*) relating three entities. By contrast, the premises *taller*(*Bob*, *Tom*) and *taller*(*Tom*, *Peter*) are binary relations.

However complexity can be reduced by segmentation (decomposition into smaller segments that can be processed serially) and conceptual chunking (recoding into representations of lower relational complexity), but there is a loss of access to relations between chunked or segmented variables. For example, *speed* = *distance/time* relates three variables and is ternary relational, but we can chunk speed into a single variable (e.g., our car's speed = 60 kph). However, the relation between speed, distance and time is not accessible in the single-variable representation so that, in order to answer a question such as: "How is speed affected if we travel the same distance in half the time?" we must return to the three-variable representation.

Performance is influenced by both WM capacity and knowledge, and the former has most influence when task structure constrains decomposition (Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998; Halford, Bunch, et al., 2007), whereas the latter has most influence where tasks can be decomposed into simpler subtasks, enabling serial processing strategies that do not exceed processing capacity (Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998). This fact establishes an important boundary condition for cognitive complexity effects, and the method for analysis of relational complexity (MARC) is founded on these principles (Halford, Bunch, et al., 2007; Halford, Cowan, et al., 2007).

Some of the classical findings in cognitive development reflect capacity limits due to constraints on decomposition. Transitive inference has been the subject of considerable controversy (Halford & Andrews, 2006), but the difficulty that children have with the explicit transitive inference task is not adequately explained without reference to processing load, as shown earlier. Transitive inference imposes a load on WM because conceptual chunking and segmentation are constrained as both premises must be processed to assign an element to a slot. For example, the premise *Tom is taller_than Peter* shows that Tom should be assigned to the top or middle slot in the ordering schema, but the second premise *Bob is taller_than Tom* is required to determine which slot. The result is that premise integration, or assignment of elements to slots in a schema consistent with the premises, imposes a load on WM in children (Halford et al., 1986) and adults (Maybery et al., 1986). It also activates regions of the prefrontal cortex (Fangmeier, Knauff, Ruff, & Sloutsky, 2006).

Another example is class inclusion, which is resistant to decomposition because it is defined by relations between a superordinate class, a subclass and (at least) one complementary subclass, and is inherently ternary relational (Halford, 1993). A source of difficulty in the Dimensional Change Card Sort task (Zelazo et al., 2003) is that the presentation tends to constrain decomposition of the task (Halford, Bunch, et al., 2007).

### 1.7. Processing capacity limits

The boundary conditions for the relational complexity metric were used to estimate the limits to processing capacity by Halford, Baker, McCredden, and Bain (2005). The method was based on interpretation of interactions, because decomposition into component variables is constrained due to the influence of each variable on the effects of others (as with statistical analyses where main effects cannot be interpreted if there is a significant interaction). It was found that adult humans can process approximately four variables in parallel. Cognitive developmental studies (Andrews & Halford, 2002; Halford, 1993) have shown that the median ages at which relations of a given complexity are attained are: unary relations at 1 year, binary relations at 18 months to 2 years, ternary relations at 4–5 years and quaternary relations at 11 years. However, attainment of relational concepts is experience-based, and so there is no suggestion that all the concepts of a given level are attained concurrently. Our data do show that attainments within a domain are influenced by complexity. For example, we found that binary relational (i.e., simpler) versions of tasks including transitive inference, class inclusion, hierarchical classification, cardinality, and sentence comprehension were attained by 3–4 years whereas more complex, ternary relational versions were attained at a median age of 5 years (Andrews & Halford, 2002).

## 2. Computational models

In this section we examine how the properties outlined above have been captured in our computational models. Two other models are examined for comparison.

### 2.1. Formal models of structural commonalities

Our efforts to model cognitive development have been designed to capture the processing of relational knowledge, rather than to exploit any one type of architecture. Our approach is to match symbolic models to the empirically substantiated conceptual account of higher cognition that we have outlined. The criterion is that the models implement the established properties of relational knowledge.

We began with mathematical category theory models at a high level of abstraction (Halford & Wilson, 1980) but then turned to modelling specific processes, first with a self-modifying production system model of transitive inference acquisition (Halford et al., 1995), and then with symbolic connectionist models of relational knowledge (Halford et al., 1994, chap. 7; Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998).

The category theory account of structure-based mappings between cognitive representations initiated by Halford and Wilson (1980) has undergone considerable subsequent development recently. This work was motivated in part by the observation (Andrews & Halford, 2002; Halford, 1993) that there is a correspondence between tasks such as transitivity and class inclusion, despite major differences in task characteristics.

However, it is now possible to define a more penetrating mathematical basis for this structural commonality (Phillips et al., 2009). As mentioned earlier, category theory defines mathematical structures by transformations, rather than objects, which makes it very appropriate for modelling structure-based mappings between cognitive representations. Category theory therefore defines structures at a very abstract level, but one which has been found to provide important insights in computer science (MacLane, 1998) and now in psychology. This approach essentially means that cognitive processes are categorised by structural properties rather than elements. The core of our theory is the way structure is defined, which differs in many respects from that of Piaget (Halford, 1993).

Briefly, a (mathematical) category consists of a collection of objects ($A$, $B$, ...) and a collection of morphisms, or maps ($f$, $g$, ...) between objects (e.g., $f$: $A \rightarrow B$ is a map from $A$ to $B$) such that every object is associated with an identity morphism (e.g., $1_A$: $A \rightarrow A$ is a map from object $A$ to itself), and every composition of two morphisms (where defined) equals a third morphism (e.g., the composition of $f$: $A \rightarrow B$ and $g$: $B \rightarrow C$, denoted $g \circ f$, equals a morphism $h$: $A \rightarrow C$). Composition must also satisfy rules known as unity and associativity. It has been shown (Phillips et al., 2009) that the structure of transitive inference and class inclusion are formally connected by the dual structures known as product and coproduct in category theory (MacLane, 1998).

Transitivity and class inclusion are attained at a median age of 5 years (Andrews & Halford, 2002), and the derivation shows that the empirically observed correspondence has a formal mathematical basis. The simpler versions of the tasks that are successfully performed by younger children do not have this formal correspondence. Phillips et al. (2009) have demonstrated the same point for empirically observed correspondences involving matrix completion (Halford, 1993), cardinality (Andrews & Halford, 2002), the dimensional change card sorting task (Halford, Bunch, et al., 2007; Halford, Cowan, et al., 2007), weight–distance integration in the balance scale (Halford et al., 2002), and theory of mind (Andrews, Halford, Bunch, Bowden, & Jones, 2003). The mathematical derivations and the empirical observations collectively mean that we have found a way to categorise concepts by their deep structure that underlies the surface properties of tasks. Thus the correspondence between tasks as different as transitive inference and class inclusion is no accident and does not result from superficial properties.

### 2.2. Self-modifying production system model

Acquisition of explicit transitive inference has been modelled by a self-modifying production system model, the Transitive Inference Mapping Model (TRIMM; Halford et al., 1995) which incorporates mapping into an ordering schema in WM. TRIMM is based on production rules. Each rule comprises a condition–action pair, and each rule represents a single step in problem solving. A production "fires", that is, is activated, when the relevant condition is satisfied, and the action then ensues. If no production has a condition that is satisfied, the model goes into the mode of developing new productions. This process is guided by using an ordered set of three elements that serves as a template. The ordered set can be anything in the experience of the child (e.g., three siblings varying in height, the story of the three bears, etc.), the only requirement being recognition of the structural correspondence between template and the ordering produced by a set of productions. For convenience, we assume the template is the ordering schema top, middle, bottom mentioned earlier.

To illustrate operation of the model, given the premise *Tom is taller than Peter*, a production fires that creates the order *Tom*, *Peter*. When the premise *Bob is taller than Tom* is presented, *Bob* might be appended to the ordered pair, creating the order *Tom*, *Peter*, *Bob*. The error can be detected because *Tom above Peter below Bob* is not in correspondence with the template. Detection of this discrepancy requires sufficient WM capacity to represent the ordered set, that is, a ternary relation. Occurrence of an error also causes this production to receive a decrement in strength and leads to a search for a new production, which appends *Bob* to the front of the ordered pair, producing *Bob*, *Tom*, *Peter*. This production yields an ordering that is in structural correspondence with the template; it will generate a correct ordering and therefore receives an increment in strength. Where WM in insufficient, the error will not be detected and there will be a failure to develop productions that consistently produce ordered sets. The model therefore accounts for findings (Halford, 1984) that children under 5 years are likely to correctly order pairs of elements but fail to integrate the pairs into ordered sets of three or more elements. The TRIMM model demonstrates a mechanism by which representations of relations can yield reasoning strategies.

### 2.3. Symbolic connectionist models

Neural net models, usually known as symbolic connectionist models, implement the criterial properties of relational knowledge that we have defined (Doumas, Hummel, & Sandhofer, 2008; Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998). The models are designed to simulate structure-consistent mappings between two relational representations, and the mapping process is dynamic, attentional, and imposes a processing load in WM. These models implement dynamic bindings between relation symbols and the sets of ordered $n$-tuples that constitute the extension of a relational representation. The components retain their identity in the binding and are accessible to other cognitive processes.

### 2.4. Structured Tensor Analogical Reasoning (STAR) model

The STAR model is designed to represent relations of varying complexities from unary to quinary, the latter being the empirically determined limit to complexity of relations that can be incorporated in a single representation. Mathematically relations have a wide range of interpretations. For example, a proposition such as "likes(John, Pizza)" is a relational instance and is also a binary relation. The representation of this relational instance in STAR is shown in Fig. 2.

In STAR (Halford et al., 1994, chap. 7; Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998; Wilson, Halford, Gray, & Phillips, 2001), relational representations are formed by binding the relation symbol and arguments into an array, as shown in Fig. 2. For example, the relation "loves(John, Pizza)" is represented by activations on three sets of input units, representing *loves*, *John*, *Pizza*, as shown in Fig. 2A (but with the relation symbol omitted, see caption). The sets of input units are treated mathematically as vectors, and the binding is performed by computing the outer (or tensor) product of the component vectors, forming matrices as shown in Fig. 2. The activations of the binding units are formed dynamically in WM, as the direct result of activations in the input units. Thus the relational representation can be formed in a single trial, and it lasts as long as the activations persist. The symbol (e.g., likes) and roles (e.g., lover, loved-object/person) are assigned to specific ranks of the representation, so formation of the model entails dynamic mapping to a coordinate system, consistent with the WM model of Oberauer (2009). The principle of commutativity from category theory is applied to make consistent assignments (Halford & Wilson, 1980).

#### 2.4.1. Compositionality

The model implements compositionality (accessibility) as defined earlier, using the retrieval process. Retrieval is achieved by using one or more vectors as input, and the output is the remaining vectors. This is done using the dot product operator, which is roughly equivalent to a non-standardised correlation. In a representation with $n$ components, any $n$-1 components can be used as inputs yielding the remaining component. Thus given "loves(John, Pizza)," if the input is *loves*(*John*. . .) the output is *Pizza*, or given . . .. (*John*, *Pizza*) the output is *loves*, and so on. Thus, the relation symbol and its arguments all remain accessible. For a more comprehensive treatment see Halford et al. (1994), Halford, Bain, et al. (1998), Halford, Wilson, et al. (1998), or Wilson et al. (2001).

#### 2.4.2. Systematicity

Assignment of the relation symbol and roles to specific ranks in the representation, and their binding by computing the outer product of component vectors, means that structure is inherent in the representation and therefore implements systematicity. It also gives a natural correspondence to relational representations in predicate calculus expressions such as "loves(John, Sally)" and in natural language ("John loves Sally") where *lover* and *loved* occupy specific positions in the sentence.

#### 2.4.3. Superposition

The representations of relational instances can be superimposed. The relational instances "loves(John, Sally)" and "loves(Ken, Tina)" are represented by arrays with different sets of input units, giving a different pattern of activation in the binding units, as shown in Figs. 3A and B. The representations can be superimposed, forming a composite representation, as shown in Fig. 3C. Notice that this is distinguishable from the superimposed representations of "loves(John, Tina)" and "loves(Ken, Sally)" in Fig. 3D. The components, symbol and arguments retain their identity in the composite representation, and the role assignments are maintained.

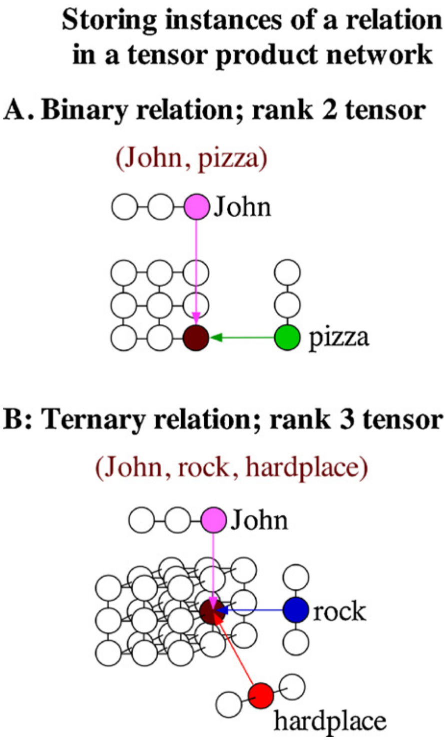


**Fig. 2.** The STAR model representation of binary and ternary relations. Part (A) represents the binary relation *John likes pizza* and part (B) represents the ternary relation *John is between a rock and a hard place*. The relation symbols (*likes* and *between*) are omitted to reduce the complexity of the figure, but each would constitute an extra rank, or extra dimension, so (A) would be Rank 3 and (B) would be Rank 4 (which cannot be shown on paper).

Superposition enables recognition of commonalities because common features have more influence on the output. It can also provide a representation of relational slots, because a slot corresponds to all the elements in a specific position in the representation. Given a structural alignment mechanism, relational representations can be formed initially without explicit representation of slots. The superposition of elements on the slots of the tensor representation represents variables. Given a representation of "larger_than{. (whale, fish), . . ., (ship, canoe), . . ." each slot can be filled in a (potentially infinite) variety of ways, effectively taking a step toward representation of a variable.

### 2.5. *Effects STAR accounts for*

We now consider how STAR implements the properties of relational knowledge outlined earlier.

#### 2.5.1. *Processing complexity*

Relational complexity corresponds to the number of roles in a representation. Thus *loves*(*John*, *Sally*) has two roles and is a binary relation, whereas arithmetic addition has three roles, *addend 1*, *addend 2*, *sum* and is a ternary relation. The number of roles affects the STAR model as follows: If each component vector has $k$ units, the number of binding units for an $n$-ary relation is $k^{n+1}$. The representation of a $n$-ary relation has a rank of $n+1$ because one rank is occupied by the relation symbol. Thus the representation of the binary relation *loves*(*John*, *Sally*) is rank 3 because it comprises three vectors representing *loves*, *John*, and *Sally*. Therefore the processing demands increase exponentially with the complexity of relations as defined by the relational complexity metric.

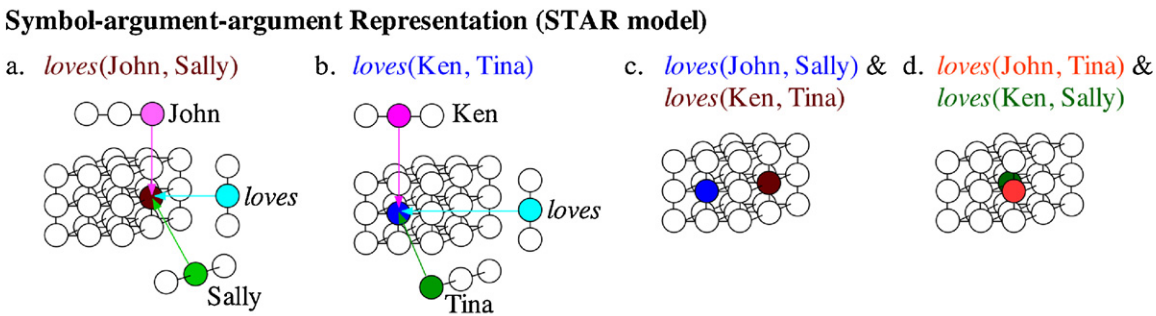


**Fig. 3.** (a), (b), Representation of single proposition. (c), (d), Superimposed representations of propositions.

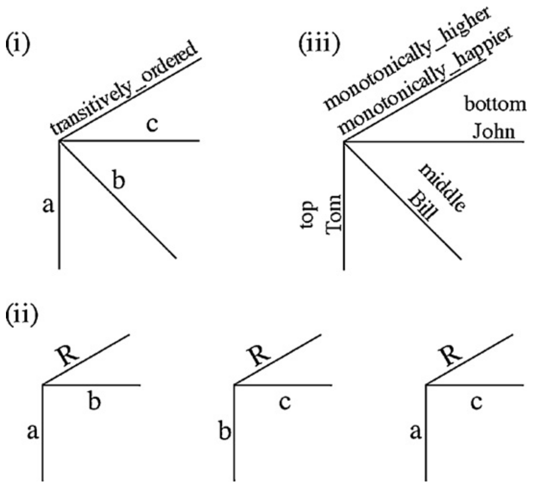


**Fig. 4.** Integration of transitive binary relations into a ternary relations.

Notice that the number of slots has more effect than the size of the components, which corresponds to the amount of information encoded in each vector. In his classic paper, Miller (1956) mentions the paradox that the limiting factor in processing capacity is number of entities, rather than their information value, which in our formulation corresponds in an approximate sense to size as more information tends to require more units for its representation. As the rank of a representation increases, the number of binding units can become unsustainably large, leading to a gradual increase in errors with increasing complexity. The capacity limit is reached when performance declines to chance level, as occurred with quinary relations in the study by Halford et al. (2005). Importantly, concepts of equal relational complexity have equal rank in the tensor representation, so STAR provides a natural account of the structural equivalence of concepts.

Capacity to process relations of increasing complexity with age is predicted to depend on representations of higher rank (with more interconnected vectors), such as a transition from binary relational to ternary relational representations portrayed in Fig. 4. This transition would depend on appropriate interconnection between the ranks; for example, each element in one vector must be connected to a set of binding units that is connected to all the other vectors. Details of interconnection are shown by Halford et al. (1994).

### *2.5.2. Integration of relations*

The way STAR models integration of premises in transitive inference is shown in Fig. 4. Premises, aRb and bRc correspond to binary relations and are represented as rank 3 tensor products, as shown in Fig. 4(ii). These can be integrated into the ternary relation transitively_ordered (a, b, c) as shown in Fig. 4(i). The conclusion aRc is another binary relation that can be retrieved from the integrated representation. The integrated representation in Fig. 4(i) can be superimposed on the ordering schema monotonically_higher(top, middle, bottom), as shown in Fig. 4(iii).

### *2.5.3. Analogy*

A simple analogy as processed by STAR is shown in Fig. 5, based on superimposed representations. The predicates *mother_of*, *protects*, *feeds*, *larger_than* are each linked to entities *woman: baby* and *mare: foal*. This yields representations of *mother_of*((*woman*, *baby*), (*mare-foal*)), *protects*((*woman*, *baby*), (*mare-foal*)), *feeds*((*woman*, *baby*), (*mare-foal*)), *larger_than*((*woman*, *baby*), (*mare-foal*)). Analogy is performed by retrieval from the representation, as shown in Fig. 6. If *woman* and *baby* are entered into the structure, the output is a bundle of all the relation symbols that

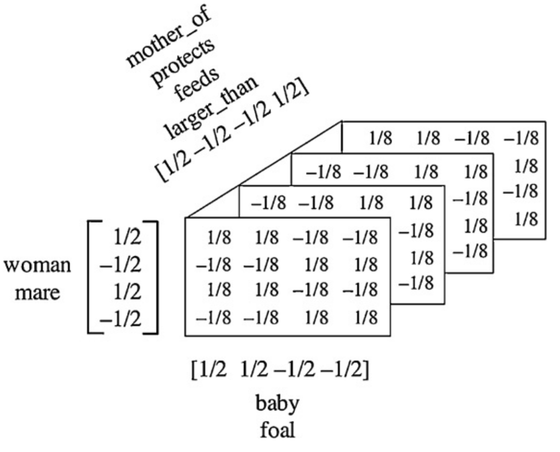


**Fig. 5.** The vectors, such as [0.5 0.5 −0.5 −0.5] representing concepts must be of length 1 and orthogonal to each other. The tensor is shown here storing a single relational instance, such as larger_than(woman, baby). In general, it would be storing several such, simultaneously, and its value would then be the sum of the tensors for the individual relational instances.

have these words as arguments – *mother_of*, *protects*, *feeds*, and *larger_than*. In the next step, the inputs are *mare* and the *bundle of relation symbols*, and the output is *foal*. Thus the simple analogy *woman:baby::mare:foal* is implemented as a set of operations on a relational knowledge base represented by the tensor product shown in Figs. 5 and 6. This example involves an analogy based on a binary relation, and it can be extended to more complex relations by using representations of higher rank. The model is capable of simulating a wide range of analogy problems (Halford et al., 1994, chap. 7, Section 5.2).

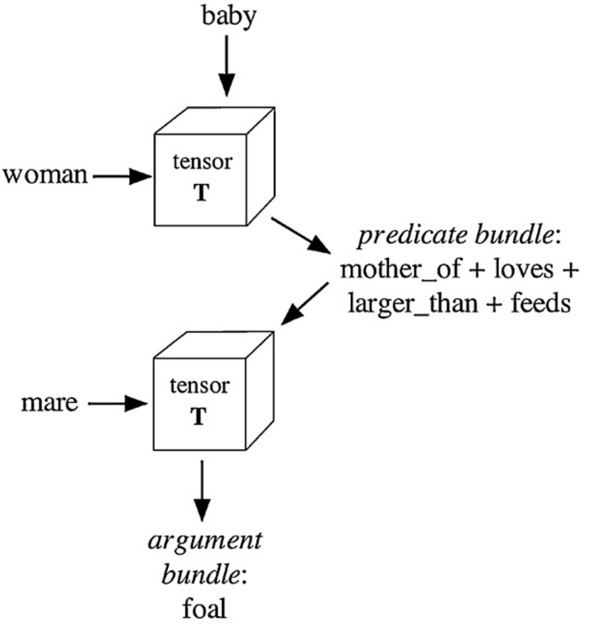


**Fig. 6.** Processing analogy woman:baby::mare:foal.

#### 2.5.4. Role-respecting similarity

This is represented in STAR by computing the dot product (roughly equivalent to a nonstandardised correlation) between vectors representing the relation. Thus *woman feeds squirrel* would be represented by the outer product of three vectors representing *feeds*, *woman*, *squirrel*. Similarly *man feeds dog* would be represented by outer product of the three vectors. The similarity of these representations is computed by the dot product of the set of three vectors (i.e., *feeds*, *woman*, *squirrel* with *feeds*, *man*, *dog*). The size of dot product would reflect the similarity of components in corresponding position, that is *woman to man*, *feeds to feeds*, and *squirrel to dog*. However *squirrel feeds woman* would produce a much lower dot product because of the dissimilarity of elements in corresponding positions, i.e., *woman to squirrel*. Thus the model captures the fact that *woman feeds squirrel* is not similar to *squirrel feeds woman*, despite identical elements.

#### 2.5.5. Truth value

The truth of a proposition can be determined by computing the dot product of vectors representing the proposition with vectors representing semantic memory (Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998, Section 4.2.2).

#### 2.5.6. Dynamic modification of representations

One property of higher cognition, captured by relational knowledge, is that representations can be modified dynamically, as in hypothesis testing. Suppose that young children tend to estimate area of a figure as a combination of length added to breadth ($A = L + B$). We can change this to $A = L \times B$ without incremental learning (i.e., we do not have to learn multiplication of dimensions) by dynamically modifying our representation. This is captured in STAR because we can superimpose $3 + 5 = 8$ on $3 \times 5 = 15$. When we change the relation symbol, we select a different set of relational instances: The switch from + to $\times$ changes ordered 3-tuple 3, 5, 8 to 3, 5, 15, etc. Thus STAR permits representations to be changed dynamically in the course of problem solving.

#### 2.5.7. Conceptual chunking

Conceptual chunking is simulated by “collapsing” two or more vectors into a single vector by concatenating, convolving or superimposing vectors (Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998, Section 4.2.4). The chunk can then be an entity in another relational representation, and hierarchical representations can be formed in this way. However, the relations between components of a chunk cannot be accessed. To access components it is necessary to recover the chunked vectors and compute a tensor product representation of their binding.

#### 2.5.8. Hierarchical structures

Fig. 7 shows how STAR represents hierarchical structures, such as the proposition *Mary believes*, (*that*) *Ken resents*, (*that*) *Jim gave Mary* (*a*) *ticket*. The number of components is too large to be processed in a single cognitive step (Halford et al., 2005), so it must be segmented into components that do not exceed capacity of WM. This can be done using the type of hierarchical representation shown in Fig. 7 (Gray, 2003), where the levels are processed one at a time. The highest level comprises the proposition *believes*(*Mary*, *chunk1*), where *chunk1* comprises a chunked representation of the proposition *resents*(*Ken*, *chunk2*), and chunk 2 is a chunked representation of *gave*(*Jim*, *Mary*, *ticket*). At the next level, chunk 1 is unpacked, yielding *resents*(*Ken*, *chunk2*). Further unpacking of *chunk2* yields *gave*(*Jim*, *Mary*, *ticket*).

The representations at each level include a representation of a chunk, which is linked to the next lower level. Thus the representation of the top level is a three-dimensional array formed by input vectors *believes*, *Mary*, *chunk1*. At the next lower level, the representation is a three-dimensional array formed by input vectors *resents*, *Ken*, *chunk2*, which is an expansion of *chunk1*. If *chunk1* is used as input, the representation of resents, *chunk2* is retrieved. At the next lower level, the representation is a four-dimensional array formed by input vectors, *gave*, *Jim*, *Mary*, *ticket*. If *chunk2* is used as input, the representation of *gave*, *Jim*, *Mary*, *ticket* is retrieved. Alternatively if, for instance, *resents*(*Ken*, *chunk2*)

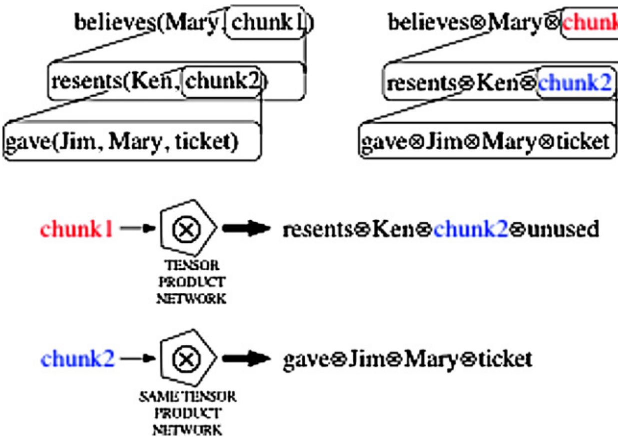


**Fig. 7.** STAR hierarchical representation of the compound proposition *Mary believes* (that) *Ken resents* (that) *Jim gave Mary ticket*. To represent these propositions within a tensor product network of bounded rank, the STAR2+ model chunks nested propositions, and expands them using a recursive matching algorithm.

is the input, *chunk1* is retrieved, and so on. Thus STAR can move naturally up and down a hierarchically structured representation using the standard retrieval processes that are inherent in the model.

### 2.6. The Relcon (Relational Concept) model

The Relcon (Relational Concept) model (Gray, Wilson, Halford, & McCredden, 2006) shows how relational categories that exhibit prototypicality and context sensitivity (Rosch & Mervis, 1975) can be formed with the STAR architecture. The algorithm works on a knowledge base represented as a set of propositions (relational instances). For example, furniture knowledge might consist of propositions including *made_of*(*chair*, *wood*), *stands_on*(*chair*, *floor*), *found_in*(*chair*, *living-room*), . . ., *made_of*(*desk*,*wood*), *stands_on*(*desk*,*floor*), . . ., *found_in*(*vase*, *living-room*), . . ., (*made_of*(*vase*, *pottery*), . . ., etc. If the first input is *chair*, the output is the set of attributes relating to chair, i.e., *made_of*( , *wood*), *stands_on*( , *floor*), *found_in*( , *living-room*). This corresponds to stored knowledge about chairs. If these attributes are now used as input, the output is all the elements bound to them, *chair* (retrieved 3 times), *desk* (retrieved twice) and *vase* (retrieved once). The algorithm iterates to a stable set of elements that share some (but not all) attributes and represents a category based on family resemblance. The typicality of a member of the category can be determined by computing the dot product of the vector representing the member with the vector representing the category, comprising all the instances superimposed. Typicality is represented in the model because chair is more typical of furniture that vase and it shares more attributes with other instances of the category and least attributes with non-instances of the category. The model therefore simulates the correlation between prototypicality and the degree to which a concept member has common properties with other concept members. The model also captures context sensitivity, that is typicality of concept instances varies with context (e.g., chicken is most typical of the bird category in the context of a farmyard, whereas robin might be more typical in the forest).

Structured knowledge can, in principle be acquired by cumulative storage of instances of a relation. This can be illustrated using the balance scale model. The balance state depends on moments on the two sides, where moments are products of weights and distances on that side. For simplicity, we consider the left side only, the same argument applying to the right side. The weights, distances, and moments on each side of the balance can be represented by vectors, e.g., on the left side, we have $V_{weightL}$, $V_{distanceL}$, $V_{momentL}$. Each instance of this relation can be represented by the outer products of three vectors, which can be superimposed to represent cumulative knowledge.

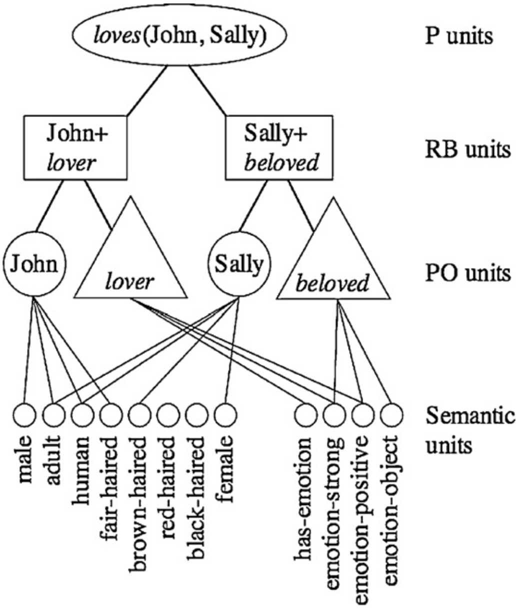


**Fig. 8.** Outline of the architecture of the DORA model.
Adapted from Fig. 2b in Doumas et al. (2008).

### 2.7. The Discovery Of Relations by Analogy (DORA) model

The DORA model (Doumas et al., 2008) can be used to illustrate how relational knowledge can be represented with a different type of architecture. Relations are represented in DORA by four layers of units, as shown in Fig. 8. At the top are P (Proposition) units that are linked to PO (Predicate and Object) units via RB (Role-Binding units). The PO units are connected to semantic units that represent features of objects (e.g. male, adult, . . ., has-emotion). RB units bind objects to roles in the proposition, so John is bound to the *lover* role in the proposition *loves* (*John*, *Sally*), and so on. This effectively binds the predicate (relation symbol) to the arguments, as required for relational representations. In DORA, the role (in this case *lover*) is explicitly represented, but this is not true in all models.

Role-filler bindings in DORA are dynamic, and are coded by firing in close temporal proximity. For example, to represent *loves*(*John*, *Sally*), the units representing the *lover* role fire in close temporal proximity with units representing *John*, while units representing *loved* fire in close temporal proximity with *Sally*. Units representing the *lover* role fire out of synchrony with units representing the *loved* role, and their fillers *John* and *Sally* also fire out of synchrony. Retrieval from long-term memory occurs due to activation that originates with the P unit, passes through RB and PO units to semantic units, which excite units in long term memory. Mapping occurs by concurrent activation of units in two analogs, aided by mapping hypotheses linking corresponding units (mapping P to P, RB to RB, etc.). These processes take account of semantic similarity effects.

The DORA model simulates many phenomena associated with acquisition of relational knowledge, most of which have developmental relevance. These include initially holistic representations of relations, learning from examples without explicit feedback, the role of structural alignment in learning relations, the relational shift or progression from featural to relational representations, progressive abstraction of representations, the influence of increasing semantic knowledge, and the role of comparison in acquisition of relational knowledge.

### 2.8. Comments on symbolic connectionist models

We regard Doumas et al. (2008) as justified in their claim that "DORA serves as an existence proof that relational representations can be learned from examples, thereby addressing one of the fundamental problems facing symbolic models of cognition." (p. 30). A similar point is demonstrated empirically by Halford and Busby (2007). DORA models analogy and relational schema induction; it provides a plausible, if limited, mechanism for cognitive development, and there is some supporting neurological evidence for synchronous activation. DORA accounts for capacity limitations based on the limits to the number of distinct phases of oscillation.

Role-filler bindings alone, however, are not sufficient to define propositions or relational instances. If we represent *loves*(*John*, *Sally*) and *loves*(*Ken*, *Tina*) solely by binding *John* and *Ken* to the *lover* role and *Sally* and *Tina* to the *beloved* role, the result would be identical to *loves*(*John*, *Tina*) and *loves*(*Ken*, *Sally*). The role assignments do not represent the *n*-tuples that comprise the relational instances. By contrast, these representations are distinguished in STAR, as shown in Fig. 3C and D. In DORA this problem is obviated because propositions are represented by four hierarchical levels of units as shown in Fig. 3, and the binding of the P unit at the top of the hierarchy to the RB, PO and semantic units serves to identify specific propositions. This problem is serious however for relations that entail many instances, as Halford et al. (1998, Sections 4.1.1.1 and 4.1.3) point out. For example, in the ternary relation *arithmetical addition*, every number is bound to every role (e.g., the number "3" is bound to the first addend role in 3 + 2 = 5, to the second addend role in 4 + 3 = 7, and to the sum role in 1 + 2 = 3). The effect is that role-filler bindings alone do not identify relational instances.

STAR also implements many properties of relational knowledge (Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998). It has accessibility because any component can be retrieved from the relational representation, it can determine the truth value of a proposition, it implements systematicity, dynamic modification, analogical reasoning, structural correspondence, and representation of variables. It can account for processing load increases as a function of cognitive complexity and for capacity limitations, and it implements higher-order relations. The properties are relevant to a wide range of phenomena. For example, accessibility and hierarchical representation are both relevant to Cognitive Complexity and Control theory (Zelazo et al., 2003) and relational knowledge plays a major role in development of reasoning (Gentner, 2010; Halford & Andrews, 2006).

Symbolic connectionist models effectively implement mapping between relational representations, which we propose to be the main driver of the transition to symbolic processes, and the consequent ability to generate inferences.

### 2.9. Semantic cognition model

One of the most advanced models of emergence of cognitive structure is the Semantic Cognition model of Rogers and McClelland (2004), which shows how representations in a feedforward net can reflect similarities within and between categories (e.g., Robin and Salmon are more similar than Robin and Daisy). Similarities can be represented even if not directly perceptible, because the activations in the Representation layer are influenced by the attributes required in the output layer. Representations come to reflect item similarities, e.g., the representations activated by "robin" and "canary" are similar because both are used to compute outputs such as "bird."

The model provides an existence proof for mechanisms that can categorise entities on the basis of input–output links. The activation patterns in representation and hidden units reflect similarities in the input–output functions computed. This feature can be applied to children's categorisations based on events in which entities participate. For example, animals and vehicles can be categorised by events in which they take part, somewhat independently of perceptible similarity (Mandler, 2000). This is potentially an important mechanism in the development of children's categories. However, by contrast with some symbolic connectionist models, such as DORA (Doumas et al., 2008), or STAR (Halford, Bain, et al., 1998; Halford, Wilson, et al., 1998), the Semantic Cognition model does not fully represent the properties of relational knowledge outlined earlier (Halford et al., 2010).

Multi-layered models form representations incrementally by incremental adjustment of connection weights or, as in the Semantic Cognition model, by adjustment of representations (backpropagatation to representation). They do not implement dynamic mapping in WM, which we argue to be essential to formation of relational knowledge. The Semantic Cognition model implements limited accessibility (Rogers & McClelland, 2004; Section R2), but it is not accessibility as we have defined it. Nevertheless, these models are important, even achieving major breakthroughs in many contexts, including plausible mechanisms for laying the foundations of relational knowledge. For example, the representation of *loves* could originate as a relational feature in a feedforward net, but it is the binding of the feature to the arguments (*John* and *Sally*) that enables it to function as a relation symbol. We have made more comparisons of the semantic cognition and symbolic models elsewhere (Halford et al., 2010). We believe there is scope for more hybrid models, perhaps along the lines of the CLARION model (Sun et al., 2005) but designed specifically to address the developmental questions raised by the articles in this issue.

## 3. Conclusion

Relational processes play a highly significant part in higher cognition, and to the extent that cognitive development depends on attaining higher levels of cognitive function, it must entail relational processes. Our conception of relational processes includes the major properties of higher cognition. A relational representation is a binding between a relation symbol and arguments such that the truth of the relation is maintained, and it must be possible to form maps between representations based on structural correspondence, together with other influences such as semantic similarity. Other properties include compositionality, meaning that elements retain their identity in compound representations and can be accessed, and systematicity, meaning that representations are intrinsically connected due to common structure. Relational representations are subject to processing load effects, loads depending on the number of entities that are related, with an upper limit of about four related variables for adult humans. The STAR and DORA models that have been described here represent these properties in different ways, and the Semantic Cognition model provides important acquisition mechanisms.

## References


Anderson, J. R. (1991). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Andrews, G., Birney, D. P., & Halford, G. S. (2006). Relational processing and working memory in the comprehension of complex relative clause sentences. *Memory & Cognition*, *34*(6), 1325–1340.

Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology*, *45*(2), 153–219.

Andrews, G., & Halford, G. S. (2011). Recent advances in relational complexity theory and its application to cognitive development. In P. Barrouillet, & V. Gaillard (Eds.), *Cognitive development and working memory: A dialogue between Neo-Piagetian and cognitive approaches* , ISBN 978-1-84872-036-7 (pp. 47–68). Psychology Press.

Andrews, G., Halford, G. S., Bunch, K. M., Bowden, D., & Jones, T. (2003). Theory of mind and relational complexity. *Child Development*, *74*(5), 1476–1499.

Blanchette, I., & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory & Cognition*, *28*, 108–124.

Boag, C., Neal, A., Loft, S., & Halford, G. S. (2006). An analysis of relational complexity in an air traffic control conflict detection task. *Ergonomics*, *49*(14), 1508–1526.

Boole, G. (1854). *An investigation of the laws of thought*. Walton & Maberly; Reprinted with corrections, Dover Publications, New York, NY, 1958 (reissued by Cambridge University Press, 2009; ISBN 9781108001533).

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning and cognitive neuroscience. *Advances in Child Development & Behavior*, *28*, 41–100.

Carey, S. (2009). *The origin of concepts*. New York, NY, US: Oxford University Press.

Christoff, K., & Owen, A. M. (2006). Improving reverse neuroimaging inference: cognitive domain versus cognitive complexity. *Trends in Cognitive Science*, *10*(8), 352–353.

Couvillon, P. A., & Bitterman, M. E. (1992). A conventional conditioning analysis of "transitive inference" in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*(3), 308–310.

Daneman, N., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 405–438.

Doumas, L. A., Hummel, J. E., & Sandhofer, C. A. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, *115*(1), 1–43.

English, L. D., & Halford, G. S. (1995). *Mathematics education: Models and processes*. Hillsdale, N.J.: Erlbaum. ISBN 0-8058-1457-4 0-8058-1458-2.

Evans, J. S. B. T. (2008). Dual-Processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278.

Fangmeier, T., Knauff, M., Ruff, C. C., & Sloutsky, V. (2006). fMRI evidence for a three-stage model of deductive reasoning. *Journal of Cognitive Neuroscience*, *18*, 320–334.
Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1–2), 3–71.
Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science: A Multidisciplinary Journal*, *34*(5), 752–775.
Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*, 306–355.
Goel, V. (2007). Anatomy of deductive reasoning. *Trends in Cognitive Sciences*, *11*(10), 435–441.
Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences*, *12*, 397–403.
Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychological Review*, *112*, 468–493.
Gray, B. (2003). Relational models of feature based concept formation, theory-based concept formation and analogical retrieval/mapping. Unpublished PhD Thesis, University of Queensland, Brisbane.
Gray, B., Wilson, W. H., Halford, G. S., & McCredden, J. (2006 July). Relcon: A tensor model of relational categorization. In *Paper presented at the 28th Annual Conference of the Cognitive Science Society, in cooperation with the 5th International Conference of Cognitive Science* Vancouver, BC.
Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? *Cognitive Psychology*, *16*, 65–93.
Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, N.J.: Erlbaum.
Halford, G. S., & Andrews, G. (2004). The development of deductive reasoning: How important is complexity? *Thinking and Reasoning*, *10*(2), 123–145.
Halford, G. S., & Andrews, G. (2006). Reasoning and problem solving. In D. Kuhn, & R. Siegler (Eds.), *Handbook of child psychology: Cognitive, language and perceptual development* (6th ed., Vol. 2, pp. 557–608). Hoboken, New Jersey: John Wiley & Sons.
Halford, G. S., Andrews, G., Dalton, C., Boag, C., & Zielinski, T. (2002). Young children's performance on the balance scale: The influence of relational complexity. *Journal of Experimental Child Psychology*, *81*, 417–445.
Halford, G. S., Bain, J. D., Maybery, M. T., & Andrews, G. (1998). Induction of relational schemas: Common processes in reasoning and complex learning. *Cognitive Psychology*, *35*, 201–245.
Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science*, *16*(1), 70–76.
Halford, G. S., Bunch, K., & McCredden, J. E. (2007). Problem decomposability as a factor in complexity of the dimensional change card sort task. *Cognitive Development*, *22*, 384–391.
Halford, G. S., & Busby, J. (2007). Acquisition of structured knowledge without instruction: The relational schema induction paradigm. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *33*(3), 586–603.
Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences*, *11*(6), 236–242.
Halford, G. S., Maybery, M. T., & Bain, J. D. (1986). Capacity limitations in children's reasoning: A dual task approach. *Child Development*, *57*, 616–627.
Halford, G. S., Smith, S. B., Dickson, J. C., Maybery, M. T., Kelly, M. E., Bain, J. D., et al. (1995). Modelling the development of reasoning strategies: The roles of analogy, knowledge, and capacity. In T. Simon, & G. S. Halford (Eds.), *Developing Cognitive Competence: New Approaches to Cognitive Modelling* (pp. 77–156). Hillsdale, NJ: Erlbaum.
Halford, G. S., & Wilson, W. H. (1980). A category theory approach to cognitive development. *Cognitive Psychology*, *12*(3), 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 Connnectionist and neural computation Theory, Vol. 2: Analogical connections* (pp. 363–415). Norwood, NJ: Ablex.
Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, *21*(6), 803–831.
Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences*, *14*(11), 497–505.
Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*, 295–355.
James, W. (1890). *The principles of psychology*, with introduction by George A. Miller, Harvard University, 1983, ISBN 0-674-70625-0.
Johnson-Laird, P. N. (2005). Mental models and thought. In K. Holyoak, & R. G. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 185–208). Cambridge University.
Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. New York and Cambridge: Cambridge University Press.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, Massachusetts: MIT/Bradford.
Kokinov, B., Holyoak, K., & Gentner, D. (Eds.). (2009). *New frontiers in analogy research: Proceedings of the second international conference on analogy*. New Bulgarian University Press.
Kroger, J. K., Sabb, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex*, *12*(5), 477–485.
MacLane, S. (1998). *Categories for the working mathematician* (2nd ed.). New York, NY: Springer.
Mandler, J. M. (2000). Perceptual and conceptual processes in infancy. *Journal of Cognition and Development*, *1*, 3–36.
Markovits, H., & Dumas, C. (1992). Can pigeons really make transitive inferences? *Journal of Experimental Psychology: Animal Behavior Processes*, *18*(3), 311–312.
Maybery, M. T., Bain, J. D., & Halford, G. S. (1986). Information processing demands of transition inference. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *22*, 2l8–223l.
McGonigle, B. O., & Chalmers, M. (1977). Are monkeys logical? *Nature*, *267*, 355–377.
Miller, N. (1956). The magical number seven, plus or minus two. Some limits to our capacity for processing information. *Psychological Review*, *101*(2), 343–352.
Naigles, L. R., Hoff, E., & Vear, D. (2009). Flexibility in early verb use: Evidence from a multiple-n diary study. *Monographs for the Society for Research in Child Development*, *74*(2).

Newell, A. (1980). Physical symbol systems. *Cognitive Science*, *4*, 135–183.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), *Psychology of learning and motivation* (pp. 45–100). Elsevier.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*(2), 109–130.

Phillips, S., Halford, G. S., & Wilson, W. H. (1995). The processing of associations versus the processing of relations and symbols: A systematic comparison. In J. D. Moore, & J. F. Lehman (Eds.), *Proceedings of the seventeenth annual conference of the cognitive science society* (pp. 688–691). Pittsburgh, PA: Lawrence Erlbaum.

Phillips, S., Wilson, W. H., & Halford, G. S. (2009). What do transitive inference and class inclusion have in common? Categorical (co)products and cognitive development. *PLoS Computational Biology*, *5*(12), e1000599. http://dx.doi.org/10.1371/journal.pcbi.1000599

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

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA, USA: MIT Press., p. 425.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605. Reprinted in DePaul, M. R., & Ramsey, W. (Eds.). (1988). Rethinking intuition: The psychology of intuition and its role in philosophical inquiry. Lanham, MD: Rowman & Littlefield Publishers.

Spearman, C. E. (1923). *The nature of intelligence and the principles of cognition*. London: MacMillan.

Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: A dual-process approach. *Psychological Review*, *112*, 159–192.

Tomasello, M., & Brandt, S. (2009). Commentary on Naigles, L.R., Hoff, E., & Vear, D. (2009). Flexibility in early verb use: Evidence from a multiple-N diary study. *Monographs for the Society for Research in Child Development*, *74*(2, No. 293), 113–126.

Van Gelder, T., & Niklasson, L. (1994). Classicalism and cognitive architecture. In A. Ram, & K. Eiselt (Eds.), *Proceedings of the sixteenth annual conference of the cognitive science society* (pp. 905–909). Atlanta, Georgia: Erlbaum.

Wertheimer, M. (1945). *Productive thinking*. New York: Harper.

Wilson, W. H., Halford, G. S., Gray, B., & Phillips, S. (2001). The STAR-2 model for mapping hierarchically structured analogs. In D. Gentner, K. Holyoak, & B. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 125–159). Cambridge, MA: MIT Press.

Wynne, C. D. L. (1995). Reinforcement accounts for transitive inference performance. *Animal Learning and Behavior*, *23*(2), 207–217.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function. *Monographs of the Society for Research in Child Development*, *68*(3), 1–155.

Zielinski, T. A., Goodwin, G. P., & Halford, G. S. (2010). Complexity of categorical syllogisms: An integration of two metrics. *European Journal of Cognitive Psychology*, *22*(3), 391–421. First published on 07 September 2009.