Opinion TRENDS in Cognitive Sciences Vol.11 No.6

Separating cognitive capacity from knowledge: a new hypothesis

Graeme S. Halford[1], Nelson Cowan[2] and Glenda Andrews[3]

[1] Griffith University, Mt Gravatt campus, 170 Kessels Road, Nathan, QLD 4111, Australia
[2] Department of Psychological Sciences, University of Missouri, 18 McAlester Hall, Columbia, MO 65211, USA
[3] Gold Coast campus, Griffith University PMB 50 Gold Coast Mail Centre, QLD 9726, Australia

We propose that working memory and reasoning share related capacity limits. These limits are quantified in terms of the number of items that can be kept active in working memory, and the number of interrelationships between elements that can be kept active in reasoning. The latter defines the complexity of reasoning problems and the processing loads they impose. Principled procedures for measuring, controlling or limiting recoding and other strategies for reducing memory and processing loads have opened up new research opportunities, and yielded orderly quantification of capacity limits in both memory and reasoning. We argue that both types of limit might be based on the limited ability to form and preserve bindings between elements in memory.

Introduction

The techniques that humans have for making the best use of available information-processing capacity are of immense value but they have to be controlled to study capacity limitations and the effects of complexity. New, principled procedures for measuring, controlling or limiting recoding and other strategies by which subjects reduce memory and processing loads have enabled complexity and capacity effects to be investigated independently of, and in interaction with, knowledge. Here, we present the hypothesis that this development enables a new, integrated treatment of reasoning and working memory (WM), including an orderly quantification of capacity limits, and that this has opened up new research opportunities.

WM and reasoning

Developments in both theory and methodology have strengthened the links between WM and reasoning, and some salient points are summarized in Box 1.

We propose that the essential link between WM and reasoning is in the common requirement to bind elements to a coordinate system. Consider, first, short-term serial recall of the words, 'Fido, Rover, Cleo'. The words are assigned to ordinal positions when presented (Figure 1a) but this assignment must be maintained for later recall, and this requires attention. Even in free recall (not shown), items on a trial must be bound to the present trial concept or node in memory; binding might be even more extensive, inasmuch as an associative network between items would greatly facilitate recall. Now consider a choice reaction-time task, in which participants press a different button in response to one of several lights, and the buttons are assigned to lights randomly (noncompatible mappings; Figure 1b). Maintaining the binding of button positions to light positions in the WM requires attention [1]. Finally, consider a transitive inference problem, such as 'Jane is taller than Wendy, Amelia is taller than Jane'. This can be solved by mapping premise elements into an ordering schema, as shown in Figure 1c.

Maintaining bindings between elements and slots using attention is common to WM and to reasoning. This is not the only common factor but there is substantial evidence that WM capacity (WMC) accounts for a sizeable proportion of the variance in reasoning [1,2] and intelligence [3]. WM and reasoning differ in whether the binding is supplied with the input (as in short-term serial recall) or has to be constructed by the reasoner, as in syllogistic (including transitive) inference, where premise elements have to be mapped to slots in a mental model in a way that is consistent with the premises. There are intermediate cases, such as where the mapping of responses to inputs, although predetermined, must be coded by the performer, or where participants construct their own mnemonics. Coordinate systems in the WM are less likely to include explicit relational representations than those in reasoning. Use of an explicit symbol that differentiates different kinds of links is a feature that distinguishes relations from associations [4]. In Figure 1c, the relation 'taller than' is explicit, whereas the ordering schemes used in the WM might not be.

We further propose that the common demand for attention when binding elements into slots is a possible explanation for common capacity limitations in WM and reasoning (Box 2). Here, rather than presenting a detailed model of these processes [4], we focus on methods for controlling chunking, recoding, rehearsal and other strategies, so that capacity limits can be measured.

WM capacity

Immediate memory capacity is roughly constant if measured in meaningful units, termed 'chunks' [5]. For example, FBICIAIRS comprises up to nine chunks, but only three chunks if the observer recognizes three familiar acronyms (FBI, CIA, IRS). WMC can only be estimated if the number of chunks can be determined, and techniques for this are summarized in Box 2.

Corresponding author: Halford, G.S. (g.halford@griffith.edu.au).

1364-6613/$ – see front matter © 2007 Elsevier Ltd. All rights reserved. doi:10.1016/j.tics.2007.04.001

Box 1. Capacity effects in WM and reasoning

- The core of WM is the temporary binding of elements to a coordinate system [16,34] that is closely related to relational representations used in reasoning [1,4]. Temporary binding to structural representations possibly accounts for the strong relationship between WMC and Gf [1].
- Capacity limits in both WM and reasoning can be attributed to the number of bindings to slots in a coordinate system or relation. WM is limited to approximately four items that can be kept active [6], whereas representations in reasoning are limited to four interrelated variables [43].
- Latent variable constructs of WMC account for approximately 0.60 of the variance in reasoning and Gf [2].
- WM has a domain-general component that is crucial to its prediction of reasoning and Gf [2].
- New assessments of WMC measure how many elements fit in the focus of attention [3] or capacity-limited region [34] more explicitly than traditional sentence and operation spans. These include: computer-paced reading of numbers, or performing simple operations of +1 or –1, while retaining words or letters for later recall [44]; or presentations too rapid and unpredictable to enable rehearsal [3].

A review [6] of disparate procedures in which items cannot be grouped or rehearsed (Box 2) converges on a range of 2–6 items in adults (usually 3–5), and fewer in children or the elderly [7]. In cases of exceptional or expert memory, it is apparently the size of recalled chunks, rather than their number, that is extraordinary [8].

Measuring complex chunks

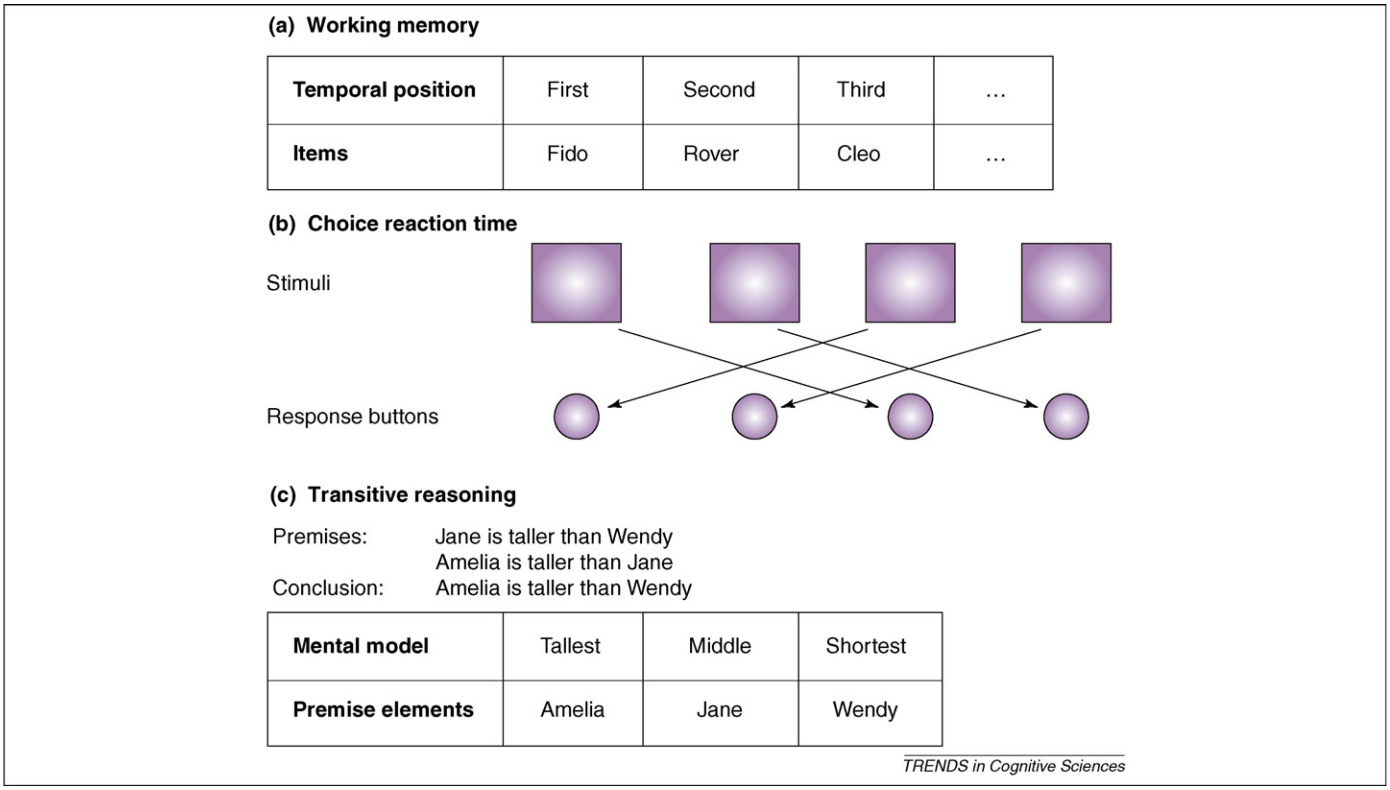
A second approach to assessing capacity limits in chunks (Box 2) is to manipulate and estimate chunking processes [9]. In a recent study using serial recall [10], pairs of words were taught with 0–4 paired exposures (and a complementary number of unpaired exposures, resulting in four exposures in each case), and cued recall was used to assess pair learning. A mathematical model that used cued recall and the order of items in serial recall to estimate items per chunk suggested that pair learning increased chunk size but left the number of chunks recalled constant at ~3.5 on average.

Follow-up work [11] extended this method. With long lists and free recall, or free scoring of serial recall, chunk capacity limits determined recall (e.g. six well-learned pairs were recalled with similar accuracy to six singletons; ~3.5 units in each case). With shorter lists and strict serial scoring, however, length limits predominated (e.g. four well-learned pairs and eight singletons were recalled equally well but more poorly than four singletons). Thus, participants might rehearse ~2 seconds of information without taxing the WM chunk limit.

Central capacity

Theoretically, capacity limits could occur separately in different parts of the processing system. Recent evidence begins to address questions such as (i) whether there is a central capacity that can be allocated across different modalities and codes, and (ii) where the capacity might reside in the brain.

One can combine a spatial memory task with a verbal task to determine whether the amount retained from one task affects the amount retained from the other. Sometimes,



Temporal position	First	Second	Third	…
Items	Fido	Rover	Cleo	…

Premises: Jane is taller than Wendy
Amelia is taller than Jane
Conclusion: Amelia is taller than Wendy

Mental model	Tallest	Middle	Shortest
Premise elements	Amelia	Jane	Wendy

Figure 1. Binding of elements into coordinate systems in **(a)** WM, **(b)** choice reaction time and **(c)** reasoning. (a) In WM, the coordinate system corresponds to spatial or temporal position. The identity of items in each location or position must be maintained in the presence of distraction or updated during the task. (b) The coordinate system corresponds to spatial location. In the non-compatible arbitrary condition, the left–right position of the stimuli is not consistent with the position of the response buttons. Temporary bindings between stimulus and response representations must be constructed and maintained in the presence of irrelevant information (colour). This imposes demands on attentional capacity. (c) Relational premises are integrated to generate an ordered triple. The coordinate system corresponds to the ordinal position. Each position functions as a variable.

Box 2. Some techniques to examine WMC limits in chunks

The following methods have been used to examine WMC limits and show, we believe, 3–5 chunks, on average. Various manipulations help to identify the number of items per retrieved chunk and, therefore, the number of chunks retrieved.

- *Spatial arrays*. A visual array (spatial arrangement) is briefly presented for comparison with a probe array identical to the first array or differing in the identity of one item [45]. It is too fast for grouping, so presumably each item is a chunk.
- *Perfect serial recall*: The length of list enabling errorless recall reflects the chunk capacity, assuming that grouping strategies cannot be used consistently [46].
- *Running span*. Items are presented rapidly, in a list with an unpredictable endpoint. The participant then recalls items from the end of the list. Given this presentation, it is impossible to group items as in ordinary serial recall, so each item is presumably one chunk [47].
- *Span with distraction or suppression*. A visual or auditory list is presented. A secondary task prevents grouping and rehearsal, so each item is assumed to be one chunk [21].
- *Categorical retrieval.* Items are to be recalled from a meaningful category (e.g. states). Each burst of items recalled is assumed to reflect WM filled to capacity from long-term memory once [48].
- *Memory updating.* Items are to be remembered, some of which are eligible to be changed (updated). Responses slow markedly as eligible items increase from one to three [34].
- *Probed recall*. Participants indicate if a probe occurred in a just-presented list. At least three semantically related items near the end of the list can show especially fast retrieval [49].There is interference from other, similar lists if the number of items exceeds basic capacity [50].
- *Multiobject tracking* (examines attention limits related to WM). An array of items includes several cued as targets. The cues disappear; all items then move in different random trajectories. When they stop, the participant must identify the targets. Targets up to the capacity limit can be tracked together as one changing object [51].
- *Reconstruction by chunks.* A structured configuration (chess setup) is examined and reconstructed from memory. The pieces reconstructed without pause presumably form a single chunk; capacity is estimated as chunks recalled per look [52].
- *Recollection by chunks.* Word series with varying interitem association strengths are presented; words recalled in the presented order are taken to reflect a single chunk [9,11].

little interference between verbal and spatial tasks is observed [12,13], consistent with the view that storage is modality specific [14]. However, recent studies show that verbal and spatial tasks interfere with one another under some circumstances. It is not because both tasks share verbal rehearsal. Morey and Cowan [15] obtained interference between a spatial memory task and a random seven-digit list, but not between spatial memory and a memorized seven-digit number or a random two-digit number. Most of the interference occurred only when the verbal items were spoken aloud. This suggests that some mnemonic processes are not part of the store used for visual memory (e.g. rehearsal or auditory memory) but that other processes are (e.g. overt verbal retrieval). In unpublished research, J. Scott Saults and N. Cowan presented concurrent arrays of spoken digits and coloured squares, and showed that, if sensory memory cues were eliminated, a one-to-one trade-off between modalities emerged. Subjects could retain either ~4 squares, in an attend-visual condition, or ~4 items including squares and digits together, in an attend-bimodal condition. Thus, the central capacity could be allocated across modalities.

Another study [16] examined memory for verbal-spatial associations (names shown in schematic houses at different screen locations). Adults could perform this task by combining modality-specific memories for the name sequence in the trial and for the spatial path between the houses in which the names appeared, provided that there was a one-to-one correspondence between names and houses. When this strategy was precluded, crossmodal associations presumably had to be retained. This occurred when some houses contained two names and others contained none, when verbal rehearsal of the names was suppressed or when the subjects were children too young to rehearse.

There is some evidence regarding where central capacity might reside. The parietal lobes take part in an attention system that integrates information across modalities, and might correspond to the seat of attention, separate from frontal areas involved in controlling attention [17]. Although both areas are often active together in WM tasks that involve the storage and manipulation of information, transcranial magnetic stimulation (TMS) to the frontal lobes disrupts only the manipulation, whereas parietal stimulation disrupts storage as well [18]. Some parietal areas show functional magnetic resonance imaging responses to simple visual arrays that change across memory loads in a way similar to behavioural responses [19]. When feature complexity is added to the stimuli by requiring memory for several features (e.g. colour in addition to shape), some areas respond according to the number of objects, and others respond according to complexity [20]. The brain data seem to be compatible with a central store that reflects the focus of attention used as a store [21], although it has not yet been proven that parietal storage does reflect the focus of attention.

Chunk capacity and chunk size limits

People recall lists best if they are separated into groups of 2–4 items, the optimal typically being three [22,23]. It could be viewed as coincidental that the optimal group size is approximately the same as the maximal number of chunks that can be recalled. However, it might be more than coincidental, if the capacity limit is attention related. Formation of a new chunk might be possible only if the elements are retained in the capacity-limited store (focus of attention?) concurrently, enabling rich associations to form between the items [21]. The optimal group size is therefore the largest number of items that can be reliably stored to be interassociated at once. Chunks of any size can be built up, of course, using the capacity-limited region reiteratively. The fact that larger chunks can be built up with practice and attention, and that the WM no longer seems to limit performance [8], does not mean that chunk capacity limits cease to have a role; expertise leads to larger chunks [24] (like the configurations in a chess game) but, when independent estimates of capacity are obtained, they still equal ~3–4 chunks in adults, and fewer in children [3,25].

Links between WM and reasoning

Accumulated evidence indicates that estimates of WMC have some degree of domain generality and have a central tendency of ~3.5 items. We hypothesize that the WM limit

Table 1. Arity of relations, number of slots or variables, an example of each level and approximate median age of attainment[a]

Arity of relation	Slots or variables	Example	Median age
Unary	One	Class membership, e.g. cat(Marcus)	One year
Binary	Two	Larger(elephant, mouse)	1.5 years
Ternary	Three	Addition(2,3,5)	5 years
Quaternary	Four	Proportion(2,3,6,9)	11 years

[a] In RC theory, processing load depends on the complexity of relations processed in any step of a task. There are two mechanisms for reducing processing loads.

reflects capacity for attention, which determines the number of elements, whether items or chunks, that can be bound into a coordinate system. This hypothesis links WMC directly to a theory of capacity limitations in reasoning, relational complexity (RC) theory [4], considered next, which proposes that reasoning is limited by the number of variables or slots that can be related in a single representation.

Complexity in reasoning

Measures of the number of symbols required to define a concept [26,27] have achieved significant successes, as have metrics based on rule hierarchies [28]. Complexity metrics have also been important in cognitive development [29]. However, we focus on a metric that discriminates between complexity and knowledge effects, and which incorporates binding of problem elements to mental models.

The RC metric

Complexity can be defined by the number of slots or 'arity' of a relation that must be represented to perform a specific cognitive process (for example, a binary relation, such as 'larger than' has two slots, one each for the larger and smaller entities) [4,30–32]. Each slot of a relation can be filled in a variety of ways, and corresponds to a variable or dimension, and an n-ary relation is a set of points in n-dimensional space. RC norms are summarized in Table 1. The relation is a coordinate system, as noted earlier, so RC corresponds to the number of bindings to a coordinate system, consistent with our suggestion of how capacity is defined in the WM.

'Conceptual chunking' involves recoding concepts into less complex relations. For example, speed = distance/time, is a ternary relation but speed can be recoded into a unary relation, speed(60 kph), as when speed is indicated by a pointer on a dial. Although chunking reduces processing load, chunked variables cannot be accessed. Thus, the chunked representation of speed does not enable us to answer questions such as, 'how does speed change if we cover the same distance in half the time?', and we must revert to the ternary relation to calculate the answer. (Similarly, if one needs to access part of a chunk in the WM or its internal structure, the chunk must be unpacked.) Conceptual chunking is analogous to mnemonic chunking, in that it compresses a representation into fewer variables, but explicit relations are defined between the variables, and they are not simply associated (e.g. velocity is not an association between speed, distance and time but a specific relation defined between them).

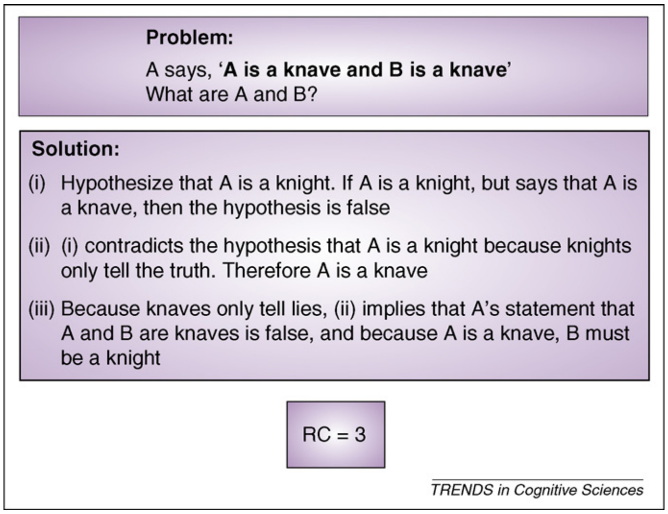
'Segmentation' entails breaking tasks into less complex steps, which can be processed serially. Strategies and algorithms are common ways of doing this (e.g. adding one column at a time in multidigit addition). Representing complex structures hierarchically, and processing one level of the hierarchy at a time, is also an effective way to segment tasks [28]. Chunking and segmentation skills are important components of expertise.

There is common ground between WM and reasoning. First, both are limited by the ability to map elements into coordinate systems, which in turn depends on attention [33,34]. Second, both involve compressing material into chunks, mnemonic or (for reasoning) conceptual in nature. Conceptual chunks contain more relational information than do mnemonic chunks because the coordinate systems in reasoning are explicit relations. Third, the limitation is in the number of independent components, irrespective of their size. Chunk size in the WM is analogous to the number of possible instantiations of a relational slot in reasoning (i.e. whether a slot can be instantiated in 2, 3, . . ., or arbitrarily many ways). In both cases, this factor has less influence on capacity than does the number of chunks in the WM, or the number of slots in reasoning.

Method for analysis of RC (MARC)

MARC principles that have been found to be consistently valid across many domains are shown in Box 3. Application of these principles shows some tasks are difficult because they resist decomposition. Two examples are discussed here (Figures 1c,2).

Transitive inference imposes a processing load [30] because both premises need to be considered to assign premise elements uniquely into slots in the representation in the WM (Figure 1c). There are three slots, so RC = 3. Premise integration in transitive inference involves neural circuits in the prefrontal and parietal cortices [35,36]. Prefrontal regions are also activated in other reasoning



Problem:

A says, '**A is a knave and B is a knave**'
What are A and B?

Solution:

(i) Hypothesize that A is a knight. If A is a knight, but says that A is a knave, then the hypothesis is false

(ii) (i) contradicts the hypothesis that A is a knight because knights only tell the truth. Therefore A is a knave

(iii) Because knaves only tell lies, (ii) implies that A's statement that A and B are knaves is false, and because A is a knave, B must be a knight

RC = 3

TRENDS in Cognitive Sciences

Figure 2. An application of the MARC to the knight and knaves reasoning task.

Box 3. Principles of the MARC

Complexity analyses begin with established models of reasoning processes, taking account of knowledge and strategies likely to be available to the reasoner.

Assessments must be appropriate for the participant sample, and control tests are essential (e.g. closely matched binary-relational tasks are important controls for assessing ternary relations) [30].

1. Effective RC for a cognitive process is the least complex relation required to represent it. The required mental models are determined from established process models, supported by empirical research.
2. When tasks entail more than one step, the processing complexity of the task is the relation that must be represented to perform the most complex step, using the least demanding strategy available to humans for that task [4].
3. Variables cannot be chunked or segmented if relations between them must be processed because relations between chunked or segmented variables become inaccessible. This leads to several corollaries:
   - Chunking sets. The relation between A and both B and C is binary if the relation between B and C can be chunked, which is possible if the relation(s) between them need not be processed. In oddity tasks, a strategy of noting that red is different from three blues has RC = 2 [53].
   - Interaction between variables. Variables cannot be chunked or segmented if they interact because the influence of each variable is modified by the others. Therefore, interacting variables must be processed jointly.
   - When two arguments of a relation function as a unit for the cognitive process being performed, they can be chunked. In Figure 2 (of the main text), ‘A is a knave and B is a knave’ is one conceptual chunk for the first step. Similarly, if A is a knight and A says (some proposition), then the predicates ‘is a knight’ and ‘says’ can be chunked because they are a unit, equivalent to ‘The knight, A says’ [41].
4. When a cognitive overload cannot be reduced by conceptual chunking or segmentation, it is likely to be handled by default to a simpler representation, usually by ignoring one or more variables, leading to fragmentation of knowledge.

and fluid intelligence (Gf) tasks that require relational integration, including modified Ravens Matrices problems [36,37] and analogy problems [38]. Moreover, as task complexity increases, more anterior regions seem to be recruited [39,40].

Knight and knave problems are segmented into separate inference steps (Figure 2) but each inference entails compound propositions. In step 1, the proposition knight(A) contradicts the proposition knave(A), which implies that knight(A) is false. Applications of the MARC [41] result in three conceptual chunks: ‘Knight A says’, ‘A and B are knaves’ (as a chunked argument of ‘says’), and ‘the hypothesis is false’. Therefore, for this step RC = 3.

The need to interpret inputs jointly constrains decomposition into simpler subtasks. Some tasks that are difficult or complex (e.g. according to cognitive complexity and control theory [28]) incorporate this factor [42]. Interpretation of interactions was used to determine how many variables humans can process in parallel, while controlling techniques for reducing processing load. Performance declined as complexity increased, and was at chance level on five-way interactions [43]. Given that an n-way interaction is defined on n variables, so content is assigned to n slots, the data indicate that humans are limited to processing four variables.

Summary

We have proposed similar limitations in WM and reasoning, which seem to reflect a central capacity limit. WM is limited to ~3–4 chunks, and reasoning is limited to relations between four variables. One priority is to identify the common underlying mechanisms. A promising hypothesis is that both WM and reasoning require items or concepts to be the focus of attention concurrently. This would enable them to be interassociated to form chunks or bound into coordinate systems, including relational representations that enable inferences to be made. Another priority is to determine the underlying brain systems. Frontoparietal networks have been implicated. Parietal regions seem to be most important for focus of attention and storage

Box 4. Outstanding questions

Techniques for the quantification of reasoning complexity and the assessment of WMC imply potential advancements in several areas. Some of these have begun to be realized but there is scope for further research. The following questions outline some possibilities.

- *WMC*. Does the WM limit of 3–5 items reflect the focus of attention? Is there a pool of capacity that is general across domains? Is there also a limit on how long ideas remain active when no longer attended, or not [54]? If not, then what is special about the attended information? Is it that only information in the focus of attention can be bound together [6,55]?
- *WM and reasoning*. Are all possible interrelations formed between items concurrently in the focus of attention? Is that why WMC is limited [6,55]? Does such mutual binding enable new chunk formation and, simultaneously, comprehension?
- *Locus of capacity*. Where in the brain might central capacity reside? In the parietal cortex [18]?
- *Theoretical integration*. Complexity analyses are bringing a more orderly interpretation to cognitive developmental findings [56]. Does this resolve apparent conflicts, such as those between precocity and developmental change?
- *Classical findings*. Hirst *et al.* [57] demonstrated an improvement in concurrent tasks with practice, apparently indicating that there was no capacity limit, but this might not hold if expertise and WMC were assessed independently. Chi’s finding of superior retention by child chess experts, as compared with adult chess novices, demonstrates a powerful effect of knowledge on short-term memory [58]; however, when capacity is measured, it has been found to increase with age. Should these findings be reassessed in the light of contemporary knowledge of WM?
- *Assessment criteria in reasoning*. Given that the norms of logic are no longer seen as appropriate criteria for human reasoning, does cognitive complexity provide an alternative basis for assessment [59]?
- *Transition processes*. Increasingly complex concepts can be acquired by adding variables to representations (e.g. density helps to account for variations in size and weight). Acquisition depends both on knowledge and on capacity to process the extra dimension; however, chunking and segmentation can reduce processing loads. Do these factors have implications for transition processes in cognitive development?
- *Improving experimental design*. The effects of complexity are not always recognized. For example, the difference between categorization by rules and by information integration could be accounted for by complexity differences [60]. Should the effect of complexity be considered in experimental design, to avoid confounding by other variables?
- *Behavioural and brain imaging research*. Recent research suggests that brain regions such as the rostrolateral prefrontal cortex are more selective to cognitive complexity than to cognitive domain [39]. Can precise manipulation of cognitive complexity, with other factors controlled, increase the precision of brain imaging research?

functions of WM, whereas prefrontal regions seem to be more important for manipulation functions in WM and reasoning. If manipulation increases RC, then TMS applied to prefrontal regions should be more disruptive as the RC of a task increases. Procedures for controlling simplifying strategies have opened the way for these and other lines of investigation into WM, reasoning, the links between them, and the underlying brain systems (Box 4).

References

1 Wilhelm, O. and Oberauer, K. (2006) Why are reasoning ability and working memory capacity related to mental speed? An investigation of stimulus–response compatibility in choice reaction time tasks. *Eur. J. Cogn. Psychol.* 18, 18–50

2 Kane, M.J. *et al.* (2004) The generality of working memory capacity: a latent variable approach to verbal and visuospatial memory span and reasoning. *J. Exp. Psychol. Gen.* 133, 189–217

3 Cowan, N. *et al.* (2005) On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cognit. Psychol.* 51, 42–100

4 Halford, G.S. *et al.* (1998) Processing capacity defined by relational complexity: implications for comparative, developmental, and cognitive psychology. *Behav. Brain Sci.* 21, 803–831

5 Miller, G.A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97

6 Cowan, N. (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–185

7 Naveh-Benjamin, M. *et al.* Age-related differences in immediate serial recall: dissociating chunk formation and capacity. *Mem. Cognit.* (in press)

8 Ericsson, K.A. *et al.* (2004) Uncovering the structure of a memorist's superior 'basic' memory capacity. *Cognit. Psychol.* 49, 191–237

9 Tulving, E. and Patkau, J.E. (1962) Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Can. J. Psychol.* 16, 83–95

10 Cowan, N. *et al.* (2004) Constant capacity in an immediate serial-recall task: a logical sequel to Miller (1956). *Psychol. Sci.* 15, 634–640

11 Chen, Z. and Cowan, N. (2005) Chunk limits and length limits in immediate recall: a reconciliation. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1235–1249

12 Cocchini, G. *et al.* (2002) Concurrent performance of two memory tasks: evidence for domain-specific working memory systems. *Mem. Cognit.* 30, 1086–1095

13 Fougnie, D. and Marois, R. (2006) Distinct capacity limits for attention and working memory: evidence from attentive tracking and visual working memory paradigms. *Psychol. Sci.* 17, 526–534

14 Baddeley, A.D. (1986) *Working Memory. Oxford Psychology Series #11,* Clarendon Press

15 Morey, C.C. and Cowan, N. (2005) When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 703–713

16 Cowan, N. *et al.* (2006) Development of working memory for verbal–spatial associations. *J. Mem. Lang.* 55, 274–289

17 Posner, M.I. and Peterson, S.E. (1990) The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42

18 Postle, B.R. *et al.* (2006) Repetitive transcranial magnetic stimulation dissociates working memory manipulation from retention functions in the prefrontal, but not posterior parietal, cortex. *J. Cogn. Neurosci.* 18, 1712–1722

19 Todd, J.J. and Marois, R. (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754

20 Xu, Y. and Chun, M.M. (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440, 91–95

21 Cowan, N. (2005) *Working Memory Capacity,* Psychology Press

22 Ryan, J. (1969) Grouping and short-term memory: different means and patterns of groups. *Q. J. Exp. Psychol.* 21, 137–147

23 Wickelgren, W.A. (1964) Size of rehearsal group and short-term memory. *J. Exp. Psychol.* 68, 413–419

24 Chi, M.T.H. (1976) Short-term memory limitations in children: capacity or processing deficits? *Mem. Cognit.* 4, 559–572

25 Burtis, P.J. (1982) Capacity increase and chunking in the development of short-term memory. *J. Exp. Child Psychol.* 34, 387–413

26 Feldman, J. (2000) Minimization of Boolean complexity in human concept learning. *Nature* 407, 630–632

27 Leeuwenberg, E.L.L. (1969) Quantitative specification of information in sequential patterns. *Psychol. Rev.* 76, 216–220

28 Zelazo, P.D. *et al.* (2003) The development of executive function in early childhood. *Monogr. Soc. Res. Child. Dev.* 68, vii–137

29 Halford, G.S. (2002) Information processing models of cognitive development. In *Blackwell Handbook of Childhood Cognitive Development* (Goswami, U., ed.), pp. 555–574, Blackwell

30 Andrews, G. and Halford, G.S. (2002) A complexity metric applied to cognitive development. *Cognit. Psychol.* 45, 153–219

31 Andrews, G. *et al.* (2003) Theory of mind and relational complexity. *Child Dev.* 74, 1476–1499

32 Halford, G.S. *et al.* (2002) Integration of category induction and hierarchical classification: one paradigm at two levels of complexity. *J. Cogn. Dev.* 3, 143–177

33 Baddeley, A. (2000) The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423

34 Oberauer, K. (2002) Access to information in working memory: exploring the focus of attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 411–421

35 Acuna, B.D. *et al.* (2002) Frontal and parietal lobe activation during transitive inference in humans. *Cereb. Cortex* 12, 1312–1321

36 Waltz, J.A. *et al.* (1999) A system for relational reasoning in human prefrontal cortex. *Psychol. Sci.* 10, 119–125

37 Christoff, K. *et al.* (2001) Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *Neuroimage* 14, 1136–1149

38 Bunge, S.A. *et al.* (2005) Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration mechanisms. *Cereb. Cortex* 15, 239–249

39 Christoff, K. and Owen, A.M. (2006) Improving reverse neuroimaging inference: cognitive domain versus cognitive complexity. *Trends Cogn. Sci.* 10, 352–353

40 Kroger, J.K. *et al.* (2002) Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: a parametric study of relational complexity. *Cereb. Cortex* 12, 477–485

41 Birney, D.P. and Halford, G.S. (2002) Cognitive complexity of suppositional reasoning: an application of the relational complexity metric to the knight-knave task. *Thinking and Reasoning,* 109–134

42 Halford, G.S. *et al.* Problem decomposability as a factor in complexity of the dimensional change card sort task. *Cogn. Dev.* (in press)

43 Halford, G.S. *et al.* (2005) How many variables can humans process? *Psychol. Sci.* 16, 70–76

44 Lepine, R. *et al.* (2005) What makes working memory spans so predictive of high-level cognition? *Psychon. Bull. Rev.* 12, 165–170

45 Luck, S.J. and Vogel, E.K. (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281

46 Broadbent, D.E. (ed.) (1975) *The Magic Number Seven After Fifteen Years*, John Wiley & Sons

47 Bunting, M.F. *et al.* (2006) How does running memory span work? *Q. J. Exp. Psychol. (Colchester)* 59, 1691–1700

48 Graesser, A., II and Mandler, G. (1978) Limited processing capacity constrains the storage of unrelated sets of words and retrieval from natural categories. *J. Exp. Psychol. Hum Learn* 4, 86–100

49 McElree, B. (2001) Working memory and focal attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 817–835

50 Cowan, N. *et al.* (2005) Capacity limits in list item recognition: evidence from proactive interference. *Memory* 13, 293–299

51 Yantis, S. (1992) Multielement visual tracking: attention and perceptual organization. *Cognit. Psychol.* 24, 295–340

52 Gobet, F. *et al.* (2001) Chunking mechanisms in human learning. *Trends Cogn. Sci.* 5, 236–243

53 Chalmers, K.A. and Halford, G.S. (2003) Young children's understanding of oddity: reducing complexity by simple oddity and 'most different' strategies. *Cogn. Dev.* 18, 1–23

54 Lewandowsky, S. *et al.* (2004) Time does not cause forgetting in short-term serial recall. *Psychon. Bull. Rev.* 11, 771–790

55 Davis, G. *et al.* (2001) Can attention select only a fixed number of objects at a time? *Perception* 30, 1227–1248

56 Halford, G.S. and Andrews, G. (2006) Reasoning and problem solving. In *Handbook of Child Psychology: Volume 2, Cognitive, Language and*

Perceptual Development (Kuhn, D. and Siegler, R., eds), pp. 557–608, Wiley

57 Hirst, W. *et al.* (1980) Dividing attention without alternation or automaticity. *J. Exp. Psychol. Gen.* 109, 98–117

58 Chi, M.T.H. (1978) Knowledge structures and memory development. In *Children's Thinking: What Develops?* (Siegler, R.S., ed.), pp. 73–96, Erlbaum

59 Halford, G.S. and Andrews, G. (2004) The development of deductive reasoning: how important is complexity? *Thinking and Reasoning* 10, 123–145

60 Nosofsky, R.M. *et al.* (2005) Procedural interference in perceptual classification: implicit learning or cognitive complexity? *Mem. Cognit.* 33, 1256–1271