Authors' Reply

Relational Complexity Metric Is Effective When Assessments Are Based on Actual Cognitive Processes

Abstract. The core issue is how relational complexity should be assessed. The target article proposes that assessments must be based on actual cognitive processes used in performing each step of a task. Complexity comparisons are important for the orderly interpretation of research findings. The links between relational complexity theory and several other formulations, as well as its implications for neural functioning, connectionist models, the roles of knowledge, and individual and developmental differences, are considered.

The target article proposed that relational complexity, defined as the number of arguments of a relation, provides the best measure of complexity in higher cognitive processes. The relational complexity metric *per se* does not seem to have been challenged in most of the commentaries, but they do raise many other issues which we will consider in turn. First, to avoid misconceptions, we would like to briefly summarise our position, placing particular emphasis on what we are *not* attempting to explain.

It is important to emphasise that performance on cognitive tasks depends on a number of factors, including familiarity with the task, domain knowledge, availability of appropriate strategies, storage and retrieval of relevant information, and so on. We fully accept the importance of these factors, and a lot of our work in other contexts has been devoted to them. See, for example, work on learning (Halford, 1995), induction (Halford, Bain, Maybery, & Andrews, in press), and strategy development (Halford et al., 1995). However in the target article we are concerned with observed effects of *complexity* on higher cognition. There are many cognitive phenomena that require a complexity metric for their explanation. We propose that these phenomena can be accounted for by the complexity of relations that have to be processed in a single decision.

We also wish to re-emphasise that the capacity limitations we defined in the target article apply where a cognitive process operates on a representation of relations in the task. This representation should have the properties of relational knowledge defined in 2.2. Associative processes, and modular processes that are specialised for processing a limited range of inputs, do not have these properties (Anderson, 1992; Halford, 1996a,b; Leslie & Keeble, 1987; Phillips, Halford, & Wilson, 1995) are therefore are not expected to show the same complexity effects.

How complexity should be analysed

Sweller confirms that he has found in empirical work that the number of interacting variables processed by a performer is a good measure of task demand. Relational complexity subsumes number of interacting variables because, as noted in 2.0 and 2.3.5, each argument of a relation effectively functions as a variable, and an n-ary relation is a set of points in an n-dimensional space. Sweller has also found that analysis of complexity depends on what dimensions are chunked, and this depends on expertise.

We agree, and in 2.1 we make the more general point that complexity depends on the cognitive processes that are being employed. The processes employed by a particular person in performing a task undoubtedly depend on expertise, age, problem presentation, goals etc. All these factors affect the way a task is chunked and/or segmented. We said in 6.0 that analysis of relational complexity depends on having a process model of the way the task is performed. The process model will need to be verified independently to make predictions based on complexity testable. Some of the commentaries indicate that the implications of these requirements have not been fully recognised.

Formal similarity of tasks does not guarantee that processes are similar, so Sweller's statement that isomorphic tasks have the same relational complexity is incorrect. This can be illustrated by contrasting two isomorphic tasks, one based on days and the other based on numbers: *Suppose five days after the day before yesterday is Friday. What day of the week is tomorrow?* and $x + 5 - 2 - 1 = 6$. If we number the days of the week consecutively, so Sunday is 1 and Friday is 6, these problems are isomorphs, but they differ markedly in difficulty. At first sight this might seem to imply that relational complexity cannot account for the difficulty, but in fact the tasks differ substantially in the relations that are processed. The arithmetic task is easily segmented; we can perform $5 - 2 = 3$, which is a ternary relation (binary operation). Similarly for $2 - 1 = 3$. Now we have $x + 2 = 6$. Subtracting 2 from both sides involves simply cancelling the 2 and performing $6 - 2 = 4$, another ternary relation, so $x = 4$. Mapping this into the word problem, tomorrow is Wednesday. The arithmetic problem can be solved by a series of steps that require, at most, ternary relations. Since we have already learned the steps, and know how to apply to them to equations of this kind, the planning process imposes a negligible load. By contrast, success in the word problem depends on devising a strategy that segments the problem into a series of manageable steps. However this is difficult to do, in part because planning a suitable strategy depends on first representing the structure of the task as a whole. It is the planning process that imposes high processing loads in this case, not performance of the individual steps. For example, to translate the days problem into the isomorphic arithmetic problem requires recognition of the correspondence between the two structures. The load imposed by planning some other strategies might not be as high, but it is still substantial.

The crucial point therefore is that relational complexity analysis has a good chance of accounting for task difficulty when it is applied to the processes used. Sweller is right that determining these processes can involve a lot of work, but this is not a burden imposed specifically by our model. It is inherent in any genuine attempt to analyse cognitive complexity. On the positive side, it is a task for which cognitive psychology and cognitive science are now well equipped. There is a rich array of techniques for theoretical modeling and empirical analysis. Perhaps most important of all, clear and coherent accounts of cognitive processes are a major benefit in themselves, and it might not be unreasonable to suggest that they should be the ultimate goal of our research. Because complexity is a major factor effecting performance, albeit by no means the only factor, having a precise way of defining complexity can be of considerable benefit in our efforts to understand cognitive processes.

Chalmers and McGonigle (C&M) attribute a lot of ideas to us which form no part of our theory. One example is that, in common with Wright and Goswami, they suggest that our conception of relational complexity is not empirically founded, which is simply untrue. In fact we have examined an extensive data base in arriving at our conclusions (see for example the references cited by Halford, 1982; 1993), and this data base includes the transitivity of choice paradigm on which C&M's position is based. C&M also attribute to us the assumption that people adopt representations that maximise the amount of information processed in parallel. They apparently missed our statement in 6.1.4: "We assume that participants normally segment sentences into constituents which are processed serially as far as possible. In our modeling, in this and other contexts, we have found it a fruitful assumption that participants tend to minimise processing demand, implying that they never process more information in parallel than necessary." Perhaps because they missed this point, they mistakenly suggest that we regard multiply centre-embedded sentences as the norm. On the contrary, we stated that the interest in these sentences is that they constrain participants to process more of the sentence in parallel. This enables processing limitations to be observed. This logic has also been used by Henderson (1994) and Just and Carpenter (1992) whom we cited in 6.1.4.

Why tasks impose loads. While it is obvious that people will tend to minimise processing loads, high loads are imposed by the nature of some tasks. The reason is that where sources of information interact, they must be considered jointly. This principle is routinely applied when we analyse effective task complexity. It can be illustrated using analysis of variance: interacting factors cannot be interpreted singly, because the effect of any factor is modified by the others. All factors that enter into a particular interaction have to be interpreted jointly. This principle provides an objective criterion for complexity, and has been the basis for our analyses. Premise integration in transitive inference (discussed in 6.1.1) illustrates the point. Premises such as "John is smarter than Stan, Tom is smarter than John" interact in that the interpretation of one premise is influenced by the other. Therefore each premise contains some ambiguity when processed alone, but the ambiguity is removed when they are considered jointly. Notice first that "John is smarter than Stan" assigns John to first or second position, while "Tom is smarter than John" assigns "John" to either second or third position. However the two premises, considered jointly, assign John uniquely to second position. Tom and Stan are then unambiguously assigned to first and third positions respectively. Because neither premise can be fully interpreted in isolation, there is a limit to segmentation. Relational complexity theory predicts that the need to consider premises jointly will impose a processing load, which has been verified empirically (Maybery, Bain, & Halford, 1986).

The most immediate implication of this is that tasks cannot always be segmented into steps that are performed serially. Ability to segment depends on a number of factors, including task structure, conditions of performance, and expertise. The analysis above would be consistent with the expertise of most people in an industrial society aged between roughly five years and late adulthood. It would apply where the premises have to be integrated mentally after a single presentation, as occurs in many experiments with human participants. However there are other conditions which make segmentation easier. For example, we could present premises one at a time, and let children order blocks. Thus given a>b, the child can place *ab*, then when given b>c they can concatenate *c* yielding *abc*, and so on. By this procedure, they never have to consider more than one relation in any decision (Halford, 1984). Consequently the processing load is less, and performance of children is much better (Andrews & Halford, in press). We routinely use tasks that have this property as control procedures in our experimental studies of transitivity. Thus segmentation can be influenced by task conditions, but is also constrained by structure that is inherent in the task.

C&M contend that premise integration is unnecessary in transitive inference tasks, and that processing is serial. To support their claim they rely on a paradigm derived from the work of Bryant and Trabasso (1971). Participants are systematically trained on pairwise comparisons (e.g. A versus B, B versus C etc.) then tested on untrained pairs (e.g. A versus C). Transitivity of choice has been demonstrated in species from pigeons to humans. Having been trained, usually over many trials, on A+B- (i.e. to select A rather than B) and B+C- etc. they show a transitive bias when presented with non-adjacent pairs, choosing B in preference to D etc. However the transitive bias is reduced when triads such as BCD are presented. This would not be expected if participants constructed an ordered array (A,B,C,D,E), because they would have stored the ordinal position of all elements. Harris & McGonigle (1994) claim that performance of primates is consistent with production rules that select items near one end of the series and avoid items at the other end. Children perform this paradigm in essentially the same way (Chalmers & McGonigle, 1984). Wynne (1995) has shown that pigeon data are well explained by associative learning models that do not entail processing relations at all. These models do not entail representation of the relations in the task (e.g. they do not entail representation of the relation between A and B, B and C etc.), there is nothing equivalent to a transitivity principle, no ordered array is constructed, and performance depends solely on relative strengths of item preferences, acquired by learning. Therefore this research shows that at least some versions of the *transitivity of choice* task can be performed without a cognitive representation of relations between task elements; that is, without processing structure. In fact it can be performed by much lower level cognitive processes.

The processes used in this paradigm reflect task demands. The paradigm shows that if the task can be performed using serial processes that impose low demands on capacity, then most participants will perform it that way. C&M then ask why are these low demand serial processes not used all the time? The answer is that they suffice for only a very restricted range of tasks. The transitivity of choice task cannot be adopted as a paradigm for all human cognition. A lot of cognitive tasks require more information to be processed in parallel, and more elaborate cognitive representations. Even transitive inference requires this when premises have to be integrated mentally after a single presentation, as we saw above. Furthermore there is strong evidence that ordered arrays are constructed under some conditions, even by young children. Andrews & Halford (in press) presented young children with premises coded as pairs of coloured blocks, with A above B, B above C, . . ,D above E. They were asked to infer the relative position of B and D, then to place C. If children had constructed an ordered array, it would be easy to insert C between B and D. The proportion of correct placements of C, given that BD had been placed correctly, increased from 42 percent at age 4 to 95 percent at age 6. This is added to several other lines of evidence suggesting that integration of premises to form an ordered array not only occurs, but is a cognitive achievement that is related to age.

C&M's contention that mental integration of information to form reasonably complex relational representations is an unnecessary burden simply reflects their reliance on tasks that make low demands. The transitive choice paradigm can be accounted for either by associative learning models, or models that postulate preference for certain stimuli. It does not require relational processing that would conform to the principles of relational knowledge that we defined in 2.2. Furthermore McGonigle and Jones (1978) found that discriminating the middle item was more difficult than transitivity of choice. "Middle" is a ternary relation, though it can be chunked to binary in some circumstances as we will consider below. Either way it is more complex than the associative processes entailed in transitivity of choice, and the finding that it is harder is clearly consistent with our position, but C&M do not acknowledge this. Criteria that distinguish between associative and relational knowledge have been defined by Halford, Bain, Maybery and Andrews (in press). We think it is unlikely that the very restricted processes advocated by C&M would meet these criteria.

C&M are incorrect when they suggest that 8 item seriation requires an oct-ternary relation. To make this error they must have missed a major component of our argument (e.g. in 2.1), which is that processing load is determined by the complexity of relations processed in a given step, not simply by the total amount of information contained in the task. We regard it as fundamental that processing load depends on the amount of information processed in any given step, not on the total amount of information in the task. Our reduction technique for defining relational complexity of a concept in 3.4.3 (which they cite to support their claim) means that forming an 8 item series requires ternary (not oct-ternary) relations to be processed in a single decision. The point can be made more simply by considering seriation of 5 items. Suppose premises A>B, A>C, A>D, A>E, B>C, B>D, B>E, C>E, D>E, are presented in random order (noting that premises do not necessarily have to be restricted to adjacent items, nor do they have to be presented in any particular order). Given (say) premise A>C, the string AC can be formed. If A>B is presented, possible strings are ABC and ACB. We need premise B>C to decide that the correct order is ABC; that is A>B and B>C jointly determine that the order is ABC. The task can be performed by a series of steps, the most complex of which entails dealing with ordered 3-tuples. That is, it only requires processing ternary relations.

It might be asked which is the most appropriate, or most valid test of transitive inference? Our contention is that any of the tests is capable of providing information about capabilities of humans and other animals, but the less demanding tasks do not demonstrate complete mastery of the concept. Given a low demand task, such as transitivity of choice, participants utilise low-level cognitive strategies. This simply says that performers are rational in the sense of Anderson (1990) because they use the least demanding strategy that suffices for the task at hand. However we would be unwilling to concede that participants who could perform no other strategy had mastered transitivity. This contention might be unwelcome to those who want to attribute transitivity to very young children, or nonhuman animals, who show no evidence of using the more demanding strategies. It would not be appropriate to dismiss performances based on tendencies to select specific stimuli, as in the associative learning model of Wynne (1995) or the production rules of Harris and McGonigle (1994). We are glad to acknowledge that the careful and ingenious research conducted in that paradigm has yielded valuable knowledge of the cognitive processes utilised there by a wide range of participants. There is no reason however to accept those models as comprehensive accounts of human reasoning. Even relatively simple reasoning tasks can easily be demonstrated to require more complex processes, as our example with transitivity above illustrates. It also seems reasonable to suggest that ability to mentally integrate premises, forming mental representations of the relations in a task, is a landmark cognitive achievement. In the case of transitive inference, it is this which we contend requires processing ternary relations.

The foregoing discussion should demonstrate the error in Wright's commentary when he contends that the relational complexity metric cannot handle more than one version of a task. The theory does *not* imply that there is one right version of a task. There are many ways of assessing even simple cognitive functions such as transitive inference, and probably at least as great a variety of cognitive processes that can be employed. Complexity is assessed on the basis of the processes employed. This is not a problem because techniques for determining strategies have appeared in abundance in the last few decades. Let us note in passing that one reason why transitive inference is useful for complexity analyses is that we have well validated process models. Indeed, transitive inference research is arguably a great, though unrecognised, success story in cognitive psychology. Furthermore complexity of processes employed is constrained by the fact that sources of information that interact must be interpreted jointly.

This point was evidently missed by Pascual-Leone (P-L) when he contended, incorrectly, that " . . since the theory has no explicit rules constraining their occurrence, chunking and segmentation turn into theoretical loopholes for explaining away empirical anomalies." In fact general principles of chunking were given in 3.4.1, the reduction technique outlined in 3.4.3 provides an objective way to determine the effective relational complexity of tasks and principles for complexity analysis were given in 6.0. However the principle that sources of information cannot be processed serially when they interact is the core of our method. It is first necessary, as we have noted all along, to have a clear model of how a task is performed. Given this, we have consistently found it useful to analyse the number of interacting variables in a given decision. Using this technique we have predicted complexity effects before they were observed in the balance scale, hypothesis testing and concept of mind (Halford, 1993). We have also been able to analyse tasks as diverse as classification, tower of Hanoi, knights and knaves (Rips, 1989), and Raven's matrices. We have also found very good correspondences across domains.

Pascual-Leone's position does not appear to have any comparable way of removing subjectivity from complexity analyses. We consider that explicit computational models are a more objective way to determine the nature and complexity of processes used in a task. For example, the demonstration that a particular task can be performed by a typical three-layered net is a good indication that it is basically associative, and does not require representation of explicit relations. Thus the demonstration by Quinn and Johnson (1997) that prototype formation can be achieved by a three-layered net suggests that it does not require relational processing.

Pascual-Leone argues that complexity estimates for proportion, a/b=c/d must include the ratios a/b and c/d, and also suggests that we do not allow for these ratios. We contend that both arguments can be refuted. Our position is that whether the ratios must be included depends on whether they are processed by the performer. Despite his suggestion to the contrary, we do take this into account. Processing of proportion can be simplified by first computing a/b = x and c/d = y, then comparing x and y. Thus if asked whether 4/8=12/24 one could compute 4/8 = .5 and 12/24 = .5 and conclude the expression is true. This segments proportion into two ternary relations performed in succession. Notice however that to plan this procedure, or to understand why it is valid, the structure of proportion must be represented. A proportion is truly a quaternary relation, and is defined by links between four variables. It is defined in four-dimensional space, not six-dimensional space, as contended by Pascual-Leone. Notice Pascual-Leone's position implies that from age 15 humans can process seven dimensions in parallel. Pascual-Leone does not seem to offer any evidence that this is so, nor does he consider its theoretical consequences. What kind of cognitive processes could be performed by a system that was processing seven dimensions in parallel? This seems to be quite unexplored territory.

Pascual-Leone misrepresents our position in at least two other ways. He incorrectly puts us in opposition to the view that effective complexity is defined relative to an adaptive system, and can be represented by a concise description of the regularities in the task. Our position is quite closely related to this view, and also to contributions such as that of Leeuwenberg (1969) who was early in his recognition of the potential for defining psychological complexity in this way. Leeuwenberg's metric was explicitly applied in some of our earlier work (Halford & MacDonald, 1977). Because we define complexity by the space in which a cognitive process is operating, our position seems quite consistent with the one which Pascual-Leone wants to put in opposition to it. His claim that in 3.2 we assigned *played(John,cricket,oval,Sunday)* to an effective complexity of four is also incorrect. That relational instance is simply a list, and its relational complexity as defined by the reduction technique in 3.4.3 is indeterminate.

Links to other formulations

We agree with Anderson, Lebiere, Lovatt and Reder (ALL&R) that there are points of contact between relational complexity theory and ACT-R. In particular, we agree that the name for a production could be used to represent the symbol (name) of a relation, and the slots of a production can correspond to the arguments of a relation. A means of binding the relation symbol and arguments would still be required however, and it is not clear how that would be done in a production. We also agree that ACT-RN uses separate memories in a way that is analogous to our use of different sets of units in a tensor product to represent arguments of a relation. There is partial correspondence between the roles for activation in the two models. In ACT-R activation is the main cause of capacity limitations. In relational complexity theory the demand for activation increases with the rank of a tensor product (see 5.2.1.2). However it is not the only factor that limits complexity in our model, because number of arguments corresponds to the rank of a tensor which constrains the number of connections between units.

While the points of contact offer interesting potential, considerable work is still needed to achieve a genuine integration, and neither relational complexity nor any comparable metric is incorporated into ACT-R as it stands. One of the clearest differences is that in ACT-R complexity is assessed by the number of symbols (Anderson, Reder, & Lebiere, 1996), whereas in our model it is based on the dimensionality of the decision space. The difference can be simply illustrated by contrasting the following sets of problems, in both of which participants must solve for x:

Set 1: x + 4 = 7; x + 3 = 7; x + 4 = 8; x + 3 = 8.

Set 2: x + 1 = 7; x + 1 = 8; x + 1 = 5; x + 1 = 6.

A count of the number of symbols, following Anderson et al. (1996) yields 5 for both sets. Yet the relational complexity of the first set is 3 (because there are three dimensions of variation) but for the second set relational complexity is 2 (2 dimensions of variation). It seems intuitively likely that set 2 would also be easier. As we noted above, there are no simple ways to arrive at a valid assessment of the dimensionality of the cognitive processes performed in a task, and counting the number of symbols is not adequate.

Incorporation of the relational complexity metric into a model with the power and generality of ACT-R would be a very exciting development. However considerable work is required to accomplish this. One of many benefits would be that ability to estimate the amount of information that humans processed in parallel should provide a useful constraint on ACT-R, just as it has on models of analogy (Halford et al., 1995; Hummel & Holyoak, 1997).

MacLeod addresses the problem of the optimum allocation of resources to a set of activities (or tasks), given that each activity has costs and benefits. We should note that this is not the core question addressed by the target article, which is concerned with assessment of cognitive complexity, even in single tasks. MacLeod's approach may be productive as a theory of dual task performance, perhaps by extension of the work of Navon and Gopher (1979) who also applied economics theory to the resource problem in psychology. Maclead's model of resource allocation, like the commentaries by Anderson and Cowan, provides an interesting source of hypotheses for the cause of complexity at a lower level of analysis. We also note that our position is quite unequivocal on several of the points raised by MacLeod. We have proposed that the complexity of a task depends on the arity of a relation processed in parallel. This is related to dimensionality because, as we pointed out in 2.3.5, an n-ary relation is a set of points in n-dimensional space and can represent an interaction between *n* variables. Therefore complexity is related to dimensionality, not to the length of the input string. Furthermore complexity effects occur because computational cost is a function of the number of dimensions processed in parallel. We agree that a high dimensional problem can be mapped into less dimensions. This is the essence of chunking, discussed in 3.4.1, though it should be recalled that we specified limits to this process.

We do not agree that complexity is a direct result of resource limitations, though resources available affect our ability to deal with complexity. This discrepancy seems to arise because the problem addressed by MacLeod, optimum allocation of resources to activities, is not the problem addressed by us. As discussed in an earlier section of this reply, complexity effects arise because when dimensions of a task interact they need to be be processed jointly and this increases processing load. Therefore processing load results from constraints that are inherent in the process being performed, not from resource limitations *per se*.

In relation to NP-completeness, NP stands for non-deterministic polynomial time problem. Similarly, P stands for the class of problems solvable by a polynomial-time algorithm. Problems in NP have a structure that any hypothesis can be checked polynomial time, but the total number of hypotheses is exponential in the size of the problem.MacLeod is of course right in pointing out that nobody has yet proved that P and NP are distinct. It is possible that there are as yet unknown algorithms to solve NP-complete problems in polynomial time, just as more efficient algorithms have been discovered from time to time for various tasks such as sorting. However, at present there is no known efficient (polynomial time) algorithm for such problems, so the time (or space) required to compute a solution is exponential in the size of the problem.

MacLeod indicates he would have preferred "a more detailed presentation of one of the mathematical models, along with some theorems on the algorithmic complexity of the proposed mechanisms". We will be gratified if our formulation opens up opportunities for further theoretical development, but the aim of the target article was to find a complexity metric that would be broadly applicable to psychological tasks, taking account of the real problems of conducting psychological research. In 5.2.1.1 and 5.2.1.2 we offered analyses of the algorithmic complexities of the basic processes of tensor product network operations. The tasks that we currently are most interested in are those that can be completed in a few such operations, and the processes that we are most interested are the simplest ones for performing such tasks.

Frontal lobe impairment

Our suggestion in 6.5, based on the review by Robin and Holyoak (1994), that processing explicit relations might be a major function of the frontal lobes, has now received some empirical support. The commentary by Waltz, Knowlton and Holyoak reviews several empirical studies, including their own ongoing research, indicating that patients with damage to the dorsolateral prefrontal cortex show selective impairment in processing complex relations. Patients with damage to the anterior temporal lobes showed no such impairment. The possibility that relational complexity offers a new way of defining the functions of a major region of the frontal lobes is clearly quite exciting. The techniques developed in our laboratory, discussed in the previous section, for manipulating complexity while holding other factors constant, should have considerable utility in testing this hypothesis. Relational complexity measures can also be used to assess deterioration due to other causes, including ageing.

Cause or effect: speed versus capacity

The complex issue of whether speed or capacity is the causal factor influencing performance in complex tasks has been well elucidated by Cowan. He presents some interesting evidence that tends to favour capacity as cause. On the other hand there is very extensive developmental work by Kail (1991) indicating a general processing speed factor that increases with age. We have reviewed this work elsewhere (Halford, 1993, pp. 119-122). We think both speed and capacity hypotheses will remain viable for the foreseeable future and we can envisage mechanisms that are appropriate to either.

Assuming that a cognitive phenomenon can be modelled by a neural net with a settling phase, it is reasonable to expect that a network that is closer to its capacity limit will take longer to settle. For example, if more items are superimposed on a fixed set of units, the items become less discriminable. Discriminability can be increased by increasing the number of units, which is one measure of capacity in a net. A net with higher capacity will therefore have a shorter decision time. Some distributed memory theories appear to model this effect fairly directly (McNicol & Stewart, 1980; Murdock, 1983). In these models memory items are represented as vectors that are superimposed on a set of units. Recognition time increases linearly with number of items, thereby predicting the set size effect in memory scanning (Sternberg, 1975). The implication is that more items impose higher demands on the available capacity, which increases decision time.

While Cowan's example is orthogonal to the relational complexity metric (a list treated as a unary relation is still unary regardless of its length), Cowan's data may be interpretable in terms of this model. If we assume that inter-word pauses in Cowan's experiments represent decision time, then pauses would increase with list length because longer lists would impose higher demand on available capacity. Pause time would also increase with lower capacity. If we assume participants with lower span have less capacity in some sense, then the longer pauses observed by Cowan with lower span participants can be explained. Thus this model is essentially consistent with Cowan's data. It also suggests a testable hypothesis. This is that participants with a lower relational complexity limit, as defined by the measures developed by Andrews and Halford (in preparation) should have longer inter-word pauses.

In case of the symbol-argument tensor modeling, transmission speed could affect capacity in the following way: Suppose role units R1 and R2, and binding units for computing the appropriate relations. Post-synaptic activation is only possible when all pre-synaptic activations from the various role units occur with the same pre-synaptic activation window (w). Suppose the distances from R1 and R2 to the binding units are D1 and D2, respectively; and transmission speed is s. A signal sent from unit R1 at time t will arrive at the binding unit at time t+D1/s. A signal sent from unit R2 at time t+d will arrive at the binding units at time t+d+D2/s, assuming a difference *d* between signal transmission from units R1 and R2. Therefore, the binding unit will be active when $d+(D1-D2)/s < w$. Clearly, the faster the transmission speed the more likely the two signals will arrive within the activation window of the binding unit, permitting the representation of the binary relation, or other higher arity relations.

In the case of processing capacity (number of units), higher ranked tensor units may simply not be connected to represent higher arity relations. From Thatcher (1994), note that while total number of units remains relatively constant through the first decade, connectivity does not, possibly suggesting the establishment of progressively higher order connections within the cortex. These two examples are meant as illustrations. Although we cannot decide the issue at present, relational complexity does not lead to a conceptual dead-end at finer levels of analysis. Rather, it suggests more detailed hypotheses to be tested by future experiments.

Complexity versus interference

Navon is concerned with the longstanding issue whether performance is a function of outcome conflict rather than resource limitations. We have pointed out elsewhere (Halford, 1993, chapter 3) that the easy-to-hard paradigm (Hunt & Lansman, 1982) avoids the interpretational difficulties associated with dual task deficit methodologies. This is because the hard version of the primary task is never performed concurrently with a dual task, so the variance it shares with the predictor, and which is taken as evidence of a resource limitation, cannot reflect outcome conflict. Berch and Foley acknowledge that the easy-to-hard paradigm (Hunt & Lansman, 1982) is an effective way to test hypotheses about capacity limitations, but point out that it is methodologically complex. We agree that this is so, but their excellent work (Foley, 1997; Foley & Berch, 1997) has demonstrated that it is effective for this purpose. The effort of using the easy-to-hard paradigm, though significant, is worthwhile.

In general, there have been many advances in methodology since Navon (1984) made his claim that resource limitations were difficult to distinguish from outcome conflict. These include experiments (e.g. Andrews & Halford, in press; Maybery et al., 1986) where both input and output are tightly controlled while the complexity of the intervening processes are manipulated. It seems very unlikely that the effects of relational complexity, some of which are very large, that have been demonstrated in these studies could be attributed to outcome conflict. Outcome conflict only applies at best to the dual task interference paradigm. However the target article goes well beyond this paradigm and is primarily concerned with complexity-based limitations. Even if outcome conflict were a viable explanation in some studies, the number and diversity of complexity effects is so great that it is very implausible that they could all be explained by outcome conflict.

The question of relational tasks performed concurrently, raised by Navon, was addressed in 3.3. Neural net models of the type we used to implement our theory are capable in principle of dealing with similarity effects. In particular, the shortcomings that Navon sees as applying to neural network models with respect to resource conflict, are not relevant to complexity-based limitations. The complexity of the memory scanning task, raised by Navon, is addressed by Halford, Maybery and Bain (1988), referred to in 3.3. An unordered string of digits is not relational, and can be represented by a set of superimposed vectors, as noted in the previous section. An ordered string of items in short term memory is a binary task, assuming that the person's strategy is to store the list as position-item pairs; {(position1,item1), (position2,item2), . . }. There is no growth in relational complexity as the number of digits is increased, though there may be an increased load on working memory if the items are held in an active state. Criteria for this were discussed by Halford et al. (1988).

Connectionist models of relations

We agree with Plate that there is considerable scope for further connectionist proposals for relational processing. Note that as Plate's Holographic Reduced Representations (HRRs) are projections of tensor product nets (in the mathematical sense of 'projection') it is not surprising that systems based on these two approaches share some properties. We also agree that role-filler binding methods, suitably augmented with content-addressable memory, are sufficient to represent and access relational information. However, we do not agree that holographic reduced representations (HRRs) implement what we intended by the term conceptual chunking. Rather, chunking entails loss of information through dimension reduction, not noise reduction through compression. Further, in light of the double dissociation reported by Waltz, Knowlton and Holyoak, there is now a neurological reason to favour symbol-argument over role-filler binding methods: symbol-argument bindings used different ranked tensors (i.e., different unit types) for different arity relations, whereas role-filler methods use the same tensor units. The former seems more consistent with the existence of a region in the brain, the dorsolateral prefrontal cortex, that is specialised for processing high arity relations.

Plate argues that HRRs do, in fact, implement chunking. While HRRs satisfy our principles (1) and (3) in 3.4.1, they do not satisfy (2). Plate states, "every vector in a HRRs model is already a chunk and no further compression is necessary". In this case, there is no sense in which one can have both a chunked and an unchunked representation of the same concept. Thus, either satisfying principle (2) "no relations can be represented between items within a chunk" implies not being able to represent relations; or more likely, because HRRs represent relations, they cannot satisfy principle (2).

The difference is, that with an appropriate (learned) chunking strategy, chunks identify unique equivalence classes of concepts. In our $v = s/t$ example, the chunk "60" represents a class of relational instances: Div(60,1); Div(120,2), etc. A chunked version of the corresponding rank two tensor representing the relations may be the vector determined by actually calculating the velocity. In Plate's method, by contrast, these equivalence classes occur by accidental collisions between vectors, at the mercy of the statistics of vector generation and manipulation. Of course, it would be possible to augment HRRs with a chunking mechanism, and represent both the chunked and unchunked representations as HRRs. However Plate's HRR model, as we understand it, does not currently do this.

Symbol-argument versus role-filler binding

Waltz, Knowlton and Holyoak (WK&H) provide further experimental support for our relational complexity metric, but remarked that the data are neutral on the issue of algorithms (e.g., whether relational information is computed by tensor product or synchronous activation). We are excited by these findings, but suggest there is also some room for interpretation at the algorithmic level of analysis.

The double dissociation between prefrontal cortex (integration of two binary relations) and the anterior temporal lobe (single binary relation) suggests different underlying neural architectures for relations of different arities. It has been noted elsewhere (Phillips & Halford, 1997) that the tensor symbol-argument binding method, but not the role-filler method (which includes Plate's Holographic Reduced Representations) permits such double dissociations. For the symbol-argument method, binary and ternary relations require distinct unit types: for example, rank two tensor units multiply two incoming sources of activation, compared to three sources for the rank three units. By contrast, role-filler methods, and for that matter synchronous activation, use the same type of unit regardless of arity. With role-filler methods, it may be the case that additional units must be recruited for higher arity relations. But, this would only demonstrate a dissociation, not a double dissociation, since first and second positions are shared by both relation types.

However, since we do not know exactly what part of the integration process prefrontal cortex is responsible for, this interpretation is at best suggestive. We agree with WK&H in that much more work at the neurological level is required before definitive claims can be made about neural mechanisms.

Relational complexity and chaotic attractors

Three commentaries, those by Heath and Hayes, Nikolic, and Borisyuk, Borisyuk and Kazanovich, have drawn attention to some possible correspondence between relational complexity and dynamic systems theory. Both formalisms are characterised by a common limitation, in that the range of dimensionalities is low, from one to four. The basis for this common limitation is well worth further exploration. At the present time the correspondence, while intriguing, is difficult to interpret. Nikolic finds the same limitation in motor processes as we find in higher cognitive processes. It may be, as Heath and Hayes suggest, that the limitation is a general property of neural processes. However at present it is difficult to determine whether these common limitations are coincidental, or represent a genuine underlying phenomenon. The potential importance of correspondence between relational complexity and dynamic systems theory is great enough however to warrant more extensive investigation.

An integrated treatment could be very productive for another reason, which is that the approaches have complementary advantages. Dynamic systems theory has elegant ways of predicting discontinuities, whereas relational complexity theory is more closely tied to observable phenomena. The levels of complexity defined in the target article have been identified through detailed analyses of tasks across a wide range of domains, and complexities have been manipulated precisely while holding other factors constant. If these levels of complexity are found to correspond to dimensionalities of chaotic attractors, the methodology of relational complexity theory would seem to open up a lot of opportunities for empirical testing of dynamic systems models of cognitive processes.

Knowledge and higher-order relations

A number of commentators, including Wright and Gentner and Ratterman (G&R) claim we give priority to capacity rather than knowledge as an explanatory factor. In fact we consider that they interact. See, for example, Halford (1993) p. 272, Postulate 1.0: “Cognitive development depends on the interaction of learning and induction processes with growth in the capacity to represent concepts.” To the extent that these factors interact, neither can be given priority. In statistical analysis an interaction cannot be decomposed to determine the relative importance of the constituent factors. Similarly, where knowledge and processing capacity interact in determining performance of a task, it is meaningless to ask which factor is more important.

In empirical research either factor may have a larger effect in a specific study, depending on how it is designed. A popular design is to use tasks that are well within the capacity of the participants, but which demand knowledge that has only been partially mastered by the sample selected. For example, studies of analogical reasoning in young children frequently use A:B::C:D analogies which require binary relations to be mapped, and capacity to process binary relations appears to develop at a median age of two years. However relations such as melting snow are used (Goswami & Brown, 1989) which children of 3-4 years are just beginning to understand. Such studies typically show that knowledge has a large effect, and they are inherently incapable of showing an effect of capacity. Alternatively, one could use tasks in which children have thoroughly mastered the prerequisite knowledge, either by selecting familiar materials, or by using extensive training (Halford, 1980). One can then manipulate the capacity requirements of the task from (say) binary to ternary with 4-6 year old children. In these circumstances, complexity produces a large effect on performance (Andrews & Halford, in press). Neither type of study shows any general priority for knowledge or capacity. The overall picture is that knowledge and capacity both have effects and, not surprisingly, investigators design their studies to reveal the factors they want to investigate.

G&R make the more specific claim that analogical reasoning in children depends on higher-order relations, which in turn depend on knowledge. Our response to this is that we have fully acknowledged the importance of higher-order relations. For example, in 6.1.3 we analysed the Tower of Hanoi in terms of the higher-order relation “prior”, for example; Prior(shift(2,C),shift(1,B)). In this context more complex tasks entail deeper structures, with more levels of embedding of higher-order relations. G&R postulate that depth of relational structures is more important than their dimensionality. However we showed that these tasks also entail higher dimensionality. So some of the effects that G&R attribute to depth of structure can equally be attributed to dimensionality. This issue is amenable to test, at least in principle, because depth of structure can be separated from dimensionality in certain cases. Thus we can have a deep structure in which each level comprises a unary relation, for example; R3(R2(R1(A))). Depth is three because there are three levels of relations. Dimensionality depends, as usual, on the number of variables. If A is the only variable dimensionality is one. However, if any or all of R3,R2,R1 is a variable, dimensionality may be as high as four. In other cases depth may be 1 and dimensionality 4, for example; R(a,b,c,d). The implication is that higher-order relations are important, but dimensionality is the more general criterion of complexity, and can be applied to structures of any depth.

Individual differences

Sweller suggests that the theory is less useful for assessing individual differences in processing capacity. In fact we are developing a test for complexity of relations that can be processed in parallel, and results so far are very encouraging. Andrews (1996) tested children aged 4-8 years in transitivity, hierarchical classification, cardinality, comprehension of relative clause sentences, hypothesis testing and class inclusion. Relational complexity was manipulated within each domain. All tasks loaded on a single factor and factor scores were correlated with age ($r = .80$), fluid intelligence ($r = .79$) and working memory ($r = .66$). This result was replicated (Andrews & Halford, in preparation) with a slightly different set of domains in 1997.

Sweller also contends that age differences that we attribute to capacity are indistinguishable from differences due to expertise. However proponents of this, rather common, view do not appear to have considered the type of experimental design we have used. First, we have used procedures and materials that are highly familiar to the children, and in some studies we have trained to asymptote on all components of the task (Halford, 1980; Halford & Leitch, 1989; Halford & Wilson, 1980). The typical result of these studies is that the younger children completely master all aspects of the procedure. Complexity is manipulated with procedure and materials tightly controlled. The result has invariably been that complexity has had a large effect on performance, especially with younger children. It might now be argued that there might have been some undetectable residual difference in expertise. It seems common to argue that this knowledge explanation must be the right one, even though the nature of the expertise differences is unspecified. It should be obvious that such an argument makes the knowledge explanation untestable. It also fails to take account of evidence, using the easy-to-hard paradigm, that some of the tasks in question are in fact capacity limited (Foley & Berch, 1997; Halford & Leitch, 1989; Halford, Maybery, & Bain, 1986).

Ultimately the strongest argument against the view that “knowledge explains all cognitive development, therefore there is no role for capacity” is that complexity has very real, and very powerful, effects on performance of young children. If these effects are entirely attributable to lack of knowledge, why can they not be removed by adequate experience or training? In the tasks where we have identified complexity effects, training has an effect only at or above the age at which capacity becomes sufficient for task demands (Halford, 1980; Halford & Leitch, 1989; Halford & Wilson, 1980). At some point it is clearly incumbent on proponents of the *knowledge only* view to demonstrate that knowledge does account for these effects.

Developmental issues

Many of the issues considered in earlier sections have important developmental implications, but in this section we will deal with those specific developmental issues not considered earlier.

Comparison with other developmental theories

Pascual-Leone disagrees with relational complexity theory but also attempts to identify it with his position by equating schemes with dimensions: “ . . the highest number of schemes -- task relevant dimensions of variation -- that subjects must consider simultaneously to solve the task.” However his position would need major transformations before it could be considered equivalent to relational complexity theory. Pascual-Leone’s contention that our model predicts a performance asymptote at age 11 is incorrect for several reasons. It fails to take account of the fact that capacity is a soft limit, and that 11 years is a median age, some children achieving capacity to process ternary relations later than this. Most importantly, it fails to recognise that capacity is an enabling factor, and that attainment of any cognitive function depends on developing the relevant knowledge, including procedural knowledge or strategies. Any attempt to dismiss this as hypothesis-saving would have to explain away our extensive studies of learning, especially the self-modifying production system model of strategy development (Halford et al., 1995).

Coch and Fischer, like Pascual-Leone, emphasise the similarity between our position and theirs. However, to the extent that Pascual-Leone’s and Fischer’s theories are similar to relational complexity theory, they must be similar to each other, but neither commentary seems to claim this. As with Anderson's ACT-R theory, discussed earlier, we welcome these comparisons, and we see some potential for constructive integration. We were of course well aware of the theories by Fischer and Pascual-Leone when our model was being formulated, and though we have already acknowledged some similarities, we believe it is clear that the differences are very real. We would like to note in passing that Chapman's (1987) formulation is possibly the closest to relational complexity theory of all the developmental theories, but even here the differences greatly outweigh the similarities. Many of the differences between relational complexity theory and Case's position, outlined in 6.3.2, would apply to Fischer's skill theory. To consider some other instances: Does Fischer's skill theory make the predictions, discussed above, that C&M, Goswami, G&R, or H&H find so controversial? Does skill theory conform to the criteria for relational knowledge in 2.2, and can it be implemented by the neural net architectures in 4.0? We see no grounds for believing that either Fischer's skill theory or Case's model offer general conceptual complexity metrics as they stand, but it may well be an interesting undertaking to build such a metric, possibly based on relational complexity, into those models.

Capacity in cognitive development

In the last two decades it has been fashionable to dismiss theories that postulate a role for capacity in cognitive development on a number of grounds, some of which appear in the commentaries. One is that capacity limitations in young children are inconsistent with a number of observed cases of precocious development. This argument is sometimes linked to another view, which is widespread though not often stated explicitly, that capacity theories are pessimistic because they imply limits to young children's capabilities. A third argument is simply that knowledge acquisition is "preferable" to capacity as an explanation of cognitive development. We suggest that existence of capacity limitations is not inconsistent, either with observed precocities, or a major role for knowledge in cognitive development.

In order to clarify these issues, it will be helpful to first shift our orientation from "what is done" to "how is it done". It then becomes clearer that the definition of processing capacity limitations is no more pessimistic about children than it is about adults. We have argued that adults typically process a maximum of four dimensions in parallel. This does not imply that adult cognition is limited, nor does it imply that adults cannot understand complex concepts. Our position is fully consistent with the observation that adults understand a lot of concepts that entail more than four dimensions. It implies that structures of more than four dimensions are processed by segmentation and conceptual chunking. This provides insights into the way cognitive processes operate, by indicating which processes can be performed in parallel and which must be serial.

The definition of capacity limitations can be equally productive in cognitive development once certain inhibiting misconceptions are cleared away. It does not say that children cannot perform certain tasks, any more than it says adults cannot understand the concept of acceleration because it is based on more than four dimensions. Furthermore, as noted in 6.3.1, relational complexity theory actually predicts some previously unrecognised capabilities of young children. It also makes a lot of predictions about how young children must be performing tasks, and it is clear that neither Goswami nor Wright has recognised this. Goswami says our caveats limit the testability of our theory. While it may mean that the theory is not amenable to some of the more simplistic tests, it generates a lot of highly testable predictions, a few of which are indicated in the target article or in this reply.

Capacity versus precocity: How versus what

Ever since the early 1970s the plausibility of capacity as an explanatory factor in cognitive development has been apparently undermined by a stream of data indicating precocious intellectual performance by infants and young children. This work has generated much excitement, and is even regarded by some people as the *raison d'etre* of cognitive development research. Contrary to the claims of Wright that we dismiss these data, we have no reason to doubt that children perform as reported in these studies. Furthermore, contrary to the suggestion by Goswami that we have not taken into account actual results in cognitive development research, we not only knew about those studies, but have given them very serious consideration. And contrary to the claim by C&M that our position is not empirically based, if any studies of precocious intellectual performance had posed a real challenge to our claims, we certainly would have abandoned or modified them.

The conflict between capacity theory and demonstrations of precocity is more apparent than real, and it is maintained by two deficiencies in the field. The first is our ignorance of the processes entailed in a lot of these performances, particularly when they are first discovered, and the second is a tendency to interpret children's performances by attributing complex processes, comparable to those used by adults, to them. We will illustrate.

Wynn (1992) showed that infants recognise the effect of adding objects to and subtracting objects from a display (i.e. 1+1 = 2, 2-1=1, 1+1≠3). It has been shown (Wynn, 1995) that this can be accounted for by an innate accumulator mechanism that we share with other species. This work provides an important basis for development of children's understanding of number, though of course infants' number comprehension is very different from that which occurs in middle childhood (Wynn, 1995). Suppose however that we had been ignorant of the cognitive processes entailed in infants' number performance. It could then have been argued that since infants can perform addition/subtraction, they process ternary relations at 5 months, which is wildly inconsistent with our norms of 5 years. The absurdity of such a claim would be easily recognised because we have reasonably good insights into the basis of infants' number knowledge. But how would these data have ben interpreted if we had had none of these insights? Psychologists could have used our own understanding of addition as a model of infants' performance. We suggest that this is what tends to happen where we do not know the cognitive basis for precocious performances. In those cases it often seems reasonable to make very optimistic assumptions about the processes infants and young children use. This creates an apparent disconfirmation of capacity limitations, though in fact the case for disconfirmation has not been made.

Perhaps the most famous instance of precocious performance seeming to disconfirm a developmental theory is Bryant and Trabasso's (1971) demonstration of transitive inference in 3- and 4-year olds. As noted in relation to C&M's commentary, some of these tasks require only primitive, nonrelational processes. Yet, although C&M argue enthusiastically that it is performed by primitive cognitive processes, they are unwilling to acknowledge a very important implication of this, which is that the task is not performed by, and is not a measure of, processing ternary relations. The effect is to provide a misleading and confused picture of what happens in cognitive development.

We pointed out in 6.2.4.4 that Goswami (1995) had over-interpreted her data as indicating ternary relational processing when they only require unary or binary relations, because of the widespread failure to analyse cognitive processes entailed in tasks. The same kind of failure appears again in Goswami's commentary. She cites work by Cutting (1996, cited by Goswami) in which 3-year olds are shown a green crayon and a yellow crayon covered with a blue filter, so both look green. Children recognise that the green crayon is better for drawing green grass. While we have no reason to doubt the finding, at least on the basis of her description in the commentary, it does not appear to require processing of ternary relations. The yellow crayon provides one cue for yellow and (when covered with the blue filter) one for green. The green crayon provides two cues for green. Analysing the task in terms of our model in 6.2.4.3, it does not entail having a cognitive representation of a ternary relation, in which the link between object-attributes and percept is conditional on a third variable, the viewing condition. Rather the percept depends on a single variable, degree-of-greenness. So the claim that the observation contradicts the relational complexity metric is another case of failing to consider the processes entailed in a task. It has been a persistent error in cognitive development research. Relational complexity analysis, insofar as it can be made on the basis of the brief description, indicates that the Cutting task is structurally simpler, and should be performed earlier than, theory of mind tasks, precisely what appears to have been observed. Therefore the data should probably be interpreted as *support* for relational complexity theory.

In 6.2.4.4 we showed that tasks which Goswami claims show 3-4 year olds process ternary relations (Goswami, 1995) could have been performed by unary or binary relations. Goswami does not contest our demonstration, but wants ternary relations to be the default explanation, and attempts to justify this by invoking parsimony. This seems to confuse parsimonious with simplistic. The range of phenomena that can be explained by the relational complexity metric, on the basis of a fairly small number of principles, may mean that it is the more parsimonious explanation. Either way, parsimony cannot be used to escape the obligation to provide evidence. This study is presented as though it were a major disconfirmation of our hypothesis, whereas in fact the data are perfectly consistent with our position. An interesting side issue is that Goswami claims her participants processed ternary relations earlier than 5 years because familiar relations were used. The title of the paper (Goswami, 1995) specifically attributes the success to the analogy of Goldilocks and the three bears. We are in complete agreement with the importance of familiar, interesting materials when assessing children (in fact we have used the same analog in our laboratory; (Rees, 1994). We have however noticed a tendency to claim that child-appropriate formats overcome all difficulties, without supporting evidence. Goswami's claims are another instance of this because the three-bears analog did not improve performance (Goswami, 1995, p. 883). Despite this, the alternative interpretation, that children's impressive performances should be attributed to use of simple relations, was not considered.

G&R argue that it is implausible that children could code an ordered three-tuple by chunking into binary relations. We think this underestimates the degree to which chunking is routinely employed by both children and adults. In tasks that entail complex relations, it is normal to chunk components which are not needed for the current decision. These chunks are then unpacked to make further decisions, though chunked and unchunked relations cannot be processed in the same decision of course. In this way complex tasks are decomposed into a series of simpler tasks that are performed successively. Furthermore, the chunking process would be facilitated by the visual presentation used by Gentner and Ratterman, because a child can scan the display and identify element(s) about which a decision must be made, chunking the rest. The chunks can then be unpacked to enable further decisions to be made. An ordered 3-tuple of the sort used by Gentner and Ratterman can be chunked into binary relational instances *first versus rest* (A,B/C), *middle versus rest* (B,A/C) and *last versus rest* (A/B,C). Chunking in this way would lead to a success rate of about 67 percent correct, which seems consistent with what 3-4 year olds can do in mapping ternary relations without additional support. In general, this kind of chunking can be used provided the relations between chunked terms do not need to be processed in the same decision as the rest of the relation. For example, chunking A,B,C into A,B/C, we can process A versus B/C in one decision, then unpack B/C to process B versus C. We cannot do this if relations between all three entities must be processed in one decision. Therefore, as noted in 3.4.1 and 6.0, testing the theory requires careful design to ensure that relational complexity cannot be reduced by chunking and segmentation.

Infancy

Three commentaries, Goswami, (H&H) and Wright suggest that we have not taken account of the large literature indicating impressive cognitive performances in infants. In fact this literature was considered extensively in developing our position, and we have written some reviews of it (e.g., Halford, 1993). However neither the commentaries nor the papers cited in them provide sufficient evidence that our criteria for relational processing were met. The problem is that, as illustrated above, when the cognitive processes used in a task are unknown we tend to interpret them in the most adult-like way, and this interpretation has often not been supported by fuller investigation. We appreciate the importance of innate cognitive mechanisms for quantification and identification of causes, but we should not assume without evidence that they have the same conceptual structure as later concepts. We cannot review this literature thoroughly here, but we will consider some representative cases.

H&H suggest that the ingenious work of Leslie and Keeble (1987) showing that 6-month old infants recognise cause, demonstrates that they understand binary relations. To us adults cause is, at least *prima facie*, a binary relation, *cause(a,b)*. H&H then make the common error described above of attributing adult cognition to infants: to adults *cause* is a binary relation, infants recognise cause, therefore infants appreciate binary relations. In fact Leslie and Keeble themselves suggest that cause could be recognised by a modular process that is essentially perceptual (see Leslie & Keeble, 1987, pp. 283-286). Are we engaging in hypothesis saving by dismissing recognition of cause as a module? Demonstrably no. In 2.2.1 we said relational schemas represent the structure of aspects of the world, and in 2.2.12 we said they provide a basis for planning, for analogical mapping, and can be modified on-line. The kind of module that Leslie and Keeble postulate can do none of these things, and there seems no reason to believe it would have any of the other properties of relational knowledge defined in 2.2. In fact, Leslie and Keeble themselves state: “A modular process, though it may be computationally very complex, nevertheless occurs in a fixed, automatic and mechanical way without being influenced by information or reasoning abilities that lie outside the module.” (Leslie & Keeble, 1987, p. 285). Thus, rather than their data being an embarrassment to us, their position is quite consistent with ours.

In this case the argument is clear because a plausible mechanism has been specified. The mechanism involved in infants’ understanding of vanished objects is currently more obscure, but we contend our position is at least as consistent with the evidence as is Goswami’s. Our theory does not entail denial of processes such as learning detours or inhibition of responses. We also accept that infancy research in the last two decades has been very successful in assessing competence uncontaminated by performance limitations. However our position is consistent, not only with what infants can do, but also with what they cannot. How, for example, does Goswami explain the failure of 10-month olds to discriminate number of objects on the basis of non-spatial attributes (Xu & Carey, 1996)? And if infants and young children have such sophisticated relational understanding, why do they have such difficulty with conditional discrimination?

Wright states repeatedly, and incorrectly, that we dismiss the work of Baillargeon. On the contrary, we have long recognised the importance of Baillargeon’s observations of infants’ object concept (see, for example, Halford, 1989), but we do not think it has been demonstrated to require the kind of explicit relational knowledge that we have defined.We have developed a theory that fits all these observations (Halford, 1996a, 1996b) and generates a lot of new predictions. We are not asking anyone to take our arguments on trust, but we strongly believe that they should not be dismissed because of poorly substantiated claims that they are inconsistent with developmental data.

Why processes matter

It seems appropriate to ask whether processes matter. After all, aren't we interested in what children can do? Perhaps process models are matters for cognitive science and are irrelevant to cognitive development. Hasn't cognitive development done well without addressing the complex and difficult question of processes, by defining achievements in terms of observable performances? This view has considerable appeal, if only because it seems to simplify the research task. However it amounts to a kind of neo-behaviourism, which was found to be unworkable in general cognition research by 1960. By this argument our investigations would go no deeper than concluding that 5-month olds understand arithmetic, that 3-year olds, monkeys and pigeons all understand transitivity (even, apparently, understand it in the same way), and so on. It also means that we would lack objective means to determine whether two tasks measure the same thing. This can cause difficulties if, on finding that children cannot perform a particular test, we keep simplifying it until we find one they do perform. The problem may be that we are unable to decide whether the simplified test assesses different cognitive processes than the original test. Consequently we might decide that the new test is "right" or "fair" and all the others are deemed to underestimate children's performance, oblivious to the fact that the tests measure different cognitive processes. This fallacy can have a seriously misleading effect on the field (Halford, 1989, 1993).

One riposte to this might be that capacity theory, like Piaget, emphasises what children cannot do. We know of no scholarly basis for this criticism of Piaget, and we consider it is not valid for capacity theory either. Surely the scientific investigation of cognitive development requires that we adopt an impartial approach, recording children's successes and failures with equal interest, as Piaget did. Then, and only then, will we be in a position to determine the factors that influence performance, and how these change with age. This has not always been the dominant orientation in the cognitive development literature however.

Analyses of cognitive processes increase the information yielded by research paradigms, which can be compared on the basis of the cognitive processes they entail. Consider, for example, the conflicting evidence about capability for transitive inference. Rather than arguing, as does Wright, that one transitivity paradigm is "right" and the others "wrong", we suggest it is more productive to ask why, for example, the transitivity of choice paradigm advocated by C&M can be performed by a wider range of participants (pigeons to university students) than the paradigm used by Andrews and Halford (in press) which causes great difficulty for children under a median age of five years. We suggest the explanation is that the paradigms require different cognitive processes, which make different cognitive demands. The transitivity of choice paradigm can be performed by basically associative processes (Wynne, 1995; Harris & McGonigle, 1994) whereas the paradigm used by Andrews and Halford requires a representation of relations in the task, and this imposes a load that is a function of the complexity of relations performed in each decision. From this and hundreds of similar comparisons, we have carefully constructed a theory of conceptual complexity, which makes a lot of testable predictions. The essence of the theory is outlined in the target article.

There is a tendency to discount complexity effects, no matter how large they are, because of an overriding concern with precocity. There is no reason however why these goals, to discover precocious performances, and to understand the effects of complexity, need be antagonistic. It is not a matter of emphasising what children can or cannot do, but of recognising that we are dealing with two sides of the coin. A thorough understanding of complexity, and of the way both children and adults deal with it, can be a major benefit in the goal of overcoming limitations.

Frye and Zelazo (F&Z) have demonstrated complexity effects in the important area of concept of mind (Frye, Zelazo, & Palfai, 1995). Though their work was originally independent of ours, their effects are consistent with those predicted by Halford (1993) and with the analysis in 6.2.4.3. Their findings are supported by work in our laboratory (Halford, Andrews, & Bowden, 1998). F&Z prefer to model these effects in terms of Cognitive Complexity and Control (CCC) theory. We have no objection to this because there may be benefits in choosing a formalism that is best suited to a particular set of phenomena. It should however be borne in mind that a hierarchically structured control process can be expressed as an *n*-ary relation. This translation is analogous to the way we defined the relational complexity of another hierarchical task, Tower of Hanoi, in 6.1.3. The benefit of doing so is that relational complexity provides a metric that is applicable to tasks with any kind of structure. This permits complexities of tasks in different domains, and with different surface structures, hierarchical or otherwise, to be compared directly. The only other response we would make is that we do not believe that in general we have neglected the link between cognition and performance. We have modelled the acquisition of performance strategies, guided by the person's concept of the task as well as other factors (Halford et al., 1995).

Complexity metric and interpretation of developmental data

Frye and Zelazo recognised an important benefit of the complexity metric when they commented that "Developmental psychology is commonly recognised as the study of change, but without a method for ordering the changes, the phenomena become as disorganised as those in the physical sciences would be without a periodic table." Many of the arguments we have made in earlier sections illustrate the point that without an objective procedure for analysing complexity, the orderly interpretation of developmental data is virtually impossible. If we cannot analyse the complexity of the processes entailed in cognitive performances, how can we decide whether the performance of infants on quantification, cause, or representation of vanished objects is the same as, or simpler than, that of young children? And how can we decide whether the transitive inference abilities of older children and adults surpass those of 3-year olds, or nonhuman primates? In the past this issue has sometimes been resolved simply by attributing the most sophisticated cognitive processes. Without a complexity metric developmental psychology is inevitably locked into unresolvable debates about which test(s) provide the truest or fairest indications of children's abilities. Continued refinement of assessments is essential of course, but given that most investigators are well aware of this, considerable benefits can be derived from objective assessments of the complexity of cognitive processes that underlie demonstrated performances.

Conclusion

Complexity effects are very real in the cognition of all higher animals, including human adults and children. A metric is needed which permits the cognitive complexity of tasks to be analysed, in such a way that it can be manipulated experimentally, unconfounded by other factors. Relational complexity theory provides such a metric. It has so far shown potential for dealing with species, age and individual differences. Such a wide-ranging theory inevitably makes contact with a number of other theories. We acknowledge some points of contact, and welcome attempts at constructive integration, but we are not aware of any other theory at present that offers the kind of general conceptual complexity metric that we have proposed, applicable to tasks that can be performed in a single processing step as well as to tasks that entail more than one step. Complexity effects are based on the cognitive processes actually used in performing a task, so valid process analyses are a prerequisite to reliable and testable assessments of complexity.

One consequence, not always recognised by commentators, is that although we accept empirically valid demonstrations of infants' and children's performances, their implications for our position depend on analysing the processes they entail. That is, while we accept data about *what* infants and children can do, we believe the field needs to be equally concerned with *how* they do it. Lack of interest in processes seems to have been the main feature that has distinguished cognitive development from general cognition research in the last three decades. The cognitive revolution has, arguably, had much less impact on cognitive development, and this may have been to the detriment of the field. Claims of performances more precocious than our norms appear to be based on assumptions about processes used by infants and children, rather than on evidence that they are really processing relations of a given complexity. Where process analyses have been performed it has generally been found that the data are quite consistent with our suggested norms.

Reference List

Anderson, J. R. (1990). The adaptive character of thought. (Vol. 1). Hillsdale, NJ: Erlbaum.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. Cognitive Psychology, 30(3), 221-256.

Anderson, M. (1992). Intelligence and development: A cognitive theory. Oxford UK Cambridge USA: Blackwell.

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

Andrews, G., & Halford, G. S. (in preparation). Relational complexity: Evidence for domain generality and age related changes during childhood.

Andrews, G., & Halford, G. S. (in press). Children's ability to make transitive inferences: The importance of premise integration and structural complexity. Cognitive Development.

Bryant, P. E., & Trabasso, T. (1971). Transitive inferences and memory in young children. Nature, 232, 456-458.

Chalmers, M., & McGonigle, B. (1984). Are children any more logical than monkeys on the five-term series problem? Journal of Experimental Child Psychology, 37, 355-377.

Chapman, M. (1987). Piaget, attentional capacity, and the functional limitations of formal structure. Advances in Child Development and Behaviour, 20, 289-334.

Foley. (1997). Assessing conceptual complexity in hierarchical reasoning: A dual-task approach. Unpublished doctoral disertation, University of Cincinnati, Ohio.

Foley, E. J., & Berch, D. B. (1997). Capacity limitations of a classic M-Power measure: A modified dual-task approach. Journal of Experimental Child Psychology, 66, 129-143.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. Cognitive Development, 10, 483-527.

Goswami, U. (1995). Transitive relational mappings in three- and four-year-olds: The analogy of Goldilocks and the three bears. Child Development, 66, 877-892.

Goswami, U., & Brown, A. L. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. Cognition, 35, 69-95.

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

Halford, G. S. (1982). The development of thought. Hillsdale, NJ: Lawrence Erlbaum Associates.

Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? Cognitive Psychology, 16, 65-93.

Halford, G. S. (1989). Reflections on 25 years of Piagetian cognitive developmental psychology, 1963-1988. Human Development, 32, 325-387.

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

Halford, G. S. (1995). Learning processes in cognitive development: A reassessment with some unexpected implications. Human Development, 38(6), 295-301.

Halford, G. S. (1996a). Capacity limitations in processing relations: Implications and causes. Paper presented at IIAS 3rd Brain and Mind International Symposium on Concept Formation, Thinking and their Development, Kyoto, Japan, .

Halford, G. S. (1996b, August 12-16). Relational knowledge in higher cognitive processes. Paper presented in symposium entitled Relational Knowledge in Higher Cognitive Processes at the XIVth Biennial Meeting of the International Society for the Study of Behavioral Development, Quebec City, .

Halford, G. S., Andrews, G., & Bowden, D. (1998, July). Complexity as a factor in children's theory of mind. Poster session presented at the XVth Biennial Meetings of the ISSBD, Berne, Switzerland, .

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

Halford, G. S., & Leitch, E. (1989). Processing load constraints: A structure-mapping approach. In M. A. Luszcz & T. Nettelbeck (Eds.), Psychological development: Perspectives across the life-span (pp. 151-159). Amsterdam: North-Holland.

Halford, G. S., & MacDonald, C. (1977). Children's pattern construction as a function of age and complexity. Child Development, 48, 1096-1100.

Halford, G. S., Maybery, M. T., & Bain, J. D. (1986). Capacity limitations in children's reasoning: A dual task approach. Child Development, 57, 616-627.

Halford, G. S., Maybery, M. T., & Bain, J. D. (1988). Set-size effects in primary memory: An age-related capacity limitation? Memory and Cognition, 16(5), 480-487.

Halford, G. S., Smith, S. B., Dickson, J. C., Maybery, M. T., Kelly, M. E., Bain, J. D., & Stewart, J. E. M. (1995). Modelling the development of reasoning strategies: The roles of analogy, knowledge, and capacity. In T. Simon & G. S. Halford (Eds.), Developing cognitive competence: New approaches to cognitive modelling (pp. 77-156). Hillsdale, NJ: Erlbaum.

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

Harris, M. R., & McGonigle, B. O. (1994). A model of transitive choice. The Quarterly Journal of Experimental Psychology, 47B(3), 319-348.

Henderson, J. (1994). Connectionist syntactic parsing using temporal variable binding. Journal of Psycholinguistic Research, 23(5), 353-379.

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

Hunt, E., & Lansman, M. (1982). Individual differences in attention. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (pp. Vol. 1, pp. 207-254). Hillsdale, NJ: Lawrence Erlbaum Associates.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychological Review, 99(1), 122-149.

Kail, R. (1991). Processing time declines exponentially during childhood and adolescence. Developmental Psychology, 27(2), 259-266.

Leeuwenberg, E. L. L. (1969). Quantitative specification of information in sequential patterns. Psychological Review, 76, 216-220.

Leslie, A., & Keeble, S. (1987). Do six-month-old infants perceive causality? Cognition, 25, 265-288.

Maybery, M. T., Bain, J. D., & Halford, G. S. (1986). Information processing demands of transitive inference. Journal of Experimental Psychology: Learning, Memory and Cognition, 12(4), 600-613.

McGonigle, B. O., & Jones, B. T. (1978). Levels of stimulus processing by the squirrel monkey: Relative and absolute judgements compared. Perception, 7, 635-659.

McNicol, D., & Stewart, G. W. (1980). Reaction time and the study of memory. London: Academic Press.

Murdock, B. B. (1983). A distributed memory model for serial-order information. Psychological Review, 90, 316-338.

Navon, D. (1984). Resources: A theoretical soup stone? Psychological Review, 91, 216-234.

Navon, D., & Gopher, D. (1979). On the economy of the human processing system. Psychological Review, 86, 214-255.

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

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

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

Rees, B. E. (1994). Children's ability to map ordered sets. Unpublished Honours Thesis, University of Queensland, Brisbane.

Rips, L. J. (1989). The psychology of knights and knaves. Cognition, 31, 85-116.

Robin, N., & Holyoak, K. J. (1994). Relational complexity and the functions of prefrontal cortex. In M. S. Gazzaniga (Ed.), The Cognitive Neurosciences (pp. 987-997). Cambridge, MA: MIT Press.

Sternberg, S. (1975). Memory scanning: New findings and current controversies. Quarterly Journal of Experimental Psychology, 27, 1-32.

Thatcher, R. W. (1994). Cyclic cortical reorganization: Origins of human cognitive development. In G. Dawson & K. W. Fischer (Eds.), Human Behavior and the Developing Brain (pp. 232-266). New York: Guilford.

Wynn, K. (1992). Addition and subtraction by human infants. Nature, 358, 749-750.

Wynn, K. (1995). Origins of numerical knowledge. Mathematical Cognition, 1(1), 35-60.

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

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. Cognitive Psychology, 30, 111-153.