Research article

Robert Lowe*, Alexander Almér, Christian Balkenius

Bridging Connectionism and Relational Cognition through Bi-directional Affective-Associative Processing

https://doi.org/10.1515/opis-2019-0017
Received January 5, 2018; accepted May 14, 2019

Abstract: Connectionist architectures constitute a popular method for modelling animal associative learning processes in order to glean insights into the formation of cognitive capacities. Such approaches (based on purely feedforward activity) are considered limited in their ability to capture relational cognitive capacities. Pavlovian learning value-based models, being not based purely on fully connected feedforward structure, have demonstrated learning capabilities that often mimic those of 'higher' relational cognition. Capturing data using such models often reveals how associative mechanisms can exploit structure in the experimental setting, so that 'explicit' relational cognitive capacities are not, in fact, required. On the other hand, models of relational cognition, implemented as neural networks, permit formation and retrieval of relational representations of varying levels of complexity. The flexible processing capacities of such models are, however, are subject to constraints as to how offline relational versus online (real-time, real-world) processing may be mediated. In the current article, we review the potential for building a connectionist-relational cognitive architecture with reference to the representational rank view of cognitive capacity put forward by Halford et al. Through interfacing system 1-like (connectionist/associative learning) and system 2-like (relational-cognition) computations through a bidirectional affective processing approach, continuity between Halford et al's cognitive systems may be operationalized according to real world/online constraints. By addressing i) and ii) in this manner, this paper puts forward a testable unifying framework for system 1-like and system 2-like cognition.

Keywords: System 1-System 2, relational cognition, associative learning, representational rank, affective computation, habits.

1 Background

The relationship between animal learning-based connectionist models and models of so-called 'higher' cognitive capacity has been the subject of much research in the animal (and human) learning community over the last century. A variety of animal and human learning studies and models thereof have described the link between basic associative processes (e.g. relevant to this special issue, including habits and their formation) and cognitive capacities (e.g. Seger 2008, 2009; Phillips et al. 2009; Halford et al. 2014), including logico-relational based reasoning. Bridging associative processes modelled using the connectionist, i.e. neural network, approach, and 'higher' cognitive capacities (e.g. relational-based) does not necessarily entail discontinuity (Halford et al. 2014). Moreover, associatively learned processes such as habits, in

***Corresponding author, Robert Lowe,** Department of Applied IT, Division of Cognition and Communication, University of Gothenburg, E-mail: robert.lowe@gu.se
Alexander Almér, Department of Applied IT, Division of Cognition and Communication, University of Gothenburg
Christian Balkenius, Lund University

themselves, can be understood in terms of 'higher' cognitive phenomena, e.g. habits of thought (Seger & Spiering 2011). The aim of the present article, however, is to postulate the bridging of connectionist approaches and cognition according to *bidirectional affective-associative processing*. This entails the use of affective neural representations whose constituents combine encodings of value dimension states and other somatic sensory states. Their utilization as (affective) predicates in propositional relational knowledge we speculate may top-down focus attention on the featural and semantic constituents of objects/arguments that bind to those predicates. The modelling approach is constrained by a requirement to be faithful to a suite of empirical data (in humans and non-human animals) and its development should therefore be similarly amenable to empirical falsification. We discuss the extent to which such a model can be said to capture cognitive capacities of a type that concern relational knowledge acquisition and usage. It is not claimed that such affective states can facilitate the capturing of all relational knowledge but rather that affective states provide a subset of important relational (predicate) information in relating individuals to one another *John loves Mary* or in relating a subject to a particular task *John likes fishing*.

1.1 The Adaptive and Cognitive Value of Associative Processes

Associative processing provides a link between animals and humans in terms of learning and behaviour, particularly in the context of pavlovian and instrumental conditioning (Pearce 2013). Grounding cognitive architectures that utilize localist representations in associative (distributed connectionist) processes has received much focus in recent decades (Hummel & Holyoak 1997, 2001; Halford et al. 1998, 2010, 2014; Rogers & McClelland 2004; der Velde & de Kamps 2006, 2015; Leech et al. 2008; Eliasmith 2013; Sun 2015). The associative component, however, shouldn't be considered a mere add-on. Sun (2015) has suggested that computational cognitive architectures benefit greatly from incorporating a notion of 'implicit' processing, as part of a structured (dual-process) cognitive architecture. Agents utilizing cognitive architectures in the real world by necessity utilize systems that learn from (distributed) patterns of sensorimotor embodiment (Pfeifer & Scheier 2001; Montebelli et al. 2008, 2010, 2013; Lowe et al. 2008) whose interaction with value systems (e.g. pavlovian) may permit emergent activity attributable to 'higher' cognition (Braitenberg 1986, Kiryazov et al. 2013; Lowe & Kiryazov 2014; Barrett et al. 2016).

Halford et al. (2014) have likened human cognitive systems to Kahneman's (2011) notion of System 1 and System 2 processes where, in this reading, System 1 utilizes implicit processes based on associative learning and System 2 utilizes explicit relational cognitive processing. However, Halford et al. (2014) also acknowledge that the two systems need not work independently, as Halford et al. (2014, p.10) discuss in reference to Oberauer (2009): "associative and analytic systems are end points of a continuum". The first system may consist only of 'functional' structure, e.g. hidden layer representations in multi-layer perceptrons where the representations themselves are functional only within the context of the 'bottom up' activation of a given connecting input layer (their constituents or features; Fodor & Pylyshyn 1988). The second system, however, can utilize object and relation representations in a manner such that the activations of neurons or clusters thereof can occur independently of constituents and be utilized flexibly in relation with many other such objects (or relations), in this sense becoming symbolic.

Notwithstanding the dubious validity of there existing a sharp division between *implicit-based* (associative) and *explicit-based* (relational) cognitive systems (see Sun 2015 for discussion), the question of how connectionist (neural network implementations of associative learning processes) and relational systems can be bridged is an ongoing research topic (e.g. Leech et al. 2008; Sun 2011, 2015; Kolias & McClelland 2013; Phillips et al. 2017; Doumas et al. 2018).

1.2 Computational Modelling Approaches to Associative Learning and Cognition

The use of computational modelling to capture forced choice[1]-based performance of animal (and human) learning paradigms provides a critical tool into understanding the mechanisms underlying the behavior and the cognitive systems by which such learning and behavior is achieved.

In animals, associative (reinforcement) learning based models (Rescorla & Wagner 1972; Sutton & Barto 1998, 2018) have been used to capture a wide range of behavioural phenomena such as performance-based asymptotic learning curves (Seger & Spiering 2011), increased reacquisition learning following extinction[2] (Balkenius & Morén 2001) and resistance to unlearning following a previous training schedule of partial reinforcement[3] (Lowe et al. 2017). Neural network / connectionist models of animal (and human) learning paradigms abound and have been used to model memory and learning in relation to structured (i.e. not fully-connected feedforward) networks that entail multiple parallel processes. A multiple parallel processing model may consist simply of a dual-process. Such dual-process models have been used in animal learning with respect to: instrumental and pavlovian routes of learning (Mowrer 1947, Klopf et al. 1993); 'fast' versus 'slow' routes of processing (Armony et al. 1997, 2005; Lowe et al. 2009); multi-functional or dimensional-based value representations (Maki & Abunawass 1991, Balkenius & Morén 2001, Morén 2002, Doya 2008, Schmajuk 2010, Lowe & Ziemke 2013, Navarro-Guerrero et al. 2017a). Such structure may, nevertheless, still be utilized in associative, habit-like, i.e. automatic, processing yet give rise to behaviours that *masquerade* as relational, i.e. appear as though they are semantically constituted. In humans, associative learning based models have been used to explain data from hitherto explored animal learning based paradigms (Delameter et al. 2012, 2017; Lowe et al. 2016, Lowe & Billing 2017). However, humans may variably use associative, but *alternatively*, other cognitive strategies in order to complete such tasks. Frank et al. (2005), for example, showed that their associative based learning model could capture data from a paradigmatic learning task designed to require a transitive inference solution. Their model, nevertheless, predicted a profile of choice performance consistent with exposure to reinforced outcomes, i.e. a non-transitive inference memory-based solution. Notwithstanding, profiles of performance differed in subjects who reported understanding of the rules of the task suggesting that a more relational strategy was being employed in these cases (one which apparently improved performance). Experimental procedures able to tease out how and when system 1 and system 2 like knowledge is used provide a key means for furthering understanding of how associative learning brings to bear on relational cognition. For example, in some tasks research has indicated that associative learning (System 1-like) precedes relational (System 2-like) understanding (e.g. apprehension of task rules, e.g. Bechara et al.'s 2005 findings on implicit/"hunch" and explicit/"conceptual" apprehension of task rules on the Iowa Gambling Task). The two types of systems, however, may facilitate one another manifesting in an *apparent* temporal ordering of the utilization of the two systems *belying* the bidirectional structural coupling of the underlying process.

1.3 Computational Modelling Approaches to Higher Cognition – Associative Learning and Relational Cognition

More intricately structured (dual+-process) neural networks may allow for more cognitive functionality. Some such networks can permit forms of relational cognitive processing, e.g. analogical reasoning, according to either sensory similarity or semantic similarity. In the tradition of the parallel distributed processing (PDP) perspective (Rumelhart & McClelland 1986), McClelland and colleagues (Rumelhart 1990; Rogers & McClelland 2004; Kolias & McClelland 2013; Saxe et al. 2018) have produced a number of semantic relational neural networks based on feedforward structure trained with backpropagation. These

1 The experimental subject is required to choose among different presented options during a task.
2 Reacquisition entails the ability to relearn what was previously learned (choice response) following a period of non-reinforcement for 'correct choice' (extinction).
3 Experimental subjects are reinforced for correct choice probabilistically or at a low rate.

networks allow for objects and relations to be associated together through a hidden layer that provides integrated representations of object and relational semantic inputs. The network permits learning by semantic similarity – the semantic-cognition model exhibiting a "hierarchical progressive differentiation of structure" (Saxe et al. 2018). This manifests in initial learning of broader semantic categories followed by more refined categories, e.g. first animal and plant are differentiated, then bird and fish, and types of plants, then other individual attributes/features.

Whilst this notion of learning and development through semantic similarity has been acknowledged as an important contribution (Halford et al. 2010), it also suffers limitations regarding: i) the types of relational knowledge that can be acquired and retrieved (Halford et al. 2014), ii) the learning (and development) of the semantic units (items and relations) whose interactions permit the learning of attributes.

Earlier work by Halford and colleagues exemplifies this: Wilson et al. (2001) constructed a feedforward neural network, which learned through backpropagation, to test whether a connectionist approach could allow for the development of propositions with "the flexibility characteristic of certain classes of symbolic neural net models" (p. 1). A key component of such flexibility, emphasized by Wilson et al. (2001), is 'omnidirectionality', i.e. the ability of the network to access any relational component or (vectorized) components from any other relational component(s). In this case an autoencoder was used with a number of semantic (labelled) units feeding forward to hidden (representational) layers. The network was structured in such a way that relations, subjects and objects could be learned at the output layer. Use of an autoencoder had the advantage of permitting *accessibility* (Wilson et al. 2001) – so that when the query to the network is made "what does Jane like?" based on the activation of the subject (Jane) and relation (like), the object pizza, for example, can be accessed. The model, however, was shown to be limited in other respects. It cannot reliably answer the query "who likes pizza?" – and therefore lacks full omnidirectionality. Moreover, it can make category mistakes in relation to sensory similarity – stimuli whose neural encodings overlap lead to overgeneralizations, e.g. Jane could erroneously be found to like watermelons on the basis of the similarity of its encoding (of semantic constituents) with pizza.

Hummel & Holyoak (1997, 2001) produced an architecture (LISA – *Learning and Inference with Schema and Analogies*, later adapted by Doumas et al. 2008 to DORA – *Discovery of Relations by Analogy*) that in some sense deals with the above-mentioned problem of encoding similarities within the network. Neural representations of propositional (e.g. *Jane likes Pizza*) statements are decomposed into *sub-propositions*[4] ("*Jane+Likes*", "*isLiked+Pizza*") and their constituents (e.g. *Jane*, *Likes*, *isLiked*, *Pizza*), and also the semantic unit constituents thereof (e.g. *Female*, *Human*, *Bread*). A given "sub-proposition" and its constituents are synchronously activated. This sub-proposition is then bound to another sub-proposition constitutive of the propositional statement through oscillatory inhibition – one sub-proposition and its constituents are activated whilst inhibiting the activation of the other sub-proposition and its constituents and vice-versa. This serves as a form of working memory whose oscillatory frequency is of a rate higher than that of super-ordinate propositional levels and thereby provides a mechanism for propositional attentional binding. Importantly, this also resolves issues of erroneous retrieval based on sensory/semantic similarity: the constituents of sub-propositions can overlap in their encodings (contain many of the same properties) yet the temporal asynchronicity of their activations precludes *overgeneralization* of retrieval unlike for the Wilson et al. (2001) autoencoder model.

Halford et al. (1999, 2014) proposed that cognitive capacities that cover System 1-like and System 2-like processes could be ranked according to their relational representational complexity. Seven ranks were posited from ranks 0 to 6 inclusive. Ranks 0 and 1 concerned feedforward neural network architectures with either no representations (no hidden layer) or a "functionally structured" process (single hidden layer). These ranks were considered System 1-like. Ranks 2-6 on the other hand entailed relational representations implemented as symbolic neural networks ordered from unary, e.g. *Fido is a dog*, to quinary, e.g. as concerns requisite representations to solve the Tower of Hanoi problem. These ranks were considered System

4 This is the terminology of Hummel & Holyoak (1997) common to descriptions of the LISA architecture to label the neural representational layer below that corresponds to full propositional statements. In the DORA architecture (Doumas et al. 2008) this neural representation is referred to instead by its *Role* and *Binding constituents*.

2-like. The neural networks, forming the STAR (Halford et al. 1998, 2010) – Structured Tensor Analogical Reasoning – architecture, are, as the name suggests based on n-dimensional (n = representational rank -1) tensor representations. Access to elements of the tensors (queries) can be achieved through dot product calculations (similarly proposed by Smolensky 1990, and more recently Eliasmith 2013). Halford et al. (2014) suggested that whilst Rank 0 and 1 architectures can potentially produce relational and inferential outputs, such knowledge is considered implicit as it lacks the key property of omnidirectionality.

However, in its failure to elucidate learning and developmental mechanisms the STAR model has been described as 'descriptive' rather than explanatory (Heath & Hayes 1998). The bridge that connects System 1-like networks (based on associative distributive learning) and System 2-like networks (based on relational localist cognition) has not been fully clarified. Furthermore, STAR and the other above-mentioned models fail to address how the semantic units (object argument and predicate-based) emerge in learning and development, limiting them to disembodied and offline processing. Hierarchical ('deep') neural network structures provide a means to ground at least object argument units through imbuing a property of invariant representation at higher stages of the hierarchy (Hinton & Salakhudinov 2006, Eliasmith 2013, Rolls 2016). Eliasmith (2013) has referred to such deep networks as providing semantic pointers by which, given an autoencoder structure, units that provide invariant representations (Rolls 2016) may access information from (point to) their lower level featural constituents. Relatedly, Deep generative based neural network architectures (Rao & Ballard 1999; Hinton & Salakhudinov 2006; Goodfellow et al. 2016) provide promise for grounding, through a connectionist approach, some of the relational-symbolic cognitive capacities alluded to by Halford et al. (2014), in real-world online interaction.

2 Affective-Associative processing and the grounding of relational cognition

Notwithstanding the potential for deep neural network architectures to ground object (argument) units and their semantic properties as invariant representations, how exactly *predicate* representation invariance is learned is less clear (Doumas et al. 2008). In the present section, we will describe connectionist models with varying degrees of representational structure in relation to Halford et al.'s (1998, 2014) representational rank perspective. The models described are pavlovian-based and, we argue, provide a means for imbuing affective representations within a neural network structure. At the root of these models is the Rescorla-Wagner (1972) pavlovian learning algorithm considered by Halford et al. (2014) to be of representational rank 0, i.e. devoid of representational information and incapable of *explicit* relational cognition (since it lacks the key feature of omnidirectionality).

2.1 Rank 0 : The Rescorla-Wagner model

Among the most important of associative learning animal models is that of Rescorla and Wagner (1972) equivalent to the delta learning rule popular in machine learning research and a forerunner to, and special case[5] of, the temporal difference learning rule (Sutton and Barto 1998). This model when implemented as a neural network (figure 1) is considered by Halford et al. (2014) to consist of representational rank 0, i.e. there are no representations of the internal state given by a hidden layer. According to Halford et al. (2014), the Rescorla-Wagner (1972) model exemplifies a representational rank 0 model. It is considered a non-structured process since it has "no internal representation". The model is likened to a single-layered perceptron (see also Luzardo 2018) whereby external stimuli inputs are linearly combined in order to activate the output node.

5 When the temporal discount parameter gamma is set to 0 the Rescorla-Wagner and temporal difference models equate (see Niv 2009).

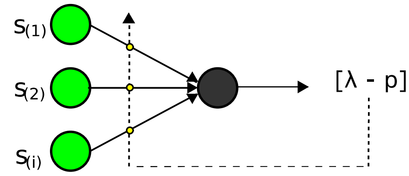

Figure 1. The Rescorla-Wagner model of learning depicted as a feedforward neural network. The black node (circle) represents the learnable value node which predicts reinforcement based on external stimuli inputs (green nodes). The black arrows represent feedforward outputs of neurons of the network. The connections between green and black nodes are learnable as denoted by the small yellow circles and dashed arrow. The connections are updated using the learning rule $\lambda - p$, which updates the weights as a function of the activation of their corresponding S nodes. Key: S(1 .. i) = external stimuli, λ = an external reinforcement signal, p = the output, or prediction, of the network. For the mathematical description of this model see Appendix A.

The Rescorla-Wagner model conflates into a single dimension of its value representation the information about multiple reinforcement properties of the stimulus. In animal learning, the use of a scalar value function has been noted as a key limitation of the Rescorla–Wagner model (Miller et al. 1995). As an example of its limitation, a reinforcer magnitude of 1.0 and presentation probability 0.5 is valued equivalently to one of magnitude 0.5 and presentation probability 1.0. Organisms may, in fact, benefit from multi-dimensional reinforcer information and value representations thereof. For example, high magnitude, low probability reinforcers might motivate learning the causal antecedents of the low presentation probability so as to increase future reward yield (Mackintosh 1971) and actively reduce prediction error (Pezzulo et al. 2015).

In spite of its simplicity, the Rescorla-Wagner model has been used to explain performance on a standard paradigm of transitive inferential learning that, according to Halford et al. (2014), requires representational rank 4 neural structure encoding for ternary relations. Transitive inference takes the form of experiencing a series of (at least) binary relations: if *aRb* and *bRC*, then *aRc*, where *R* represents an arbitrary relation (predicate), e.g. *greater than*, and where *a*, *b* and *c* are arguments, e.g. numbers 7, 5, 3. In the standard ('minimal'[6]) test – the five-term series – experimental subjects are required to learn the relations between pairs of stimuli based on their reinforcement value. The sequence to learn is: A+, B-; B+, C-; C+, D-; D+, E-. Where "+" indicates that choosing that stimulus, given the particular pairing, is reinforcing whereas the "-" suffix indicates no reinforcement is given for the stimulus. The letters themselves can be substituted for any type of stimuli, e.g. colours, odors. Transitive inference is said to occur when subjects are able to infer that given a novel pairing of B with D that B is preferable to D. That is to say, the inference is if B is rewarding relative to C and C is rewarding relative to D, then B should be rewarding relative to D. The Rescorla-Wagner model has been noted to be able to make correct transitive choices (Wynne 1995, Halford et al. 2014). However, the manner in which this type of 'connectionist' model achieves the performance owes to an artefact of memory. Stimulus A has the highest reinforcement value since it is never devalued, during learning, by non-reinforcement (A is never followed by non-reward while being paired with a rewarding stimulus). This unconditional positive reinforcement allows A to be chosen repeatedly at an early stage of learning so that the reinforcement value of B is relatively rarely devalued by A-B pairings (strong reinforcement of A leads to only rare choice of B leading to devaluation of B). B though has a lower reinforcement valuation than A (since it is occasionally devalued as a result of B not A choices). This means that C, competing against a not so strongly reinforced B, will be chosen relatively more often in B-C pairings than B will be in A-B pairings. In this way, along the sequence the associative model learns that A, then B, then C, then D, then E has the most reinforcement value purely as a function of associative memory based on rate of exposure to devaluation. Such transitivity has been deemed, therefore, *implicit* by Halford et al. (2014) since it does not permit omnidirectional relational-cognitive capacity and is limited with respect to representational rank 4 neural structures. Halford et al. (2014) also point out that by adapting the standard

6 It is considered minimal because anchor effects may be induced by A and E (A never leads to no reinforcement, E never leads to reinforcement) and so B, D provide the only pair for which to test for transitive inference.

five-term series set up to allow for E-A pairings, the transitive effect is lost in the Rescorla-Wagner model.

2.2 Rank 1+ : The Balkenius & Morén (2001)/ Morén (2002) model

Addressing the criticism above – of the use of a unidimensional representation of value in the Rescorla-Wagner model – Balkenius and Morén (2001 and Morén (2002) – see also Balkenius et al. (2009) – presented a model of learning (figure 2) that derived a computation of reinforcement omission from a reinforcement magnitude computation adaptation of the original Rescorla-Wagner rule. Although not explicitly noted by the authors, this effectively provides an omission probability when taken as a fraction of the reinforcement magnitude (standardly set to a maximum of 1.0). For every trial a reward is not presented to the network, the representation of omission increases and serves to inhibit effects of the Rescorla-Wagner based reward node (black node in the figure) on the output of the network (blue node in the figure). Since the reward node can only be updated by positive, but not negative, prediction errors[7] it provides a value representation of reward magnitude (or perhaps salience/presence) of the stimulus input.

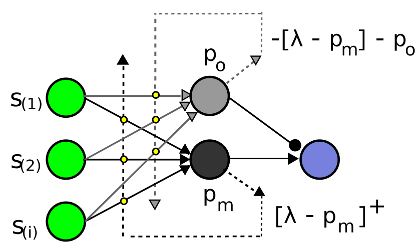

Figure 2. The Balkenius & Morén (2001)/ Morén (2002) multi-dimensional pavlovian model of value. The model embeds a Rescorla-Wagner model into a feedforward neural network but separates value into presence (or magnitude) – black node – and inhibition (or omission) – grey node – components. The output of the latter has an (contextually) inhibitory effect on the output of the former. This provides a hidden layer whose output yields a response rate (blue node). Key: p_m = magnitude node prediction, p_o = omission node prediction, λ = reinforcement value. For the mathematical description of this model see Appendix B.

The model provides a means thereby for contextual inhibition – by implementing differential rates of learning for the reward magnitude and reward omission representations (the latter being faster to learn/unlearn), it is possible for the network to inhibit output in a non-rewarding context but to rapidly re-produce the output/response when the rewarding context is re-established (as a result of rapid unlearning of the omission representation). The model thereby captures the profile of learning/unlearning characteristic of the acquisition-extinction-reacquisition paradigm while the standard Rescorla-Wagner model does not. For stimuli of magnitude 1, the model also preserves the property of the Rescorla-Wagner model that output conforms to a scalar representation of conflated reward magnitude and omission. Thereby the model should, under certain parameterizations (see Balkenius & Morén 2001), be able to replicate the Rescorla-Wagner model findings regarding implicit transitive inference. In the Balkenius and Morén (2001) version this entails setting the learning and unlearning rates of reward acquisition to the same value while setting the omission learning rate to zero. In the Morén (2002) model this entails setting the omission learning/unlearning rate to that of the reward learning.

As a fully connected feedforward neural network, the model might be viewed as a multi-layered perceptron with a single hidden layer (Halford's representational rank 1). Halford et al. (2007) suggested that rank 1 can transition to rank 2 "by imagining the hidden layer at Rank 1 ... being divided into two components which are then connected so as to form a matrix" (p.2). The splitting of the value representation 'hidden' layer and employment of such multi (dual)-process structure provides a key thereby for Halford

7 This is the case for the Morén (2002) version of the model.

to bridge associative and relational forms of neural networks. However, the Balkenius & Morén model lacks the "omni-directional access property of relational knowledge, which is considered basic to higher cognition" (Halford et al. 2007, p.2).

2.3 Rank 1+: Affective-Associative modelling

Many animal learning theories have posited the existence of (at least) dual routes for memory, learning and decision making. In animal learning models this has often manifested in systems that utilize pavlovian and instrumental conditioning (Mowrer 1947, Cardinal et al. 2002, de Wit et al. 2009). In pavlovian (or classical) conditioning, associations made between stimuli and outcomes are not contingent upon behavioural (instrumental) intervention. By contrast, in instrumental (or operant) conditioning such stimulus–outcome associations are contingent upon behaviour. Two-(or dual-) process theories have emphasized the interdependency of these two purportedly distinct processes (e.g. Overmier & Lawry 1979).

Two-process theories tend to emphasize one or other of the i) energizing or motivational component, essentially pavlovian, or ii) the directional control of responding (cf. Mowrer, 1947; Amsel 1958, 1992; Braver et al. 2014), i.e. where specific responses are selected. Early models of such two-processes emphasized the former component (within which the Rescorla-Wagner 1972, and Balkenius & Morén 2001, models would fit), while a perspective on two-process models entailing directional control as well as illuminating an important associative component, have since received growing focus. This initially took the form of *Associative Two-Process Theory* (Trapold 1970, Trapold & Overmier 1972) and then *Associative Mediational Theory* (Overmier & Lawry 1979, Kruse & Overmier 1982) where the latter identified a mediating role of differential reward expectancies on behavioural responding that could be embedded within the former (Lowe, Almér et al. 2017).

Underlying *directional* two-process theories is the use of a three-term contingency of instrumental learning: S–R–O, where S = stimulus, R = response, and O = goal based outcome, and where the pavlovian process (through S–O associations) is embedded within the instrumental process.

2.3.1 Associative Two-Process Theory

The theory of the Associative Two-Process (ATP) identifies S–R and S–E–R routes ('processes') where E represents an expectation of an outcome (or mediator) – see figure 3. ATP theory indicates that the outcome expectancy route is formed according to two associatively learned components. Firstly, there are S–E associations – pavlovian associations – and secondly there are E–R associations (Overmier & Lawry 1979) whereby outcome expectations can substitute for, compete with, or facilitate, the external stimulus in guiding instrumental responding. The division of this route into two components has been verified by use of *transfer-of-control* paradigms wherein the original, learned S-E and E-R contingencies are experimentally manipulated leading to testable hypotheses concerning the pattern of initial responding to these new contingencies (see Peterson & Trapold 1980; Lowe & Billing 2017).

By way of example, figure 4 schematizes a *transfer-of-control* scenario. As is typical for the paradigm, there are three phases. Each of these phases consists of a number of independent trials for learning: presentations of a stimulus, response options, and then a non-negative 'outcome' if the correct response is chosen.

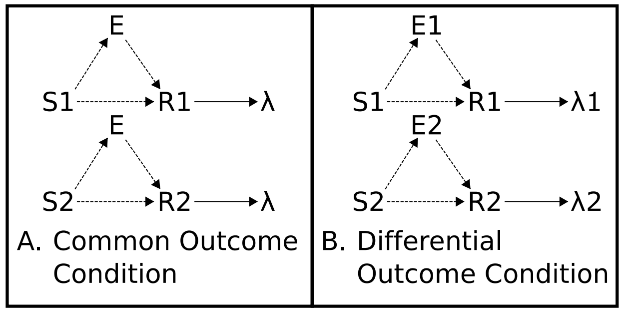

Figure 3. Associative Two-Process Theory. Response choice is guided by the interaction of two processes: i) a stimulus-response (S-R), or habit-enabling process; ii) a stimulus-outcome expectation-response (S-E-R) process. (A) Common Outcome Condition. Reinforced S-R associations (mappings) cannot be distinguished by outcome (λ). (B) Differential Outcome Condition. Reinforced S-R associations can be distinguished (λ1, λ2), and cued, by differential outcome expectancies (E1, E2). Directional arrows indicate causal links. Dashed lines indicate learnable connections.

The phases break down as follows: Firstly, there is an initial instrumental learning phase where the two components (S-E and E-R) of the 'goal-directed' route can be learned as well as the S-R ('habit-enabling') route. Secondly, a pavlovian (contingency change) learning phase is presented where new S-E associations are made. Finally, a second instrumental phase is utilized, which uses previously experienced stimuli and responses but introduces novel stimulus-response pairings.[8] This serves as a test of transfer of the knowledge of the components (S-E and E-R) learned in the first two phases that provide the relevant building blocks for the S-E-R process to select the 'correct' response in phase 2.

In the specific *transfer-of-control* example given in figure 4, over the first two phases outcomes (O1 and O2) are common to S1 and S3, and S2 and S4, respectively (given that in phase 1 the correct responses are made to obtain those outcomes). As a result, when Phase 3 (transfer test) occurs, since the animal/human has learned S1 and S3 according to the same outcome (O1)—that is, it has formed S1-E1 and S3-E1 associations—S3 automatically cues the response associated with E1 (learned in Phase 1), in this case R1 substituting for the external stimulus. No new learning is required for this in spite of the fact that the subject has not been exposed to the particular (external) stimulus-response pairing (S3-R1) previously.

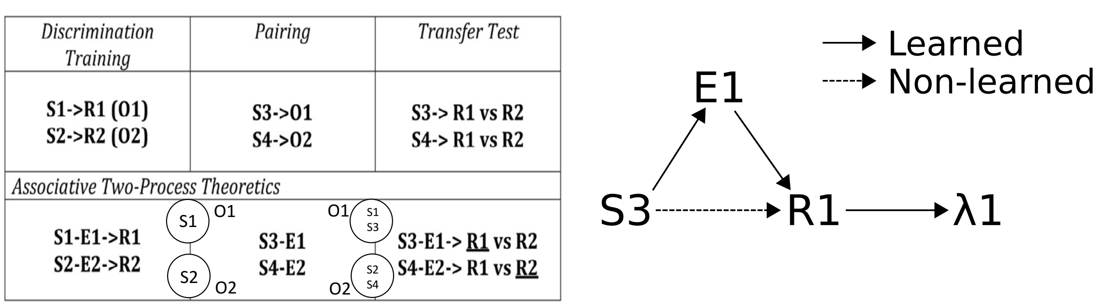

Figure 4. Transfer-of-control paradigm. The conditioning consists of three phases: Phase 1 (Discrimination Training) —an initial instrumental phase where different stimulus-response (S-R) pairings (S1-R1, S2-R2) yield different outcomes (O1, O2); Phase 2 (Pairing) — a pavlovian learning phase where new stimuli are presented and associated with previously experienced outcomes; Phase 3 (Transfer Test) —an instrumental transfer phase where the stimuli from phase 2 are re-presented as are the response options from Phase 1. ATP theory predicts that responding in the transfer test (phase 3) will be based on already existing S-E and E-R associations learned from the first two phases where the theorized preferred selections (underlined Rs) are shown in the left diagram and the S3->R1 choice process is schematized on the right. Left diagram adapted from Urcuioli (2005).

8 The first and second phase of the transfer of control paradigm can, in fact, be presented in any order though more standardly the initial instrumental phase is used first.

ATP postulates, therefore, that by way of a (dual-route) structured learning process, a type of transitive inference is possible to find correct responses in the test phase without the requirement of learning. S3-R1 associations have not been learned at the beginning of the test phase, but previous experience allows for a transitive performance of the form A->C (S3-R1) derived from A->B (S3-E1), B->C (E1-R1).

The transfer-of-control problem does not entail a designed transitive inference problem but animals and humans appear to resolve this problem by utilizing internal hidden stimuli (states) through which inference can be made.

However, in the schematic (figure 4), the associative learning processes are not omnidirectional and thus the property of retrievability given by Halford et al. (2014) is lacking for such a network to be considered to imbue higher (relational) cognition. It would instead conform more to what Halford et al. (2012) terms *implicit transitive inference* (as opposed to *explicit transitive inference*) – a term also used in animal learning circles (e.g. Goel 2007). This provides an example whereby relational behavior does not necessarily imply relational knowledge or representation.

2.3.2 Affective-Associative Two-Process Modelling

The Affective Associative Two-Process model that we developed (Lowe et al. 2014; Lowe et al. 2016; Lowe & Billing 2017; Lowe, Almér et al. 2017) merges Associative Mediational Theory (Overmier & Laury 1979; Kruse & Overmier 1982) and Associative Two-Process theory (Trapold 1970; Trapold & Overmier 1982). It does so by modelling the differential expectancies of ATP ("E") in terms of differential reinforcement outcomes. In such cases, differential outcomes can take the form of differential reinforcement magnitudes (Peterson & Trapold 1980; Delameter et al. 2012, 2017) or differential omission rates/probabilities (as studied by Urcuioli 1990; Kruse & Overmier 1982).

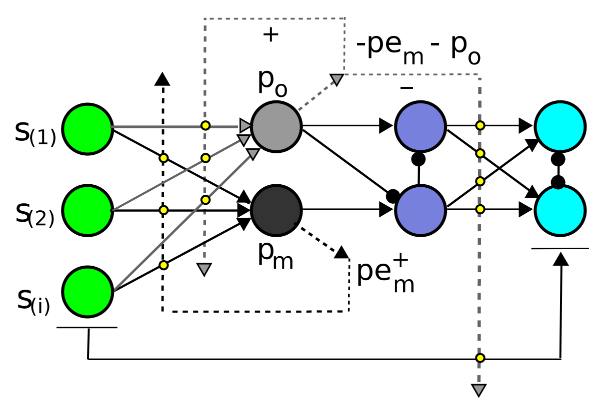

Figure 5. Neural computational model of the theorized Affective-Associative Two-Process (Lowe et al. 2017). This model extends that of Balkenius & Morén (2001)/ Morén (2002) by a) adding an output node (blue) to which the omission probability node (grey) provides excitatory (rather than inhibitory) input and which also receives inhibitory input indirectly from the (blue) output node to which the magnitude node (black) excitatorily connects, b) have a differential response layer (cyan nodes) for which reinforcement associations bias selection, c) a dual-response process whereby responses can be directly biased by reinforced stimulus associations. Key: pe_m = magnitude prediction error, p_m = magnitude prediction, p_o = omission prediction. For the mathematical description of this model see Appendix C.

The Affective-ATP neural network model (figure 5) has previously been depicted as an adapted Actor-Critic architecture (e.g. Lowe et al. 2017) but here is depicted as a feedforward ANN. This depiction allows us to compare with the earlier-mentioned models but also show the ANN as a type of representational rank 1 cognitive process (Halford et al. 2007, 2014). The model embeds the Balkenius and Morén (2001) model (which in turn embeds the Rescorla-Wagner 1972, model) into an Associative Two-Process (ATP) via adding a ‘pessimistic’ computation of the output of the value computation (reward omission probability expectation) to the ‘optimistic’ computation of the Balkenius and Morén model (reward acquisition probability expectation). These two affective value state representations can then be associated with different responses and are updated dependent on the reward (and omission of reward) outcomes the responses yield. These pavlovian-affective representations of expectancies serve to implement Associative Mediational theory embedded within the ATP.

The feedforward ANN depiction of the Affective-ATP model highlights how transitivity of choice (Kahneman & Tversky 1986, Regenwetter et al. 2011) is computationally processed. The direct/habitual route S-R (horizontal arrow at the bottom of the figure) provides the relation to be inferred in the absence of explicit learning of this association (see previous section on transfer of control). The connections between successive layers provides the means for ‘inference’ (when associatively learned). This process implements the S-E and E-R route illustrated in figure 4 (right hand side) and occurs as follows: i. the omission and magnitude value dimensions of the external stimuli (S1, S2, etc.) are learned and processed, ii. these values are input into affective value states (‘optimistic’ reward acquisition probability inputs and ‘pessimistic’ reward omission probability inputs) and are non-linearly transformed (via differentially parameterized logistic functions) so at to allow for ‘categorized’ semantic outputs to iii. form associations with responses. This ‘categorization’ disambiguates the control that affective states can have over responding. In this model, the E (expectancy) component can thus be seen as having two stages: i. value dimension computation, ii. affective value computation. The model, as it builds on, and can collapse to, the Balkenius and Morén (2001) model (and in turn that of Rescorla-Wagner 1972) is capable of resolving the sequential transitive inference problem mentioned in section 2.1. through *implicit transitive inference*. As mentioned above in this sub-section it also carries out another form of transitive inference using its structured hidden state.

However, as is made clear by the feedforward ANN depiction (figure 5) it does not satisfy the omnidirectional criterion for higher cognition necessary for retrieval of propositional (and sub-propositional) components. Delameter (2012) – see also Delameter et al. (2017) – has also proposed an ANN model to capture differential outcomes data. However, this is also a feedforward ANN (does not permit omnidirectionality) and has not been used to capture *transfer-of-control* data – it is not clear, therefore, that the model allows for implicit transitivity (transitivity of choice). Furthermore, for the purposes of this article, the model does not represent affective value.

2.3.3 Extending Affective-ATP Processing: Beyond Dual-processing

While we have grounded the Affective-ATP model in the pavlovian learning mechanisms of the Rescorla-Wagner model, naturally models abound of pavlovian processes. Amsel (1958, 1992) provided a motivational-pavlovian model centred on his frustration theory of invigorated responding – subjects will work harder for rewards that are not immediately forthcoming. From Amsel’s frustration theory was derived the anticipatory frustration *directional model* of Overmier and Laury (1979), Kruse and Overmier (1982) as previously described. Other models exist that have been used in the context of Pavlovian-Instrumental Transfer (PIT) experimentation, e.g. Balleine and Ostlund (2007). PIT is a phenomenon whereby a conditioned stimulus brings to bear on the rate of conditioned responding. Cardinal et al. (2002) – see also Cardinal (2006) – described a model of affective / emotional processing in relation to pavlovian conditioning. Within this model affective states, described as “pure value states” by Cardinal et al. (2002, p.324) are constituted by learned associated external (neutral) stimuli and unconditioned stimuli. These affective states in turn are able to directly (without learning) bias responding. Such responses can be considered preparatory (non US-specific), e.g. orientation, or consummatory (US-specific), e.g. salivation to food. In figure 6 (right hand side) we have included the additional links (US-> R, S->Affect) to our Affective-ATP model (from figure 5) with the difference that the Affective-ATP model requires associative (learned) links between affect and response representations/nodes. The response options can be considered preparatory, e.g. differential orientation responses (often used in differential outcomes experiments, e.g. pointing to a matching to sample stimulus in space).

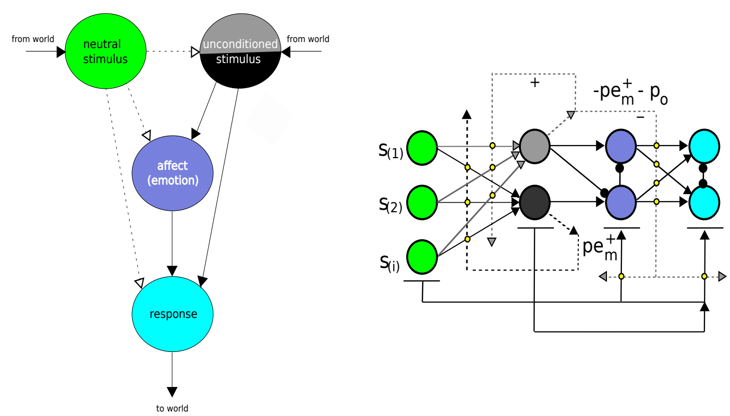

Figure 6. Pavlovian-Affective Value Model of Cardinal et al. (2002). Left. Cardinal et al. (2002) model with colour-coded adaptation. Right. Adapted Affective-ATP model to accommodate Cardinal et al. connections. Key: dashed arrows, with yellow circles, represent (associative) learnable feedforward connections; solid arrows represent non-learnable feedforward connections.

In figure 6 (left) is depicted the Cardinal et al. (2002) model whereby to-be-conditioned stimuli can be associated with both responses (thereby implementing the traditional S-R habit-forming route) and S-Affective route.

To the authors' knowledge systematic testing of the validity of each of the routes to differential outcomes data has not been undertaken. In our current implementation of the model, the S-affect route learning would necessarily be lagged relative to the S-US learning since US outputs are constitutive of the affective states and therefore, we speculate that this route might not influence *individual* differential outcomes learning and transfer of control. In the next section, however, we will discuss a role for this route in implicit transitive inference based on social learning contexts.

A further conceptual extension of the Affective-ATP model concerns incorporating punishment-based (or nociceptive) representations of stimuli, rather than just the reward and reward omission-based representations currently modelled. While work has been carried out assessing how punishment and reward representations might be combined associatively in active decision making (Lowe & Ziemke 2013; Navarro-Guerrero et al. 2017a, b), we seek inspiration from the work of Rolls (1999, 2013, 2018) concerning stimulus-reinforcer association learning theory of affect and emotions. In this theory, schematized in figure 7, emotions can be elicited as the results of experienced (primary conditioned) or anticipated (secondary conditioned) positive rewards (e.g. excitement) or omission/termination thereof (e.g. frustration, anger) but also as the results of experienced or anticipated punishers (e.g. fear) or omission/termination thereof (relief).

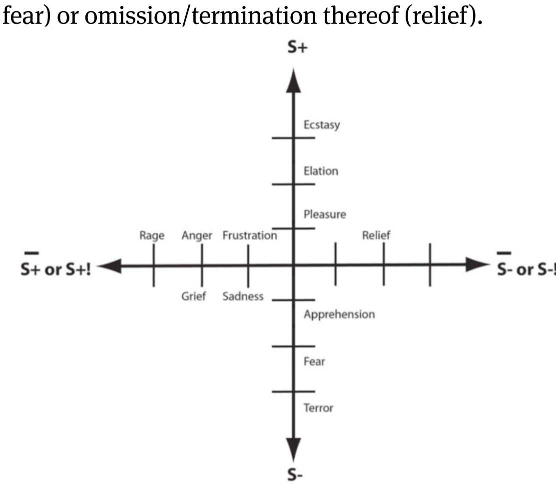

Figure 7. Rolls (1999) stimulus-reinforcer associative learning model of emotion. Acquisition of reward elicits positive emotions scaled by reward intensity (S+). Punishment (S-) elicits negative emotions similarly scaled by intensity. Omission (¯) or termination (!) or reward or punishment elicit a different suite of emotions scaled by intensity.

Our Affective-Associative perspective is consistent with Rolls' (1999, 2018) insofar as we emphasize the importance of reward contingencies rooted in the value dimensions of acquisition (magnitude) and omission. We claim that these dimensions then lead to *optimistic* and *pessimistic* affective states, respectively, whose influence on behavioural responses are designed to maximize future reinforcement.

3 Bridging Affective-Associative Processing and Relational Cognition

Thus far we have discussed associative feedforward neural networks, their ability to imbue affective representational states and their abilities to perform implicit relational cognition in the form of implicit transitivity. The feedforward depiction of the previous models promotes a view that those networks provide 'structure sensitive' cognition (Fodor & Pylyshyn 1988). However, as Fodor and Pylyshyn (1988) note in reference to such feedforward relational semantic networks (e.g. that implement John->loves->the girl: "the links in Connectionist diagrams are not generalized pointers that can be made to take on different functional significance by an independent interpreter, but are confined to meaning something like 'sends activation to' " (p.17). Therefore, the nodes of the network do not have value beyond their activation by fed-forward input constituents and cannot engage flexibly in alternative relational contexts. Eliasmith (2013) utilized the term 'semantic pointer' to allude to bidirectional networks (e.g. autoencoders) for which relatively invariant representations (for clusters of nodes/neurons at higher levels in the hierarchy) are able to: i) retrieve information from their constituents, but also ii) be used independently of their constituents as symbols in relational activity. With respect to the Affective-Associative models referred to in the previous section, this suggests that relational cognitive functionality may be limited to temporal relations, e.g. *S1 succeeds E1*, *R1 succeeds E1*.

3.1 Context-Specific Value Relational Representations

Notwithstanding the limitations of feedforward connectionist models for imbuing cognitive (logico-relational) properties, explanation is needed for how symbolic-relational models are grounded in (or connected to) associative processes that extract correlative patterns in the world (Harnad 1990) from which 'meaning' may be constructed. Grounding relational knowledge in the real world is critical for seamless interaction in the world for physically embodied systems (humans, but also robots). But it may also be the case that such relational systems can't be fully understood without recourse to how they are shaped by the dynamics of the world. The spatial and temporal dynamics of higher cognition (e.g. forming semantically constituted relations) is at the very least constrained by those of the world but may also exploit such dynamics in order to learn and develop higher cognition.

Leech et al. (2008) provided an example of how relational knowledge might be developed via exploiting the temporal ordering of propositional information input to a connectionist architecture. In their proposed recurrent (bidirectional) model of relational priming and analogy, priming is done by learning the temporal order or relational transformations (before and after states), e.g. apple & knife is a before state and a cut apple & knife, is an after state. The network was trained in accordance to this temporal ordering whereby the 'before' state inputs were clamped (fixed temporal presentations to the network) while transformation weights (for activating after states) were learned. Subsequently, before and after states were clamped in order to learn hidden (transformative) representations that would allow for retrievability of before (from after) and after (from before) states. Such relational learning that is grounded in the spatial and temporal dynamics of the world may provide the building blocks for learning through structural alignment (Halford et al. 2014) whereby analogies can be made based on learned structural relationships in the world (taller than, faster than, precedes, succeeds). It must also constrain the sorts of spatial-temporal dynamics that may occur in neural systems dedicated to analogical reasoning that have hitherto been considered only in a disembodied context (Hummel & Holyoak 1997, Doumas et al. 2008).

3.2 Bridging Implicit and Explicit Relational Cognition Neural Networks

As mentioned in the previous section, the Affective-Associative Two Process model, as a feedforward artificial neural network, structurally permits linguistically formulated transitive inferential logic, i.e. the premises *succeeds(A,B)*, *succeeds(B,C)* and inference *succeeds(A,C)*, is inherent in the neural network connective structure. However, the inferential process entails *learning* of the premises in order to arrive, without learning, at the inference. This deviates from Halford et al. (2010, 2014) and Fodor & Pylyhsyn's (1988) conceptions of transitive inference who view connectionist schemas as limited regarding what they can relationally represent. For Halford et al. (2010) the premises are not learned[9] but concern 'one shot' manipulation of symbols, i.e. representations that are independent of process (e.g. unlike feedforward neural network activation). Halford et al. (1998, 2010, 2014) have postulated a tensor product structure for neural networks whose *n*-dimensional (vector) complexity represents *n* minus 1 relational complexity, e.g. the relation *loves*(John, Mary), consisting of 3 vector representations (2 objects, 1 relation), has *binary* relational complexity and is of representational rank 3. Fodor & Pylyshyn (1988), on the other hand, likening connectionist models to graphs and in reference to a relation "John -> loves -> the girl" suggest: "Connectionist graphs are not structural descriptions of mental representations; they're specifications of causal relations. All that a Connectionist can mean by a graph of the form 'X → Y' is: states of node X causally affect states of node Y … the graph can't mean 'X is a constituent of Y'" (p.17).

3.2.1 Case 1: Rank 2+ Autoencoding relational cognition model

In Halford et al. (2007) it was suggested that in order to (developmentally) transition from a representational rank 1 neural network to a rank 2 network, it would be required to divide the hidden layer of a multi-layered network (with one hidden layer) into (at least) two partitions: "The transition from Rank 1 to Rank 2 can be envisaged by imagining the hidden layer at Rank 1 … being divided into two components which are then connected so as to form a matrix" (Halford et al. 2007, p. 2). Such a connectivity schema might be envisaged as a recurrently connected neural network (e.g. of a form related to Leech et al. 2008). Halford et al. (2014) has suggested that some forms of recurrent, or auto-associative, neural network may alleviate such problems, e.g. autoencoders that are able to (auto-associatively) re-present inputs as idealized values (based on the statistics of previous learning) at an output layer. Wilson et al. (2001) later demonstrated the extent to which such an auto-associative (on an autoencoder based multi-layered perceptron – one hidden layer) could permit 'accessibility' of items in a binary relational proposition, e.g. *John Loves Mary*.

The non-standard[10] autoencoder was structured according to its having *object*, *subject* and *relation* items selectively connected to a hidden layer, where *object* and *subject* could be viewed as arguments to the predicate (*relation*). It was structured so that two of the three components projected to one of three separate hidden layer partitions. These partitions then provided direct output to a partitioned output layer representing the single other component, e.g. *object* (pizza) and *subject* (Mary) input representations were represented in a hidden layer that projected to the output *relation* (likes) – see figure 8, left. The network connectivity, therefore, was such that it had the potential for one-to-one, one-to-many, many-to-one and many-to-many queries. For example, for the latter query type *Who is frustrated by what?* (in predicate calculus: *frustrated(X,Y)*?) the dot product computation could produce an effective rank 2 tensor product network, i.e. a matrix giving all frustrated people and all frustrated situations (see figure 8, right). However, this network was found to have 'limited accessibility' (Wilson et al. 2001). It was best able to access single elements from one-to-one queries, e.g. *What is John frustrated by?* (*frustrated(John, Y)?*) yielding in the example in figure 8 the output 'task'. Where one-to-many (vector) outputs provided the target output, the network was more limited and suffered from incorrect generalization – if the encodings of different item

9 Or at least the learning of the premises is not a major concern from the modellers' perspective.
10 The autoencoder did not have full connectivity between layers nor was the output layer trained to re-present the outputs of the input layer.

(*relation*, *subject* or *object*) vectors overlapped, i.e. had constituents similar to each other, there was greater scope for erroneous output, i.e. accessing items that weren't appropriate to the relation. This limitation is apparently not the case for tensor product networks that are able to resolve many-to-many queries and are thereby considered omnidirectional (have full accessibility to all possible queries made on the stored propositions). This occurs since each proposition has a unique symbolic representation whose querying (through dot product computations) can yield the appropriate outputs whether single values, vectors or rank 2 tensor product networks.

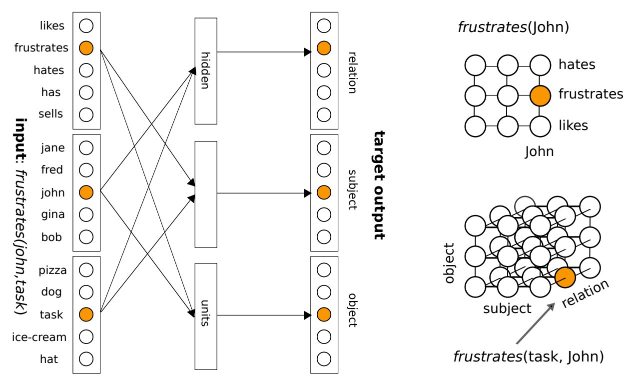

Figure 8. Connectionist and Symbolic Neural Network Implementations of Unary (Rank 2) and Binary (Rank 3) Relations. Left: Adapted autoencoder of a binary relation of the form *relation(subject, object)*. Each input element connects to corresponding hidden layer partitions that connect to an output layer representing the two other element types (e.g. a subject element connects to hidden layer partitions that output to object and relation targets). Adapted from Wilson et al. (2001). Right: Tensor product networks (Wilson et al. 2001, Halford et al. 2014 following Smolensky 1990), also known as the STAR architecture. Input vectors representing subject, object and relation (rank 3, lower network) provide inputs to a tensor product network that captures the relation in a single symbolic neuron. From this neuron it is possible to access the other input values in the relations (omnidirectionality) using a dot product query. A binary relation provides the output of the query "who feels what about the task?", given in predicate calculus by *P(X,Y)* where *P* and *X* are the terms being queried. The upper figure shows a tensor product network for a unary (rank 2) relational neural network implementing a subset of the relations and objects of the Wilson et al. autoencoder. Adapted from Halford et al. (2014).

3.2.2 Case 2: Rank 2+ Hierarchical structuring of relational cognition

The lack of semantic similarity (via implementing overlapping semantic units as relational constituents) of STAR (figure 8, right) as a model of analogical mapping has been criticized by Hummel and Holyoak (1997). Semantic similarity, while potentially providing a problem of overlapping constituents leading to incorrect generalization, provides a property through which analogies may be learned.

An alternative ANN approach to the Wilson et al. (2001) model and to STAR (to which Wilson et al. 2001, Halford et al. 2014 attribute *omnidirectionality*) is that provided in the LISA/DORA framework (Hummel & Holyoak 1997; Hummel & Holyoak 2001; Doumas et al. 2008; Morrison et al. 2011; Holyoak 2012; Knowlton et al. 2012; Doumas et al. 2018). The problem of overlapping activation of constituent elements to propositional component representations, referred to in Case 1, is overcome by activating each sub-proposition (and its constituents) of a proposition (e.g. *John loves*; *Mary is loved* of the proposition *John loves Mary*) one at a time. So, all constituents of *John loves* are synchronously activated and this pattern inhibits the *Mary is loved* pattern. The *John loves* pattern, through self-inhibition, loses activation and simultaneously disinhibits the *Mary is loved* pattern (Knowlton et al. 2012). A high frequency oscillation (relative to the neural states representing full binary propositions) between the two patterns is said to allow for *role-filler* binding (of the "sub-propositions") so as to input a stable pattern to the compound neural representation that encodes the full proposition *John loves Mary* that is of representational rank 3 (Halford et al. 2014). This process is schematized in figure 9. Role-filler binding constituents may overlap in their representation – they might even involve *the same* objects (fillers) and predicates (roles) but LISA/DORA exploits time (temporary representations involving re-use of components) so as to bind position-sensitive (and thereby meaning-sensitive) role-fillers to propositional statements. Hummel et al. (2004) has suggested that a critical feature that LISA (and DORA) possesses that STAR lacks is role/filler *independence* – that is to say that the roles (predicates), e.g. loves, and fillers (object, subject), e.g. John and Mary, that are used in a relational statement should not be dependent upon the particular statement. John could be both ‘the lover’ and ‘the beloved’ and should not be *re-presented* in these different roles, which is a requirement for STAR.

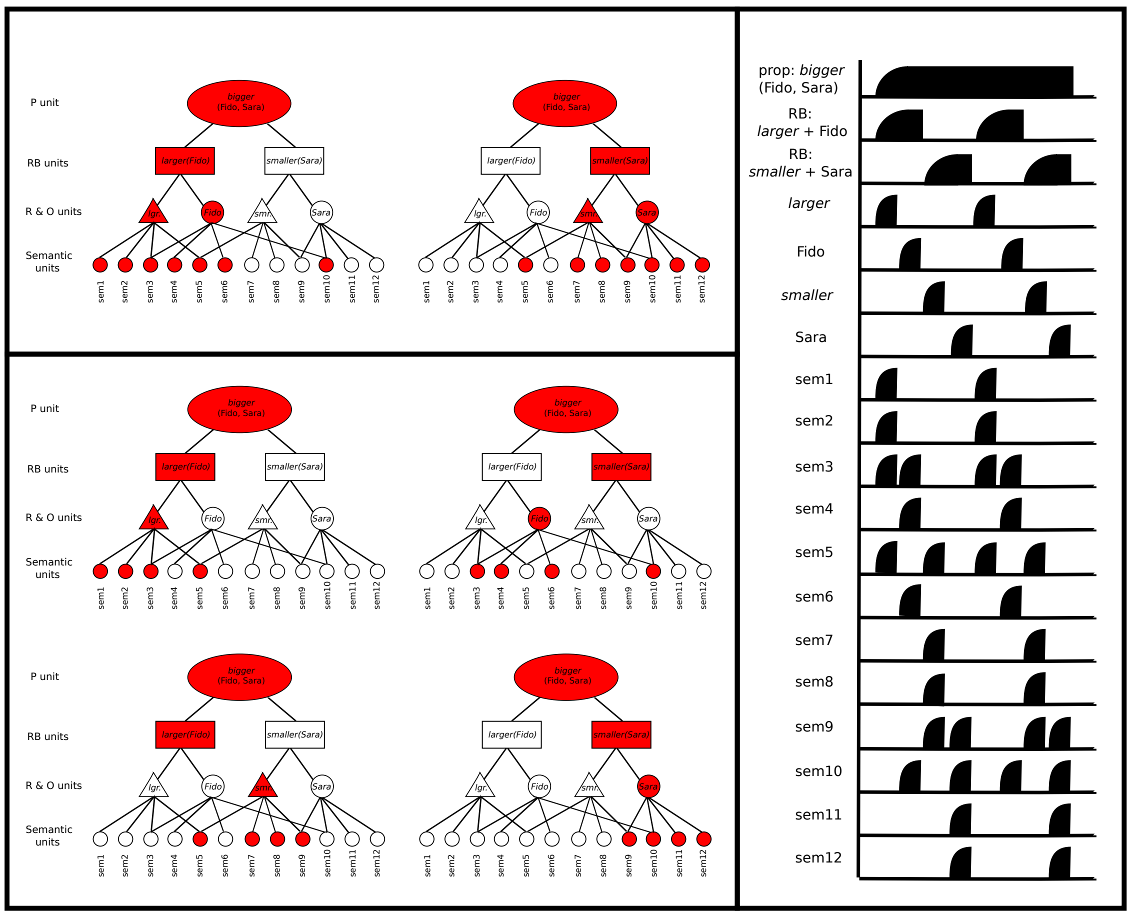

Figure 9. The LISA/DORA Connectionist-Symbolic Architecture of Relational Cognition. Left. Top – LISA architecture, role-binding (RB) units of the proposition (P unit) neural representation of *bigger*(Fido, Sara) fire synchronously with all constituent neurons (object/O unit, predicate/role-filler or R unit and their semantic constituents/units) but asynchronously with each other, i.e. the sub-proposition *smaller*(Sara) and its constituents are inhibited by *larger*(Fido) and its constituents. Bottom – DORA architecture, the object/predicate constituents of role-binding (RB) units fire asynchronously with each other. Right. The DORA firing patterns follow hierarchical oscillatory frequencies allowing for higher units to entrain lower units relevant to analogical learning.

Learning and retrieval (from LTM) by analogy in DORA is guided top-down by a propositional unit and occurs in reference to semantic similarity (amount of overlap) between a given sub-proposition’s constituents (semantic units) and those of another that may give rise to analogous propositions but can also allow for the learning of new predicates. Oscillatory (dynamic binding) activation of overlapping representations provides the means to learn analogies since distributed semantic representations form the basis of analogies (semantic similarity) that can be disentangled (through temporal asynchronous activation) when learning or retrieving specific role-filler bindings.

The historically earlier LISA uses the oscillatory binding (also referred to as dynamic binding) mechanism to drive retrieval, and reinforcing in memory, of analogous propositions. This requires a *Driver* proposition unit to top-down activate its hierarchically ordered constituents as illustrated in figure 9 (top). Then analogous (*Recipient*) propositions are driven bottom up through activation of semantically similar units (the constituents of the analogous proposition). Activation of these units in the analogue then feed forward through the relational hierarchy – see figure 10. Activation in Object (O) and Predicate (P) units feeds forward to a given role-binding (RB) unit that is invariant to sub-optimal activation in either unit and is passed forward to the proposition unit (P). In turn activation feeds back down the hierarchy so as to enable those O and P unit constituents of the RB unit to more cleanly win the competition against rival object and predicate units allowing in turn for stronger representations in the (higher) relational levels. This is achieved through a global inhibition mechanism and guards against multiple RB units being simultaneously activated. The Driver oscillates between its synchronous RB unit constituent activations thereby permitting the propositional statement (P) to be retrieved using feedforward and feedback dynamic entrainment of the RB unit constituents of the (*Recipient*) analogous proposition. All of the analogue recipient's local units remain active for a period (i.e. have a prolonged oscillatory phase) along with those of the 'driver' analogue allowing hebbian learning (LTM) for the like-for-like local units (mapping). This is such that future inducements of Driver activity allow for direct activation of the local constituents of the analogous proposition rather than requiring feedforward activation via semantically similar units (see figure 10). LISA thereby achieves analogue retrieval and analogue learning (local unit mapping) through use of its conceptually critical oscillatory binding mechanism.

The LISA/DORA framework has been criticized on account of not providing a description as to how semantic constituents (above all predicates) are learned in the first place (Halford et al. 2010)[11].

In order for any form of unary or binary relational knowledge to be acquired to bring to bear on analogical learning and reasoning, some degree of associative learning must precede it. Refined object and relational knowledge may be limited such that inaccurate or very holistic (Doumas et al. 2008, O'Reilly et al. 2017) knowledge of objects may still have a top-down influence on what is learned through association. In essence, any top-down knowledge that serves to focus attention on the constituents of the objects and predicates relevant to relational knowledge may benefit the acquiring of both relational knowledge and knowledge of the featural constituents of objects and predicates.

Bottom-up associative learning ('System 1') requires the use of space and time to resolve the binding problem of apprehending the separateness and identity of individual objects. For example, an unfamiliar object occluded by another may appear to young infants (Baillargeon 2004, O'Reilly et al. 2017) as a blend of the two whereas more distal objects might be easier to process as being separate. However, apprehending their co-occurrence in a particular situation requires some kind of embodied oscillation (moving towards, orienting, foveating) between one object and another necessitating the use of *time*, i.e. sequential processing of the stimuli/objects relevant to the scene. In this sense the bottom-up process may recapitulate the top-down oscillatory binding process proposed by LISA/DORA as constrained by the embodiment of the individual and the physical characteristics of the world. Notwithstanding, the possibility to learn the constituents of objects and predicates that engage in relational knowledge, structural separation of the constituents of propositional knowledge (binary relations) is required: "Simply jointly activating patterns representing "John", "Mary", and "Loves" cannot distinguish "John loves Mary" from "Mary loves John" (or even from a description of a narcissistic hermaphrodite)" (Hummel & Holyoak 1997, p.13).

11 It has also been criticized regarding the lack of evidence for oscillatory dynamics in the brain encoding for formation of relational information (see Eliasmith 2013).

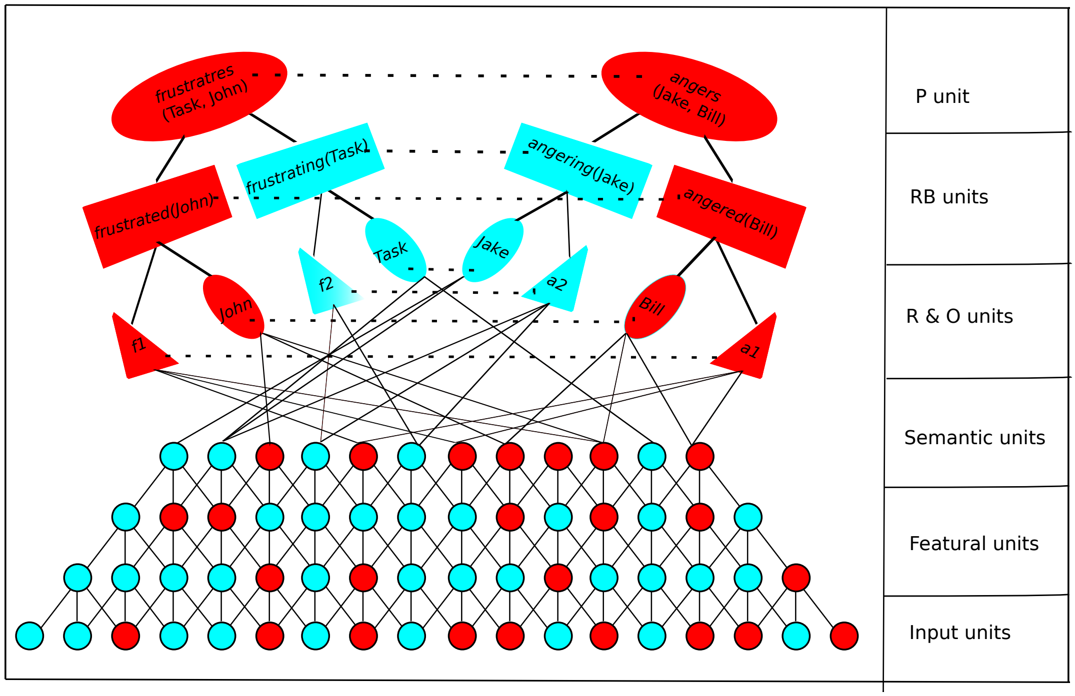

Figure 10. Analogical Retrieval and Learning LISA/DORA. The figure depicts the LISA/DORA architecture following analogical retrieval where the proposition *frustrates(Task, John)* and its constituents is analogous to *angers(Jake,Bill)*, for example, as it applies to a related task or interaction. Prior to learning, the *Driver* (left-most Proposition) top-down activates semantic units. *Recipient* (right side) R & O units propagate activity upwards as a result of semantic similarity (i.e. where Driver activated semantic units concurrently activate the Recipient's R&O units). The dashed horizontal lines indicate that following analogical retrieval the individual unit layers of the analogues are associated (stored in long-term memory) permitting top-down retrieval in future iterations. In this depiction we add featural (and input) units so as to highlight the possible interface of parallel distributed (System 1) and localist (System 2) systems (adapted from Hummel & Holyoak 1997).

Top-down driven processing, starting either from propositional (binary) relational knowledge (LISA) or (during learning) from unary relational and even object/predicate based representations (DORA) are viewed as being critical for the formation of analogical relational structures. The mechanism of oscillatory binding precludes simultaneous processing of overlapping semantic constituents *for a given proposition*, e.g. where objects John and Mary share semantic features – both are human, have noses, etc. This enables role-filler bindings (e.g. *John is frustrated*, *Mary frustrates*) to be kept distinct (Knowlton et al. 2012). As for the bottom-up associative learning binding problem, relational knowledge is constrained by memory storage capacity entailing overlapping constituents of objects and predicates. LISA/DORA resolves this through the use of time, i.e. the sequential processing of unary relations and their constituents, and additionally in DORA, objects and their constituents followed by predicates and their constituents. In this manner, the semantic constituents of unary relations, objects and predicates maintain their identity in spite of potentially being semantically similar but nevertheless are apprehended as belonging to the same scene (relation in this case).

Top-down driven processing, insofar as it activates constituents of objects and predicates, might also serve to activate (through asynchronous oscillation) featural (non-semantic) constituents potentially providing the means for refining the learning of such featural constituents. This would occur through providing a discriminative attentional mechanism to facilitate bottom-up associative learning means of resolving the binding problem (i.e. using physical space and time as opposed to that contrived by top-down oscillatory activation). Relational context can provide a means of disambiguating between objects that are sensorially (not just semantically) similar. Such refinement could be viewed through the lens of the predictive coding framework (Friston 2010, O'Reilly et al. 2017).

3.3 Bridging Connectionism and Relational Cognition through Affective-Associative Processing.

As acknowledged by Doumas et al. (2008): “The model, as it stands, does not speak to where the semantic invariants ... come from.” (p.33). To the authors’ knowledge work to the present-day using DORA (and LISA) has not addressed this. DORA and LISA have been concerned with how relational knowledge can be acquired through analogy rather than how the semantic constituents of objects and predicates for a given relation may be acquired as a result of interaction with the outside world. Furthermore, Halford et al. (2010) cites as a future question for the LISA/DORA connectionist-symbolic framework to address: “What is the precise nature of the link between dynamic binding in working memory and acquisition of relational knowledge”, (p. 503).

From this we can derive two research questions required to be addressed in order to arrive at a fuller account of how a connectionist-relational architecture can explain the integration of system 1-like and system 2-like processing in cognitive agents whilst accounting for the *omnidirectionality* property (through retrieval from long-term memory – figure 10 – noted by Halford et al. 2010):

1. *How are the semantic – object and predicate – constituents learned within a connectionist-relational architecture?*
2. *How do top-down driven relational activations and bottom-up associatively learned activations interact through dynamic binding?*

In figure 11 is presented a hypothetical architecture integrating an affective-associative representational rank 1+ model (section 2) with the LISA/DORA architecture. The affective-associative components (from left to right of each plot up to but not including the R & O units) concern figure 6 (Cardinal / Affective-ATP modelling) where for clarity of visualization arrows indicating learnable or non-learnable connections are conflated into the same type. The affective mechanism is also inspired by Rolls (1999) stimulus-reinforcer associative learning perspective on emotion elicitation (see figure 7).

Additional to the affective-associative processing network is depicted a replicated pair of affective units representing perceived affective states (e.g. through facial expression) that, through mirror neuron activation (De Gelder 2009) may be vicariously experienced by the perceiving individual. Such a social dimension is a key element of affective processing but also potentially for relational knowledge concerning other subjects. For details on vicarious learning using mirror neuron systems the reader is referred to Lowe et al. (2016) – space precludes discussion here. Constituent units may now be considered as hierarchical somatic (top) and object-based stimuli (bottom) comprising semantic (invariant) and non-semantic featural (non-invariant) components. Lower down the hierarchy neurons encoding simpler features are more sensitive to variation in the inputs and for different objects are expected to increasingly overlap based on the sensory properties of the presented stimuli.

The figure represents a connectionist-relational architecture of an adult human that deploys the driver – *frustrates* (Task, John) – proposition to exert top-down synchronous entraining of activation to all its (semantic and featural) constituents. The affective-associative component left-to-right illustrates a reward-based affective neural network blending that of the Affective-ATP and Cardinal et al. (2002) models. It could be naturally extended by having additional reinforcer units that represent punishment (or nociception) – and omission thereof – allowing for a Rolls (1999)-relevant model of affect to bring to bear on relational-cognitive processing.

This constitutes a form of *constituent refinement*. In the above case, however, the activation can be viewed as bidirectional. Bottom-up sensory processing of incoming stimuli (John and Task) whose constituent features may overlap serve to activate object nodes in the R & O layer. Naturally, these object nodes may misrepresent the *subject* and *object* ordering of the proposition *Task frustrates John*, and even more intuitively in its retrieved analogue (figure 10) *Bill angers Jake*. Thus, discrimination benefits from top-down activation focusing processing on one or other object (via oscillation) and its semantic and featural constituents. Such a process would be of less obvious utility in the absence of sensory feedback (during which time *precision tuning* to error might be dampened, see Friston et al. 2010 who provide a predictive coding argument for the existence of such a mechanism). In the case of there being sensory input a particular type of hierarchical processing should occur. Here the level of top-down oscillatory binding should be constrained to that which is sensorially available. For example, if sensory feedback for *John + looking frustrated* is available, asynchronous oscillations between the *John* object and the *is frustrated* predicate (role) would be attended to in order to refine the constituents of both – for *John* the semantic and featural components (e.g. nose, but also non-semantic features of the nose may be important), for *is frustrated* the visual (and extra-visual, e.g. prosodic) sensory properties, such as facial action units (Ekman 2003), may be attended to and refined. All the while, the higher relational level of *John is frustrated* entrains the attentional focus on its object and role constituents (asynchronously activated). The propositional unit – *John is frustrated by Task*, would entrain the R & B units' activations and their constituents oscillating between them in order to disambiguate the overlapping constituents. This level is 'motivationally' significant since constituent refinement (for both semantic and non-semantic features) for similar objects is important insofar as the objects have different roles in a relation. Constituent refinement of predicates (*frustrating* versus *is frustrated*) would similarly be important so as to distinguish my frustration (*frustrating*) from that of another (John's) and to what extent.

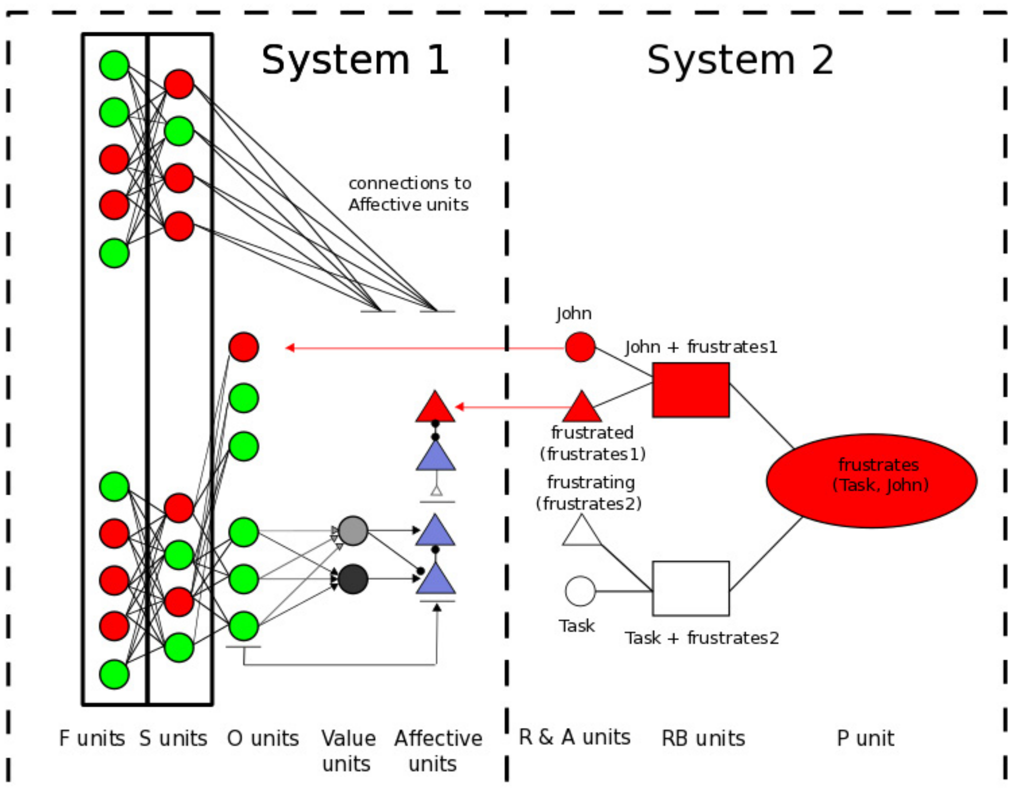

Figure 11. Affective-Associative System 1-System 2 Network. Affective-Associative network extended through representation of its 'deep' (external) stimulus and affective constituents. F units (featural units as in figure 10) and S units (semantic units) are boxed to indicate that they are generative layers. Driver activation from the P unit flows through Role-Binder (RB) units and their object (O unit) and affective unit constituents. These units in turn generate activation (or bias activation) in the S units (that are then used to retrieve analogical propositions, e.g. *angers(Bill, Jack)*). S units simultaneously generate activation in their featural constituents (F units). Attention may be focused on any RB unit constituents but will be biased by simultaneous bottom-up activation. This attention will also help refine constituent representations, e.g. the affective expressive constituents of John (and by analogy Jack) and the facial features of Jack (or task stimuli features). Generative activation may also undergo *precision tuning* (Friston 2010) whereby attention to semantic and featural constituent details may be greater or less depending on *unpredicted* fed forward activation. In the case of lowered attention achieved through such a precision tuning mechanism, analogical retrieval may occur relatively unencumbered. In the case of high attention (to the external stimuli) analogical retrieval may be deprioritized.

The above provides an approach for implementing constituents of object and predicate units, essentially using deep generative neural networks, whereby top-down activations allow for the generation of activations of constituent units (as for autoencoders or deep Boltzmann machines – Hinton & Salakhutdinov 2006) or the biasing of activations of the constituents (as for 'VizNet', Rolls 2016).

The affective representations thereby serve to i) learn predicate and proposition formation as grounded in real world interactions, ii) top-down entrain agents to focus on affective relational constituents (featural and semantic), i.e. to refine and learn features of objects and predicates (internalized affective states and those expressed by another). The latter mechanism serving to better categorize those very objects and predicates constitutive of the 'predicted' relation. The interaction between, on the one hand, this bidirectional activation that serves learning about propositions in the world and, on the other hand, analogical learning and retrieval, is then potentially enabled by an attentional mechanism (such as precision tuning, Friston 2010) wherein unpredicted external stimuli/events focus agents on the outside world (activation flows down the hierarchy, see figure 10), and predicted stimuli/events permit activation to flow upwards from semantic units to recipient analogue localist units (desensitizing activation from inputs units).

The experimental testing of such a bidirectional mechanism would not be without challenges. As for the Frank et al. (2005) experiment mentioned in section 1.3., performance measures combined with subjective reporting of the understanding of the rules of a given task could distinguish associative 'strategists' from relational 'strategists'. Doumas et al. (2018) have sought such evaluations through manipulating task difficulty whereby associative strategizing is less likely to pay off for more complex relational problems. An experimental set up that constrains sensory input to participants so that one object (John) is presented followed by another (Task) in an oscillatory fashion and at different rates might also bring to bear on how easily relational knowledge can top-down entrain associative processing and how easily analogical knowledge can be utilized on a given task.

4 Discussion

We have presented a view on the bridging of connectionist and relational cognitive architectures using affective-associative neural network modelling and a review of connectionist models and their representational ranks (Halford et al. 2007, 2014). Halford et al.'s (2007, 2014) ranking system broadly distinguished between System 1 (associative-based rank 0 and rank 1 representations) and System 2 (relational-based representations) like knowledge where the former entails the use of associative learning mechanisms and is tied to interaction in the world and the latter permits higher cognitive functions concerning relational knowledge subject to the property of omnidirectionality. We have focused our attempts at bridging these two considered types of representational knowledge according to the encoding of affective states as predicates – a subset of predicate knowledge, but an important subset nevertheless. A fuller connectionist architecture should account for where all elements 'come from'.

In presenting feedforward affective-associative neural networks imbuing representational ranks 0 and 1 we considered how such forms of processing can allow for implicit relational knowledge and that it may permit the *grounding of* symbolic connectionist states (Harnad 1990) by exploiting the spatial and temporal dimensions of physical embodied interaction (Leech et al. 2008). We finally discussed the possibility that System 1 and System 2 like processing may be part of a single unified generative process. In this view, bottom up processing is entrained by top down processing driven by 'looking for' object and predicate relations in the world. Such processing can focus embodied attention whilst simultaneously refining the semantic and featural constituents of those relations (as well as those of analogue relations).

A major motivation for our theoretical work is practical application and dealing with the hard problem of engineering physically embodied agents (e.g. robots) so that they may make sense of the patterns of sensorimotor activity that impact them (e.g. Li et al. 2013, 2014). Affective states, grounded in dimensions of stimuli valuations, can take many forms – facial (Ekman 2003), prosodic (Schröder 2001), tactile (Andreasson et al. 2018). Such diverse dimensions can all be argued to entail a predictive / generative process in order to make sense of them (Lowe & Ziemke 2011, Morrison et al. 2013, Barrett et al. 2016, Lowe et al. 2017). In future work, we will strive to develop a predictive/generative architecture that permits affective states to bridge connectionist/associative and higher cognitive-relational capacities in the service of intelligent and adaptive robots and furthering cognitive scientific understanding.

References

Amsel, A. (1958). The role of frustrative nonreward in noncontinuous reward situations. *Psychol. Bull.* 55:102–119.
Amsel, A. (1992). *Frustration theory: an analysis of dispositional learning and memory*. Cambridge University Press, Cambridge.
Andreasson, R., Alenljung, B., Billing, E., & Lowe, R. (2018). Affective touch in human–robot interaction: conveying emotion to the nao robot. *International Journal of Social Robotics*, 10(4), 473-491.
Armony, J. L., Servan-Schreiber, D., Cohen, J. D., & LeDoux, J. E. (1997). Computational modeling of emotion: Explorations through the anatomy and physiology of fear conditioning. *Trends in cognitive sciences*, 1(1), 28-34.
Armony, J. (2005). Computational models of emotion. In Proceedings. 2005 *IEEE International Joint Conference on Neural Networks*, 2005. (Vol. 3, pp. 1598-1602). IEEE.
Baillargeon, R. (2004). Infants' physical world. *Current directions in psychological science*, 13(3), 89-94.
Balkenius, C. & Morén J (2001). Emotional learning: a computational model of the amygdala. *Cybern Syst Int J* 32:611–636.
Balleine, B. W., & Ostlund, S. B. (2007). Still at the choice-point. *Annals of the New York Academy of Sciences*, 1104(1), 147-171.
Barrett, L. F., Quigley, K. S., & Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708), 20160011.
Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and economic behavior*, 52(2), 336-372.
Braitenberg, V. (1986). Vehicles: Experiments in synthetic psychology. MIT press.
Braver, T. S., et al. (2014). Mechanisms of motivation–cognition interaction: challenges and opportunities. *Cognitive, Affective, & Behavioral Neuroscience*, 14(2), 443-472.
Cardinal, R. N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, 26(3), 321-352.
Cardinal, R. N. (2006). Neural systems implicated in delayed and probabilistic reinforcement. *Neural Networks*, 19(8), 1277-1301.
de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: A case for animal–human translational models. *Psychological Research* PRPF, 73(4), 463–476.
De Gelder, B (2009), "Why bodies? Twelve reasons for including bodily expressions in affective neuroscience", *Philosophical Transactions of the Royal Society*, vol. 364, 3, pp. 3475-3484.
Delamater, A. R. (2012). On the nature of CS and US representations in Pavlovian learning. *Learning & Behavior*, 40(1), 1-23.
Delamater, A. R., Garr, E., Lawrence, S., & Whitlow Jr, J. W. (2017). Elemental, configural, and occasion setting mechanisms in biconditional and patterning discriminations. *Behavioural processes*, 137, 40-52.
Doumas, L. A., Morrison, R. G., & Richland, L. E. (2018). Individual differences in relational learning and analogical reasoning: A computational model of longitudinal change. *Frontiers in Psychology*, 9.
Doya, K. (2008). Modulators of decision making. *Nature neuroscience*, 11(4), 410.
Ekman, P., & Friesen, W. V. (2003). *Unmasking the face: A guide to recognizing emotions from facial clues*. Ishk.
Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition*. Oxford University Press.
Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1-2), 3-71.
Frank, M. J., Rudy, J. W., Levy, W. B., & O'Reilly, R. C. (2005). When logic fails: Implicit transitive inference in humans. *Memory & Cognition*, 33(4), 742-750.
Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature reviews neuroscience*, 11(2), 127.
Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
Goel, V. (2007). Anatomy of deductive reasoning. *Trends in cognitive sciences*, 11(10), 435-441.
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. (1999). A conceptual complexity metric based on representational rank.
Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in cognitive sciences*, 14(11), 497-505.
Halford, G. S., Andrews, G., Wilson, W. H., & Phillips, S. (2012). Computational models of relational processes in cognitive development. *Cognitive Development*, 27(4), 481-499.
Halford, G. S., Wilson, W. H., Andrews, G., & Phillips, S. (2014). *Categorizing cognition: Toward conceptual coherence in the foundations of psychology*. MIT Press.
Harnad, S. (1990). The symbol grounding problem. Physica D: Nonlinear Phenomena, 42(1-3), 335-346.

Heath, R. A., & Hayes, B. K. (1998). Why is capacity limited? Missing dynamics and developmental controversies. *Behavioral and Brain Sciences*, 21(6), 839-840.
Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786), 504-507.
Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological review*, 104(3), 427.
Hummel, J. E., & Holyoak, K. J. (2001). A process model of human transitive inference. In *Spatial schemas in abstract thought* (pp. 279-305).
Kahneman, D., & Tversky, A. (1986). Rational choice and the framing of decisions. *Journal of business*, 59(4), 251-278.
Kahneman, D., & Egan, P. (2011). *Thinking, fast and slow* (Vol. 1). New York: Farrar, Straus and Giroux.
Kiryazov, K., Lowe, R., Becker-Asano, C., & Randazzo, M. (2013). The role of arousal in two-resource problem tasks for humanoid service robots. In 2013 *IEEE RO-MAN* (pp. 62-69). IEEE.
Klopf, A. H., Weaver, S. E., & Morgan, J. S. (1993). A hierarchical network of control systems that learn: Modeling nervous system function during classical and instrumental conditioning. *Adaptive behavior*, 1(3), 263-319.
Kollias, P., & McClelland, J. L. (2013). Context, cortex, and associations: A connectionist developmental approach to verbal analogies. *Frontiers in Psychology*, 4, 857.
Kruse, J. & Overmier, J.B. (1982). Anticipation of reward omission as a cue for choice behavior. *Learning and Motivation*. 13(4):505–525.
Leech, R., Mareschal, D., & Cooper, R. P. (2008). Analogy as relational priming: A developmental and computational perspective on the origins of a complex cognitive skill. *Behavioral and Brain Sciences*, 31(4), 357-378.
Li, C., Lowe, R., & Ziemke, T. (2013). Crawling posture learning in humanoid robots using a natural-actor-critic cpg architecture. In *Artificial Life Conference Proceedings* 13(pp. 1182-1190). One Rogers Street, Cambridge, MA 02142-1209 USA journals-info@ mit. edu: MIT Press.
Li, C., Lowe, R., & Ziemke, T. (2014). A novel approach to locomotion learning: Actor-Critic architecture using central pattern generators and dynamic motor primitives. *Frontiers in neurorobotics*, 8, 23.
Lowe, R., Philippe, P., Montebelli, A., Morse, A., & Ziemke, T. (2008). Affective modulation of embodied dynamics. In *The Role of Emotion in Adaptive Behaviour and Cognitive Robotics, Electronic Proceedings of SAB Workshop*.
Lowe, R., Humphries, M., & Ziemke, T. (2009). The dual-route hypothesis: Evaluating a neurocomputational model of fear conditioning in rats. *Connection Science*, 21(1), 15-37.
Lowe, R., & Ziemke, T. (2011). The feeling of action tendencies: on the emotional regulation of goal-directed behavior. *Frontiers in Psychology*, 2(Dec), Article-346.
Lowe, R., & Ziemke, T. (2013). Exploring the relationship of reward and punishment in reinforcement learning. In 2013 *IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)* (pp. 140-147). IEEE.
Lowe, R., & Kiryazov, K. (2014). Utilizing emotions in autonomous robots: An enactive approach. In *Emotion modeling* (pp. 76-98). Springer, Cham.
Lowe, R., Sandamirskaya, Y., & Billing, E. (2014). A neural dynamic model of associative two-process theory: The differential outcomes effect and infant development. In *4th International Conference on Development and Learning and on Epigenetic Robotics* (pp. 440–447).
Lowe, R., Almér, A., Lindblad, G., Gander, P., Michael, J., & Vesper, C. (2016). Minimalist social-affective value for use in joint action: A neural-computational hypothesis. *Frontiers in Computational Neuroscience*, 10. Available at: https://doi.org/10.3389/fncom.2016.00088
Lowe, R., Almér, A., Billing, E., Sandamirskaya, Y., & Balkenius, C. (2017). Affective–associative two-process theory: a neurocomputational account of partial reinforcement extinction effects. *Biological cybernetics*, 111(5-6), 365-388.
Lowe, R., & Billing, E. (2017). Affective-Associative Two-Process theory: A neural network investigation of adaptive behaviour in differential outcomes training. *Adaptive Behavior*, 25(1), 5-23.
Lowe, R., Dodig-Crnkovic, G., & Almér, A. (2017). Predictive regulation in affective and adaptive behaviour: An allostatic-cybernetics perspective. In *Advanced Research on Biologically Inspired Cognitive Architectures* (pp. 149-176). IGI Global.
Luzardo, A. (2018). *The Rescorla-Wagner Drift-Diffusion Model* (Doctoral dissertation, City, University of London).
Mackintosh N.J. (1971) An analysis of overshadowing and blocking. *Q J Exp Psychol.* 23:118–125.
Maki, W. S., & Abunawass, A. M. (1991). A connectionist approach to conditional discriminations: Learning, short-term memory, and attention. *Neural network models of conditioning and action*, 241-278.
Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological bulletin*, 117(3), 363.
Montebelli, A., Lowe, R., & Ziemke, T. (2008). The cognitive body: from dynamic modulation to anticipation. In *Workshop on Anticipatory Behavior in Adaptive Learning Systems* (pp. 132-151). Springer, Berlin, Heidelberg.
Montebelli, A., Lowe, R., Ieropoulos, I., Melhuish, C., Greenman, J., & Ziemke, T. (2010). Microbial Fuel Cell Driven Behavioral Dynamics in Robot Simulations. In *ALIFE* (pp. 749-756).
Montebelli, A., Lowe, R., & Ziemke, T. (2013). Toward Metabolic Robotics: Insights from Modeling Embodied Cognition in a Biomechatronic Symbiont. *Artificial life*, 19(3_4), 299-315.
Morén, J. (2002). Emotion and Learning - A Computational Model of the Amygdala. *Lund University Cognitive Studies*, 93.

Morrison, I., Perini, I., & Dunham, J. (2013). Facets and mechanisms of adaptive pain behavior: predictive regulation and action. *Frontiers in human neuroscience*, 7, 755.
Mowrer, O. H. (1947). On the dual nature of learning: A reinterpretation of ‘‘conditioning’’ and ‘‘problem-solving’’. *Harvard Educational Review*, 17, 102–148.
Navarro-Guerrero, N., Lowe, R. J., & Wermter, S. (2017a). Improving robot motor learning with negatively valenced reinforcement signals. *Frontiers in neurorobotics*, 11, 10.
Navarro-Guerrero, N., Lowe, R. J., & Wermter, S. (2017b). The effects on adaptive behaviour of negatively valenced signals in reinforcement learning. In 2017 *Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (pp. 148-155). IEEE.
Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53(3), 139-154.
Oberauer, K. (2009). Design for a working memory. *Psychology of learning and motivation*, 51, 45-100.
O’Reilly, R. C., Wyatte, D. R., & Rohrlich, J. (2017). Deep predictive learning: a comprehensive model of three visual streams. *arXiv preprint arXiv*:1709.04654.
Overmier, J. B., & Lawry, J.A. (1979). Pavlovian conditioning and the mediation of behavior. The *Psychology of Learning and Motivation*, 13, 1–55.
Pearce, J. M. (2013). *Animal learning and cognition: an introduction*. 3rd edition. Psychology Press.
Peterson, G. B., & Trapold, M. A. (1980). Effects of altering outcome expectancies on pigeons’ delayed conditional discrimination performance. *Learning and Motivation*, 11, 267–288.
Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioural control. *Progress in neurobiology*, 134, 17-35.
Pfeifer, R., & Scheier, C. (2001). *Understanding intelligence*. MIT press.
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.
Phillips, S. (2017). A general (category theory) principle for general intelligence: duality (adjointness). In International Conference on Artificial General Intelligence (pp. 57-66). Springer, Cham.
Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1), 79.
Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological review*, 118(1), 42.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non- reinforcement. In A. H. Black, & W. F. Prokasy (eds), *Classical Conditioning II: Current Research and Theory*. New York: Appleton-Century- Crofts.
Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. MIT press.
Rolls, E. T. (1999) *The brain and emotion*. Oxford University Press, Oxford
Rolls, E. T. (2013)What are emotional states, and why do we have them? *Emot. Rev.* 5(3):241–247
Rolls, E. T. (2016). *Cerebral cortex: principles of operation*. Oxford University Press.
Rolls, E. T. (2018). *The Brain, Emotion, and Depression*. Oxford University Press.
Rumelhart, D. E. (1990). Brain style computation: Learning and generalization. In *An introduction to neural and electronic networks* (pp. 405-420). Academic Press Professional, Inc..
Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). A general framework for parallel distributed processing. *Parallel distributed processing: Explorations in the microstructure of cognition*, 1, 45-76.
Saxe, A. M., McClelland, J. L., & Ganguli, S. (2018). A mathematical theory of semantic development in deep neural networks. *arXiv preprint* arXiv:1810.10531.
Schmajuk, N. (2010). *Mechanisms in classical conditioning: A computational approach*. Cambridge University Press.
Schröder, M. (2001). Emotional speech synthesis: A review. In *Seventh European Conference on Speech Communication and Technology*.
Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neurosci. Biobehav. Rev.* 32, 265–278.
Seger, C. A. (2009). “The involvement of corticostriatal loops in learning across tasks, species, and methodologies,” in The Basal Ganglia IX, eds H. J. Groenewegen, P. Voorn, H. W. Berendse, A. B. Mulder, and A. R. Cools (New York: Springer-Verlag), 25–39.
Seger, C. A., & Spiering, B. J. (2011). A critical review of habit learning and the basal ganglia. *Frontiers in systems neuroscience*, 5.
Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial intelligence*, 46(1-2), 159-216.
Sun, R. (2015). Interpreting psychological notions: A dual-process computational theory. *Journal of Applied Research in Memory and Cognition*, 4(3), 191-196.
Sutton, R.S., Barto, A.G., (1998). *Reinforcement Learning: An introduction*. 1st edition. MIT Press.
Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. 2nd edition. MIT press.
Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably different? *Learning and Motivation*, 1, 129–140.

Trapold, M. A., & Overmier, J. B. (1972). The second learning process in instrumental learning. In *Classical Conditioning II: Current Research and Theory* (pp. 427–452). New York: Appleton-Century-Crofts.
Urcuioli, P. (1990). Some relationships between outcome expectancies and sample stimuli in pigeons' delayed matching. *Animal Learning and Behavior*, 18(3), 302–314.
Urcuioli, P. (2005). Behavioral and associative effects of differential outcomes in discriminating learning. *Learning and Behavior*, 33(1), 1–21.
Van der Velde, F., & De Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences*, 29(1), 37-70.
van der Velde, F., & de Kamps, M. (2015). The necessity of connection structures in neural models of variable binding. *Cognitive neurodynamics*, 9(4), 359-370.
Wilson, W. H., Marcus, N., & Halford, G. S. (2001). Access to relational knowledge: A comparison of two models. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 23, No. 23).
Wynne, C. D. L. (1995). Reinforcement accounts for transitive inference performance. *Animal Learning & Behavior*, 23(2), 207-217.

Appendix A: Rescorla-Wagner Model

$$p = \sum_i s_i v_i \tag{1}$$

where p is the prediction of reward, s_i is the stimulus indexed by i, v_i is the corresponding weight (or valuation) of the stimulus.

$$\Delta v_i = \alpha s_i [\lambda - p] \tag{2}$$

where λ is the reinforcement value set in (0,1), α is a learning constant in [0,1].

Appendix B: Balkenius-Morén Model

The Balkenius & Morén (2001) model uses equation (1) so that to calculate reward magnitude. The weights update rule for valuating reward magnitude is the same as for (2) except that it computes only non-negative values (Morén 2002).

$$\Delta v_i = \alpha s_i [\lambda - p_m]^+ \tag{3}$$

$$p_o = \sum_i s_i \; w_i \tag{4}$$

where p_o is the prediction of reward omission.

$$\Delta w_i = \beta s_i (-[\lambda - p_m]^+ - p_o) \tag{5}$$

where $\beta > \alpha$, is in [0,1].

$$E = p_m - p_o \tag{6}$$

where E is the output of the network (motivational state).

Appendix C: Lowe et al. Model

The Lowe, Almér et al. (2017) model is conceived as a temporal difference instrumental learning model. Here is presented only the pavlovian component and in non-temporally discounted form (i.e. where $\gamma = 0$) thereby collapsing the model to the Balkenius-Morén model above (eqs 3-6). However, the output of the model differs in providing for *optimistic* (eq. 6) and *pessimistic* (eq. 7) affective valuations of the stimuli. Note, $pe_m = \lambda - p_m$ as depicted in figure 5.

$$E_{pes} = -E + p_o \tag{7}$$