THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 1996, *49A* (1), 5–28

# Exploring the Central Executive

Alan Baddeley

*MRC Applied Psychology Unit, Cambridge, U.K.*

The central executive component of working memory is a poorly specified and very powerful system that could be criticized as little more than a homunculus. A research strategy is outlined that attempts to specify and analyse its component functions and is illustrated with four lines of research. The first concerns the study of the capacity to coordinate performance on two separate tasks. A second involves the capacity to switch retrieval strategies as reflected in random generation. The capacity to attend selectively to one stimulus and inhibit the disrupting effect of others comprises the third line of research, and the fourth involves the capacity to hold and manipulate information in long-term memory, as reflected in measures of working memory span. It is suggested that this multi-faceted approach is a fruitful one that leaves open the question of whether it will ultimately prove more appropriate to regard the executive as a unified system with multiple functions, or simply as an agglomeration of independent though interacting control processes. In the meantime, it seems useful to continue to use the concept of a central executive as a reminder of the crucially important control functions of working memory.

In the 20 years since its publication, the model of working memory proposed by Baddeley and Hitch (1974) has continued to be useful in stimulating further research, although the three subcomponents have differed markedly in both the amount of further research evoked and also in their apparent success in accounting for available results. The most extensively explored and probably the simplest is the phonological loop (see Gathercole & Baddeley, 1994 for a review). The visuo-spatial sketchpad has proved less tractable, but it has continued to see steady progress through work that is often linked into the related question of factors underpinning visual imagery (for a recent review see Logie, 1995). Meanwhile, the central executive component of working memory remains the least studied, despite the fact that it is almost certainly the most important component in terms of its general impact on cognition.

In the early years of the model, the neglect of the central executive was intentional, as it seemed better to concentrate efforts on the more tractable problems of the two slave

Requests for reprints should be sent to A.D. Baddeley, who is now at the Department of Psychology, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, U.K. Email: Alan.Baddeley@Bristol.ac.uk

I am grateful to Pat Rabbitt and to an anonymous referee for their extensive and constructive comments on an earlier version of this paper and to the U.S. Air Force Grant AFOSR-90-0343 for financial support.

systems. By the mid 1980s, however, the degree of neglect had become an embarrassment to the model, and the attempt began to redress the balance (Baddeley, 1986). The intriguing problem of working out the functional and possible evolutionary significance of the phonological loop provided a potent distractor, but in recent years, I believe, we have begun to make progress. The evidence for such a claim is described below.

## STRATEGIES FOR ANALYSING THE CENTRAL EXECUTIVE

### The Executive as Ragbag

It is probably true to say that our initial specification of the central executive was so vague as to serve as little more than a ragbag into which could be stuffed all the complex strategy selection, planning, and retrieval checking that clearly goes on when subjects perform even the apparently simple digit span task. This still seems a sensible way of starting to explore working memory, as it accepts the complexities and the ultimate need to explain them, while concentrating on analysing the simpler and presumably more tractable slave systems. The Baddeley (1986) version of the model finally began the attempt to specify the central executive in greater detail, relying heavily on the Supervisory Activating System (SAS) component of Norman and Shallice's (1980) model of attentional control. This model has the advantage that it relates working memory to Shallice's study of frontal lobe patients (Shallice, 1982) and to Norman's concern with slips of action. However, the working memory model was still open to the objection that the central executive was just a convenient homunculus—a little man who sits in the head and in some mysterious way makes the important decisions.

### The Central Executive and the Frontal Lobes

In recent years there have been at least two dominant approaches to attempting to understand the processes underlying executive control—one principally stemming from neuropsychology, the other being rooted in the psychometric tradition—and a number of attempts have been made to combine the two. In the area of neuropsychology, there is abundant evidence that disorders of executive control are associated with damage to the frontal lobes (Shallice, 1982, 1988). One approach to understanding executive processes, therefore, is to attempt to study the functions of the frontal lobes (e.g. Duncan, 1986; Duncan, Johnson, Swales, & Freer, in preparation; Shallice & Burgess, 1991, 1993). We could therefore define the central executive anatomically, as that system that resides in the frontal lobes.

However, although I regard the neuropsychological evidence as highly relevant, this does not seem to me to be the most fruitful line for the working memory model to take. The model is principally a functional model that would exist and be useful even if there proved to be no simple mapping on to underlying neuroanatomy. To add a component that was defined neuroanatomically rather than functionally would therefore be anomalous. It would also be likely to be unhelpful, bearing in mind the fact that the frontal lobes constitute an extremely large area that almost certainly has multiple functions that are as yet poorly understood. It is entirely possible that, although the frontal lobes are often involved in many executive processes, other parts of the brain may also be involved in executive control. If we identify the central executive exclusively with frontal function, then we might well find ourselves excluding from the central executive processes that are clearly executive in nature, simply because they prove not to be frontally located. Equally, we would be in danger of describing functions as executive simply because they were based on the frontal lobes. Note that this is not an argument against exploring the role of the frontal lobes, merely against defining the functional concept of a central executive in terms of a specific anatomical location.

Extending this argument, Baddeley and Wilson (1988) suggested that it was desirable to extend this proposed separation of functional and anatomical concepts into neuropsychology. In the case of most neuropsychological functions, this separation already occurs: a patient is referred to as suffering from an *amnesic syndrome* rather than a *temporal lobe* or *hippocampal syndrome*, and it would be generally agreed that it is more profitable to talk about a patient's *dyslexia*, or *acalculia* rather than attempt to specify the dysfunction purely in terms of hypothetical anatomical locus of damage. For that reason, we suggested that the commonly used concept of a *frontal syndrome* should be replaced by the functional concept of a *dysexecutive syndrome*. The anatomical substrate of the dysexecutive syndrome represents an important question, but it does not form a component of its definition. In what follows, I shall make extensive use of neuropsychological evidence, much of it from patients who have damage to the frontal lobes, but I do not propose to use anatomical localization as a *defining* criterion for the central executive.

## Psychometric Approaches to the Central Executive

A second and related approach to the analysis of executive processes has been through the study of individual differences. This approach is reflected in two separate but related research streams, one based on the traditional concept of intelligence, and the other more directly influenced by the assumption made by Baddeley and Hitch (1974) that working memory involves the simultaneous storage and manipulation of material. This latter approach, which has been particularly influential in North America, is considered later. The more traditional psychometric approach has been based upon the assumption that intelligence measures reflect the operation of a central cognitive processor, which could potentially be identified with the central executive of working memory. This, in turn, leads to the question of whether intelligence is better considered as reflecting a single general factor or capacity, for example Spearman's *g*, or whether, as Thurstone (1938) proposed, *g* can usefully be broken down into a number of subprocesses. This general issue continues to be pursued actively using populations of normal subjects (Kyllonen & Christal, 1990), of neuropsychological patients (Burgess & Shallice, in press; Della Sala, Gray, Spinnler, & Trivelli, in preparation; Duncan, 1986), and of normal elderly subjects (Rabbitt, 1983; Salthouse, 1991). The results of such studies are clearly relevant to a concept such as working memory, but they depend crucially upon exactly which tests are included in the study, and what subject groups are tested. Although there is some encouragement for the clustering of tests, this is by no means always so. Furthermore, the clusters obtained do not at present show any consistent pattern. Some neuropsychological studies suggest a clustering of classic "frontal" tasks (Della Sala et al., in preparation), but others do not (e.g. Burgess & Shallice, in press; Duncan et al., in preparation). In the case of studies using normal subjects, the pattern is again unclear, as reflected in the present volume: Lehto finds correlations between some aspects of "frontal" tasks and some working memory tasks, and Waters and Caplan obtain patterns of correlations that appear to be largely specific to type of material and method of processing.

## The Homunculus: Friend or Foe?

Given that we adopt neither the anatomical nor the classic psychometric approach, what other methods are open? One possibility is to accept the homunculus with all its limitations, but to argue that such a concept is not only useful in defining the scope of our attempts to understand the subsidiary slave systems, but may also be productive, provided we systematically attempt to analyse the functions performed by the homunculus. If we can separate and explain just one part of the role currently attributed to the homunculus, then we shall have made progress. If we can define and analyse a range of several executive processes, then we would be in a position to begin to ask whether they are better regarded as individual and separable functions, or whether a unitary account could be given.

Using this gradualist approach, we can hope eventually to account for many—and indeed one hopes all—of the executive functions that at present are performed by our rather unsatisfactory homunculus. We would then be in a position to use psychometric methods to ask whether these are better regarded as reflecting the various operations of a single unitary controller, or whether they might be better regarded as an executive committee of interacting but independent administrators. This is clearly an ambitious and lengthy enterprise, which may prove ultimately unsuccessful. Even if this proves to be the case, however, the attempt to fractionate and understand a range of executive processes is likely to prove fruitful in raising new questions and providing new findings that will need to be accommodated by any adequate model of executive control. The sections that follow give an account of the progress we have so far made along this particular path.

Although presented as a formal strategy, in actual practice it has evolved gradually and pragmatically, starting with an attempt to capture just one necessary function of a central executive—namely, that of co-ordinating the two proposed WM slave systems. This component of the programme was initially stimulated by the attempt to test the hypothesis that patients suffering from Alzheimer's Disease showed particular impairment in the operation of the central executive, but it has more recently been extended to a broader range of patients, notably including those with frontal lobe lesions.

A second strand of research stemmed from the attempt to link the central executive to the SAS system postulated by Norman and Shallice, by means of the task of random generation (Baddeley, 1986). A third component of our approach is much more recent, being based on the assumption that one important role of the central executive should be to act as an attentional controller, selecting certain streams of incoming information and rejecting others. We have just begun to investigate this aspect of the executive's function, concentrating initially on the effects of ageing.

The final executive capacity to be considered is the ability to select and manipulate information in long-term memory. Although this is the least developed aspect of our own work, it has formed an important component of the work on individual differences in working memory that have been studied intensively elsewhere, notably by Carpenter, Just, and their colleagues and by Engle and his associates. Having reviewed progress on these four approaches, I will return to the question of whether this strategy seems sufficiently fruitful to merit continuation. Should we sack the homunculus, or continue in our long-term aim of whittling away his functions until he eventually becomes unnecessary, and can gracefully retire?

## APPROACHES TO FRACTIONATING THE EXECUTIVE

### Dual-Task Performance

Our first attempt to devise a measure of executive function stemmed from research on the memory deficit accompanying Alzheimer's Disease. A preliminary study indicated the expected substantial deficit in episodic long-term memory, but also suggested a degree of impairment extending across aspects of working memory, both verbal and visuo-spatial (Spinnler, Della Sala, Bandera, & Baddeley, 1988). We hypothesized that this might reflect the common central executive contribution to the two slave systems of working memory and attempted to devise a series of tasks that would allow us to test this more directly. One necessary feature of the model is its capacity to co-ordinate information from the two slave systems. We therefore set up a number of tests in which Alzheimer's Disease (AD) patients, normal elderly control and young control subjects were tested on individual tasks reflecting the operation of the relevant slave systems and were then required to combine performance on two such tasks (Baddeley, Logie, Bressi, Della Sala, & Spinnler, 1986).

For this first series of studies, we used pursuit tracking as the visuospatial task. The subjects were required to keep a light pen in contact with a moving spot of light on a VDU. We varied rate of movement of the spot so as to ensure that all subjects were performing at an equivalent level, 70% time on target. This was combined with each of three other tasks. The first involved articulatory suppression, the second reaction time to a tone, and the third digit span. Length of digit sequence was dependent on each subject's digit span and hence varied between groups, ensuring that error level was equivalent across subjects when the memory task was performed alone.

The results of this study indicated that articulatory suppression did not significantly impair performance in any of the three groups, although an earlier study by Morris (1986) and a later follow-up study observed that patients were significantly impaired, whereas the elderly and young controls were unaffected by suppression. When combined with the concurrent RT task, tracking in the AD patients showed a significantly greater decrement than was found in either of the two control groups. This was not simply a trade-off effect, as performance on both speed and accuracy of the reaction time task showed a similar pattern of significantly greater impairment in the AD patients. Similar results were also shown for the concurrent digit span task, with both tracking and memory performance showing a significantly greater decrement in the AD patients than in either of the control groups, who did not differ. Our results were therefore consistent with the hypothesis that the capacity to combine performance on two tasks—a capacity that, we have argued, is a necessary function of the central executive—is particularly impaired in AD patients. However, although our results are certainly consistent with this view, other interpretations need to be considered.

One possibility that has been suggested is that our results might simply reflect impaired performance on the constituent peripheral tasks rather than the cost of their co-ordination. We think this is unlikely, for two reasons: (1) We carefully titrated the level of difficulty in the case of our digit span and tracking tasks to a point at which all groups were performing at the same level. It is of course still possible to argue that achieving this level demanded more attention from the AD patients, but there is no reason to assume that any of the groups were doing anything other than devote their full available resources to the task. A second reason for rejecting the specific peripheral deficit interpretation is that the dual task deficit in AD patients found in this and other studies using other combined tasks (Greene, Baddeley, & Hodges, in press), is not found in the case of normal ageing. Despite the fact that age tends to impair performance on the constituent tasks, provided level of performance on the tasks is age-adjusted, combining them does not lead to an enhanced effect of age in this or other studies (Salthouse, 1991). As some aspects of WM do appear to decline with age (Salthouse, 1992; Welford, 1958), this pattern of results supports the view that the executive can be fractionated (see Gick, Craik, & Morris, 1988; Morris, Gick, & Craik, 1988).

A second potential explanation of our results is to suggest that the decrement observed in AD patients reflects some overall deficit, such as one in general intelligence or *g*. Rabbitt (1983) has suggested, for example, that much of the cognitive deficit observed in ageing is attributable to reduced general intelligence. I am unhappy with this interpretation, for two reasons. First of all, if both AD and ageing represent a simple reduction in *g*, which is reflected in the capacity to combine tasks, then one would expect dual task performance to be sensitive to age. As suggested above, this is not the case.

The second problem with a concept such as intelligence is that it replaces one problem, the nature of the executive, with another, the nature of intelligence. Most established tests of intelligence, such as the WAIS, Heim's AH4 or Cattell's Culture Fair Test, are based on the performance of a range of subtasks, each of which probably reflects a number of different processes. Whereas it might well be informative to study the contribution of simpler and potentially purer measures to performance on intelligence tests, it is far from clear that a general concept such as intelligence can throw useful light on either the functioning of the central executive or, indeed, on the neuropsychology of Alzheimer's Disease.

A similar problem of lack of specificity occurs in the case of an account of our results in terms of task difficulty. Dual-task performance clearly *is* difficult for AD patients; the role of the central executive hypothesis is to predict *what* will be difficult, rather than simply using the general concept of difficulty to label the tasks on which AD patients fail. Such an argument gains strength from a second experiment, wherein we varied difficulty within a task without increasing the demand for dual-task coordination (Baddeley, Bressi, Della Sala, Logie, & Spinnler, 1991).

The task in question was that of classifying words as belonging to one or more semantic categories. Earlier work by Yntema and Mueser (1960) had shown that the time to categorize a word increases as a function of the number of potential categories from which it is selected. We therefore presented the category judgement task to our patients and controls, studying the influence of number of simultaneously presented categories on time to decide whether a presented word was or was not a category member. We tested performance both immediately and after a six-month interval during which the disease had progressed.

As expected, both patient and control groups took longer with larger numbers of categories, and control subjects were consistently faster and more accurate than were AD patients, whose performance declined as the disease progressed. There was, however, no interaction between difficulty and disease stage, hence arguing against the hypothesis that simply increasing difficulty will make a task more sensitive to the progress of the disease. This contrasted with dual-task performance, where combined performance deteriorated substantially more rapidly than did performance on the constituent tasks when performed alone. It could, of course, be argued that our categorization task had varied the wrong kind of difficulty. However, such an objection simply highlights the inadequacy of a concept as poorly specified as "difficulty". We have proposed a rather specific executive process, which we have shown to be more disrupted in AD patients than either performance on its constituent parts or an alternative task that varied in difficulty. This is clearly not sufficient to settle the issue, but was enough to encourage us to investigate dual-task performance more extensively in both AD patients and in other groups of patients.

Our task seemed to be a useful and interesting one, but it was logistically far from convenient. As a standard piece of laboratory equipment, the pursuit rotor is almost extinct, and although it is possible to develop computerized tracking tasks, we found that lack of standardization was a major problem in purchasing equipment in different countries, even when purchasing machines of apparently identical specification. We therefore decided to develop a paper-and-pencil version of the test, and we began a series of pilot studies in search of a simple paper-and-pencil tracking task that could be scored easily and explained readily to AD patients. After a range of unsuccessful attempts, we ended up with a task in which a chain of 0.5-cm-square boxes is laid out to form an irregular path on a sheet of paper. The subject's task is to start at one end of the chain of boxes, placing a cross in each box in turn and working as rapidly as possible until the chain is complete. A second sheet is then presented, containing rather more boxes, and the subject is asked to fill in the chain of boxes as rapidly as possible for a period of 2 min. After establishing the subject's digit span using standard procedures, sequences of digits at that length are presented and tested continuously for a 2-min period, noting how many are completed and how many correct. The final stage involves combining box-crossing and digit span, again for a 2-min period (Della Sala, Baddeley, Papagno, & Spinnler, in press).

We validated the new method using a sample of 13 AD patients and 12 controls. The results were clear, with all but one AD patient showing clear impairment under dual-task conditions, whereas none of the controls showed a marked decrement. The one exception subsequently proved to have been initially misdiagnosed, proving not to have a progressive dementia, and was hence excluded from the data, which then showed a clear separation, with no overlap between the two groups in susceptibility to dual-task interference.

Although we were pleased that our modification was successful, it still left us with many theoretical questions, one being that of the neuroanatomical basis of performance on this executive task, and in particular its possible link with frontal-lobe function. We therefore gave the task to a sample of 24 neurological patients who were free of AD, but who all showed clear neuroradiological evidence of lesions somewhere in the frontal lobes. All 24 patients were assessed in three ways. Two of these involved standard cognitive measures that are generally assumed to be disrupted following frontal-lobe damage. The first of these is verbal fluency in which subjects attempt to generate as many words as possible in 90 sec, from a specified category such as animals. Poor performance on this task is typically associated with left frontal damage (Perret, 1974), and the task itself is typically assumed to be associated with the executive process of search and retrieval from long-term memory (Baddeley, Lewis, Eldridge, & Thomson, 1984; Engle, in press). The second classic frontal measure was the Wisconsin Card Sorting Test (Nelson, 1976), in which subjects are required to form concepts as to ways in which cards containing a range of features may be sorted. Once the initial concept (e.g. colour) has been achieved, the experimenter switches to another concept (e.g. number), until eventually the six possible conceptualizations have all been tested. Performance is measured in terms of number of concepts achieved and number of errors, with perseverative errors being particularly characteristic of frontal-lobe damage. Although no task is uniformly and unambiguously associated with frontal lobe damage, these are probably the most frequent and generally accepted measures in the field (Lezak, 1983).

Our third measure attempted to capture the characteristically disinhibited behaviour that often accompanies frontal-lobe function and is typically the aspect of behaviour that makes some frontal-lobe patients fail to cope independently, even when many aspects of cognition that are frequently disturbed following frontal-lobe damage are preserved (see Shallice & Burgess, 1993, for a discussion). The behavioural syndrome is readily recognizable, although difficult to characterize in detail. Rylander (1939, page 20) suggested that it involves "disturbed attention, increased distractibility, a difficulty in grasping the whole of the complicated state of affairs . . . while able to work along old routine lines . . . but cannot learn to master new types of task". Such patients are often inclined to facetiousness, tend to perseverate in conversation, and may show signs of confabulation. In order to identify this aspect of cognitive dysfunction, we used clinical judgement of two independent assessors, both experienced neurologists and neuropsychologists. One judge categorized the patients as showing dysexecutive behaviour or not on the basis of their performance during the test session; the second independently based his judgement on the existing patient notes. The two assessments agreed for all except two of the 24 patients, and these two were dropped from further analysis.

Normative data for the Wisconsin Card Sorting Task and verbal fluency indicate that our group were significantly impaired, though to varying degree. The assessment of behavioural dysfunction was based on a simple dichotomy, which resulted in approximately equal numbers of patients in each group. There was a clear tendency for patients showing behavioural disorder to demonstrate a clear decrement in performance when the box-crossing and digit span tasks were combined, an effect that showed up particularly strongly on the memory component of the test, with behaviourally disturbed subjects showing a drop from 88% to 65% correct sequences on adding tracking to span, whereas the behaviourally undisturbed group showed no significant decrement (88% to 84%). The behaviourally normal and behaviourally disordered groups did not, however, differ significantly either in verbal fluency or Wisconsin Card Sorting performance.

This set of findings clearly requires replication and extension, but taken at face value it appears to be most consistent with a view of the central executive that involves a number of subcomponents, possibly associated with the functioning of different aspects of the frontal lobes. More specifically, it implies that the disordered and disinhibited behaviour that is sometimes found in frontal lobe patients is associated with difficulty in distributing attention.

How should this pattern of results be interpreted? One option might be to conclude that, because of its association with behavioural manifestations of the dysexecutive syndrome, our dual task is a better measure of executive function than the more established tests such as verbal fluency. We do however, have other reasons for regarding verbal fluency, for example, as a good measure of at least some aspects of the central executive. We found that it was particularly sensitive to the effects of a concurrent heavy digit load (Baddeley et al., 1984), and recent work by Engle (in press) finds that verbal fluency performance relates closely to working memory span measures of the type devised by Daneman and Carpenter (1980). A much more plausible interpretation would seem to be that the two tests simply measure different executive processes, with the two processes being differentially associated with the behavioural problems commonly observed in dysexecutive patients. For example, impaired verbal fluency may be associated with retrieval problems, resulting in disruption in autobiographical memory, either associated with extreme poverty of recollection, as in so-called "dynamic aphasic" patients who have great difficulty in initiating retrieval, or in the apparently opposite pattern also found in dysexecutive cases, where recollection is fluent but inaccurate, resulting in confabulation (Baddeley & Wilson, 1986).

The association between impaired dual-task performance and behavioural disturbance was not predicted in advance of our study. It is, however, reminiscent of results reported by Alderman (in preparation), who studied a group of severely brain-damaged patients who were subjected to a rehabilitation regime based on a token economy system in which socially acceptable behaviour earned points that could subsequently be cashed to purchase small luxury items. The system is usually highly successful in helping seriously behaviourally disturbed patients adapt to living with others. A small subgroup of patients, however, fail to respond to the treatment, and Alderman was concerned to understand what characterized these atypical patients. He tested all patients on a wide range of neuropsychological tests, including most of the standard tests of "frontal" function, together with a number of variants on our dual-task performance test. The standard "frontal" tests showed a weak association with behavioural disturbance, and performance on each of the range of dual-task tests was strongly associated with failure to benefit from the token economy.

The evidence from two sources of a strong association between poor dual-task performance and behavioural problems, therefore, seems to merit further investigation. It is unclear what might be the mediating mechanism, but one possibility is that effective social behaviour requires simultaneous monitoring of one's own interests and desires while attending to the potentially conflicting concerns of those around, a form of dual-task or indeed multi-task activity. However, before proceeding further along such speculative lines, it is necessary to replicate these results, preferably in conjunction with more careful and quantitative monitoring of the behaviour of the dysexecutive and control patients.

Meanwhile dual-task performance is beginning to be used somewhat more widely. Dalrymple-Alford and his associates have found a mild degree of impairment on dual-task performance in patients suffering from Parkinson's disease (Dalrymple-Alford, Kalders, Jones, & Watson, 1994). A study by Greene, Baddeley, and Hodges (in press) also observed impairment in mild AD patients, although the pattern of deficit is somewhat different from that observed in the previously described AD study, in that Greene et al.'s patients decreased their performance on box-crossing, rather than showing a decrement based on increased error rate on the memory test. This raises an important methodological problem that concerns any dual-task performance measure—namely, that of how to combine the two components so as to give a single score. Such a score is necessary if one is to compare the performance of subjects who may have a somewhat different trade-off between performance on the two components of the test. This problem, which is discussed in greater detail by Baddeley et al. (in press), will clearly have to be tackled if the test is to become suitable for general clinical use.

One final point to be noted from these more recent studies is the evidence for dual-task performance as a capacity that extends across a range of different tasks. In the original AD studies, broadly similar results were obtained whether performance on a pursuit rotor was combined with articulatory suppression, reaction time to a tone, or digit span, and our later studies indicated that changing the tracking task to one involving writing crosses in boxes did not diminish the sensitivity of dual-task performance to AD. Alderman also obtained similar results from a range of combinations of tasks, and the study by Greene et al. (in press) observed a correlation of 0.49 between performance on the combined box-crossing and digit span task and the dual-task component of the Test of Everyday Attention (Robertson, Ward, & Ridgeway, 1994). In this subtest, subjects search a list resembling a page of the telephone directory, at the same time as listening for a specified message. Again characteristically, correlation of these two tests with other tests of executive function such as verbal fluency was significant but low ($r$ = 0.33 and 0.31 for the box-crossing and search tasks, respectively).

In conclusion, the capacity to carry out two tasks simultaneously appears to be a candidate for one separable feature of executive function. However, it would be premature to regard our results as conclusive, bearing in mind problems of allowing for possible differences in speed–error trade-off and of the statistical scaling problems that are inevitably found when comparing impaired and non-impaired populations. We do, however, regard progress as sufficiently encouraging to justify further attempts to tackle these thorny methodological problems.

## Random Generation

The second area of investigation of executive function was prompted by the problem of explaining a set of results published many years earlier, in which subjects had been encouraged to generate sequences of letters, making the order as random as possible (Baddeley, 1966). The results were highly consistent and broadly fitted the conceptualization that the process of random generation depends on a system of limited informational capacity—hence the more rapid the rate, the less random the output, and the larger the set of selection alternatives, the slower the maximum generation rate. The proposed limited-capacity system appeared to have features in common with the system that limited performance in reaction-time studies, as when random generation was combined with a card-sorting task paced at a rate of 2 sec per response, the greater the number of response alternatives, the less random and more stereotyped was the concurrent sequence of letters generated.

Despite its lawfulness, this pattern of results remained difficult to explain until the arrival of the Norman and Shallice (1980) model. It will be recalled that this involves two sources of control of action, schemata that channel behaviour into well-learnt habitual patterns, together with the Supervisory Attentional System, an attentional controller that is capable of overriding habitual response patterns when it is necessary to initiate new behaviour. Looked at from this viewpoint, random generation could be seen as reflecting a series of habitual letter-retrieval schemata that were based on processes such as alphabetic recitation or the production of common acronyms. The requirement to make the sequence random, however, demanded the constant intervention of the SAS in order to break up these stereotyped sequences. As the SAS was presumably also required for the decision process involved in sorting cards into different categories, card sorting interfered with the randomness of the letter sequence generated (Baddeley, 1986). This initial suggestion has led to the adoption of random generation as a secondary task that might be assumed to disrupt the operation of the central executive. Disruption of a range of tasks, from the learning of simple contingencies (Dienes, Broadbent, & Berry, 1991) through the disruption of strategic thought in chess (Robbins et al., in press), suggests that this assumption is a reasonable one. However, although there have over the years been many speculations as to what underpins the task of random generation, there was until recently little attempt to relate the task more clearly to the functioning of a hypothetical executive component of working memory. An opportunity to remedy this came with a joint grant with Duncan to investigate executive processes.

An important concept in the study of attention over the last 20 years has been that of automaticity. Shiffrin and Schneider (1977) argued that the repeated pairing of a specific stimulus with the same response will gradually reduce the attentional demand of responding to the stimulus, up to a point at which the stimulus will evoke the response virtually automatically. Random generation could be regarded as the opposite end of this continuum from automaticity, the aim being to generate a response that is minimally associated with what went before, hence producing a task that continues to demand attention even after much practice. The pattern of results produced by Baddeley (1966) is consistent with the idea that a common mechanism is involved in studies of reaction time and of random generation but leaves the underlying mechanism unexplained. We decided to attempt a more detailed analysis, which, we hoped, would lead to an explanation.

The vast majority of studies of random generation have used verbal output, typically involving letters or numbers. This is logistically somewhat cumbersome, because it requires copying down the subject's output and then keying these into a computer for analysis of randomness. Number and letter generation also lend themselves to a range of simplifying strategies, reciting telephone numbers or dates, spelling words, or providing the initial letter of names of objects around the room. Furthermore, having only a verbal output mode limits the use of random generation as a secondary task. We therefore decided to explore the possibility of using random keypressing as an alternative generation procedure (Baddeley, Emslie, Kolodny & Duncan, in preparation).

Our first study required subjects to generate random sequences of a hundred numbers at each of three rates—0.5, 1, or 2 sec per response. One condition required subjects to generate random sequences of spoken digits. In a second condition, subjects were seated at a keyboard containing 10 keys, one located under each finger or thumb. Subjects were required to generate a random sequence of presses, again at each of the three rates. Results indicated that the degree of randomness was somewhat less for keypressing than for digits but showed an equivalent decline as generation speed increased. We also observed characteristic stereotyped response sequences, which, as in the case of alphabetic stereotypes, became more frequent as rate of generation increased. In short, keypressing appeared to be broadly equivalent to digit generation. From this point on we used keyboard generation almost exclusively, typically selecting a rate of one response per second.

Our next study explored further the assumption that random generation reflects the limited capacity of a general purpose executive system. We combined keypressing with performance on a memory span task in which subjects recalled sequences ranging in length from one item to eight. If performance depends upon a general purpose system, then there should be interference between the verbal memory task and the visuo-spatial generation task. Furthermore, if the system reflects a limited-capacity working memory, then the degree of disruption of random generation should increase with concurrent memory load. As Figure 1 shows, this is what we observed.

We went on to investigate the influence on keyboard generation of a range of further tasks, selected on the basis of their expected loading on the central executive component of working memory. We found that articulatory suppression—for example, counting repeatedly from 1 to 6—had no significant effect on random generation, which was, however, substantially disrupted by a category generation task in which subjects attempted to produce as many items as possible from a specified semantic category such as *animals* or *fruit*. Such verbal fluency tasks do seem to depend relatively heavily on executive resources, as evidenced both by their susceptibility to concurrent digit span (Baddeley et al., 1984), and to impairment in patients suffering from the dysexecutive syndrome (Baddeley & Wilson, 1988). An even greater degree of impairment was produced by a concurrent requirement to perform the AH3, a demanding measure of fluid intelligence (Heim, 1975). Duncan (1993) has argued that performance on such intelligence tests is an index of executive function and depends upon the operation of the frontal lobes. The overall pattern of results, therefore, was broadly consistent with the

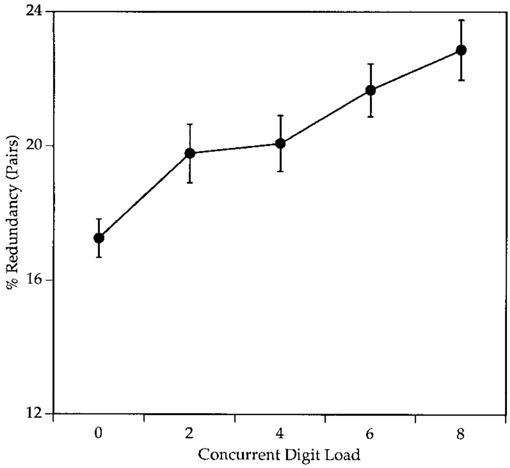


FIG. 1. The effect of concurrent digit-span load on the randomness of key pressing.

assumption that random generation competes for the same limited capacity as is necessary for performance of a range of tasks that depend to a greater or lesser extent on central executive functioning.

However, although this pattern of results was broadly supportive of the initial hypothesis, none of the results so far described places any major constraints upon the possible models of the underlying process. An exception to this was provided by one condition in which we asked our subjects to combine random number generation with random key-pressing, both being required at a 1/sec rate. What might one expect? In comparison to performing either of the generation tasks alone, by combining them we were asking for double the amount of information to be generated. Given that we are moving from a single-task to a dual-task mode, this seemed likely to create even more problems, and we suspected that subjects would simply be unable to perform the task. In fact they coped very well; concurrent digit generation reduced the randomness of keypressing by about the same amount as concurrent category generation or holding a sequence of digits of span length. The reciprocal effect of concurrent keypressing on digit generation was rather less.

In order to explain these results, we had to think in greater detail about the possible processes underlying random generation, and we came up with a model that is somewhat similar to a simplified version of retrieval in Raaijmakers and Shiffrin's (1981) SAM model, as represented, not entirely sympathetically, by Roediger (1993). It involves setting up a retrieval plan, running it, and checking the output, which, if judged to be suitably random, is then emitted at the appropriate time. We assumed that the decrease in randomness at higher speeds occurs principally because of the time taken to shift from one retrieval plan to another. If this were not time-limited, then the subject could presumably switch every time and would not need to bother checking the randomness of the output. On the other hand, if the same retrieval plan is used repeatedly, then the stream of responses is likely to be stereotyped and non-random. Anything that interferes with the capacity to switch retrieval plans will tend to increase the degree of redundancy.

Note that although such a system may be quite sensitive to concurrent activity, degradation is gradual rather than catastrophic; the subject simply makes fewer switches of retrieval plan, resulting in more stereotyped responses. In the case of combined keyboard and digit generation, if the performance limitation reflects the switching process, then the result will simply be that generation will continue, but with fewer switches of retrieval plan. Provided the retrieval plan for keypressing is separate from that for digit generation and can operate simultaneously, the outcome will be a general reduction in randomness rather than catastrophic breakdown.

Our final two experiments attempted to test the switching hypothesis directly. If the need to switch retrieval plans is the source of the disruption in performance, then it should be possible to devise a task that places minimal load on memory and other executive processes but has heavy switching demands. We were encouraged in our search by evidence from neuropsychology, where the Trails Test from the Halstead–Reitan battery is regarded as sensitive to frontal-lobe damage. The test measures speed of performance on a task in which subjects use a pencil to join together a series of numbered squares (Trails A). A second subtest (Trails B) requires subjects to alternate between numbered and lettered squares, to produce a sequence that connects squares A–1–B–2–C–3, etc. Patients with frontal-lobe damage are particularly disrupted on this alternating test (Lezak, 1983).

As our subjects were pressing keys, it would obviously not be appropriate to give them the Trails Test, so we developed our own verbal equivalent. In the first study, subjects pressed keys at a rate of 1/sec, either alone or in combination with one of three concurrent tasks, involving either reciting the alphabet at a 1/sec rate, counting at the same rate, or alternating letters and numbers (A–1–B–2–C–3, etc.). Whereas neither counting nor reciting the alphabet had a detectable effect on the randomness of keypressing, the concurrent alternation task markedly reduced randomness. We were therefore encouraged to carry out a replication, this time starting the subjects on each trial with a different initial number–letter pair, for example F–9, and requiring them to repeat and then continue (F–9–G–10–H–11–I–12, etc.). Despite the minimal memory load imposed by this task and its entirely predictable nature, it led to a substantial reduction in redundancy of keypressing and was itself disrupted, as indicated by the number of occasions on which generation ceased and had to be prompted by the experimenter.

We have provisionally concluded therefore that random generation disrupts the operation of the central executive by its demand for the constant switching of retrieval plans. We believe that our simple alternation task performs a similar function and propose to explore it further in studies using both normal subjects and neuropsychological patients.

Although our results are encouraging, we are still some way from an adequate model of random generation. It is, for example, still unclear to what extent the load imposed by generation stems from (1) the need to switch strategies, (2) the problem of accessing new strategies, or (3) the monitoring of the response output. The initial concept of a limited-capacity general processor responsible for both input and retrieval is also clearly oversimplified. Baddeley et al (1984) observed that a concurrent sorting task had a marked effect on the learning of a list of words but little or no effect on retrieval, suggesting a degree of automaticity in the retrieval process. A subsequent series of experiments by Craik (1995) replicated the learning effect but did in addition find a consistent but smaller effect of a secondary RT task on number of words retrieved. However, when the RT task itself was studied, the effect of concurrent word retrieval on mean RT was substantial. RT did not, however, vary as a function of number of items retrieved per unit time, suggesting a general effect rather than a direct effect of item retrieval on checking or production. Furthermore, RT was not affected by whether subjects were instructed to focus principally on retrieval or on RT. Craik suggests that whereas learning depends on amount of attention available, retrieval behaves differently, depending on the cognitive demand of operating in a general retrieval mode. Craik suggested that this difference is linked to data from recent PET studies that suggest differential cortical localization of encoding and retrieval processes in the left and right frontal regions, respectively.

In conclusion, although our understanding of the role of WM in random generation is far from complete, it seems likely that further analysis may have important implications for the more general question of the role of attentional processes in retrieval.

## Selective Attention

The third proposed component of executive processing resulted from speculation as to what further capacities might be likely to be required by a general executive processing system, in addition to the capacity to timeshare and to switch retrieval plans. One obvious candidate is the capacity to attend selectively to one stream of information while discarding others—the classic phenomenon of selective attention. This was investigated in a series of experiments carried out in collaboration with Duncan Godden, Elizabeth Maylor, Ian Robertson, and Tony Ward. Middle-aged and elderly subjects were studied, on the assumption that age is a variable that influences executive processes in ways that are important but far from straightforward.

The suggestion that the elderly may have an impairment in working memory extends back at least as far as Welford (1958), and data from studies of ageing were incorporated in initial discussion of the concept of a central executive in working memory (Baddeley, 1986). Available evidence, however, is somewhat equivocal and appears to depend, in ways that are not fully understood, on the exact nature of the task. For example, studies by Morris et al. (1988) and Gick et al. (1988) used the technique developed by Baddeley and Hitch in which a concurrent digit load accompanies a secondary task—in this, case sentence verification. Working memory load was manipulated in two ways, either by increasing the concurrent load or by increasing the syntactic complexity of the sentence being processed. Both of these impaired performance, but whereas syntactic complexity had a greater effect in the elderly than in the young, there was no such interaction betwen age and concurrent digit load. This inconsistency characterizes much of the literature (see Craik, Anderson, Kerr, & Li, 1995, for a review), suggesting that ageing may be an interesting and productive variable to study within the context of working memory.

One problem in attempting to carry out theoretically driven studies on ageing stems from the fact that almost every physical and cognitive function shows some decline. Consequently, showing that the elderly perform poorly on any given task cannot be regarded as evidence for the task's peculiar vulnerability to ageing, unless other factors are ruled out. This has led to a series of attempts to capture the essence of the ageing deficit in terms of a single measure such as general intelligence (Rabbitt, 1983), speed of processing (Salthouse, 1991), or reduced capacity for inhibition (Hasher & Zacks, 1988). We therefore adopted a strategy of measuring the first two of these and using them as part of a multi-variate analysis that allowed us to ask whether any of our working memory conditions led to age-related deficits that were not attributable to general intelligence or simple speed of processing. If we were to find such additional effects, they would argue in favour of a multi-component executive rather than a monolithic system based on a single underlying capacity.

Our basic paradigm was one in which subjects were instructed to press a key as rapidly as possible whenever a specified stimulus occurred. In the first two of our experiments, subjects were required to count the stimuli and report the total when cued, thereby, we felt, providing a more demanding task than simple reaction time (Wilkins, Shallice & McCarthy, 1987). However, exclusion of counting from the third experiment did not change the pattern of results, so the counting requirement will not be discussed further. We used two manipulations to increase the attentional demand of the task. The first of these was to present irrelevant stimuli, which the subject was instructed to ignore; in Experiment 1 these were always in a different sensory modality, while in Experiments 2 and 3 stimuli within the same modality were also included. Our second attentional manipulation was produced by an occasional instruction, which required the subject to switch from responding to signals in one modality to those in the other. We measured both the overall effect of this and its short-term effect, as Allport, Styles, and Hsieh (1994) have presented evidence suggesting that this initial response is particularly markedly slowed.

The first experiment, therefore, was one in which subjects watched a VDU and responded by pressing a key when a circle appeared. All subjects were tested on four conditions in which visual detection was performed (1) alone, (2) with irrelevant tones, (3) with instruction to respond to *both* circles and tones, and finally (4) with a requirement to switch between circles and tones on a given cue. We tested 24 middle-aged subjects (mean age = 42, range 35 to 50) and 24 elderly subjects (mean age = 72, range 66 to 83), all of whom were members of the APU Subject Panel and all of whom were required to perform a paper-and-pencil test of fluid intelligence, the Cattell Culture Fair Test (Cattell & Cattell, 1960).

The results were disappointingly straightforward. Reaction time was slowed by the presence of an irrelevant signal on the other dimension and by the instruction to switch channels—an effect that, as Allport et al. reported, is particularly marked in the case of the first response after switching. The elderly subjects were consistently slower than the young and, as expected, were lower in performance on the fluid IQ test. When IQ was taken out as a covariate, the age differences disappeared—a result that mirrors closely many similar findings reported by Rabbitt (1983).

Our second experiment utilised the same group of subjects and a design that was similar, apart from the inclusion of conditions in which irrelevant stimuli occurred within the *same* dimension as the target stimuli. Hence, subjects responded to circles but ignored triangles, or responded to low tones but ignored high.

Comparing Experiments 1 and 2 indicated that subjects were slower in responding when they had to ignore irrelevant stimuli, particularly when these occurred within the same sensory dimension. Switching from one modality to another slowed responding, particularly on the initial stimulus on the new modality, and again the elderly responded consistently more slowly than did their middle-aged counterparts. When the analysis was performed using IQ as a covariate in Experiment 2, the age difference disappeared when subjects were ignoring a stimulus on the other dimension or were switching from one dimension to the next. There did, however, remain a significant age effect for the condition in which subjects were required to ignore irrelevant stimuli within the same dimension as the targets.

Bearing in mind the potential theoretical interest of the last finding, we decided to replicate, using auditory stimuli that varied on dimensions other than simple pitch, adding differences in timbre to produce two stimuli that could best be described as a squeak and a grunt. For visual stimuli we used a square and a diamond. In order to simplify the design, we omitted the condition in which subjects switch modalities; this had consistently failed to show a differential sensitivity to age. This is broadly consistent with the findings of Allport et al. (1994), which seem to point to a phenomenon based on negative priming rather than a reflection of some active limited-capacity attentional control mechanism.

As expected, there were overall main effects of age and of whether the irrelevant stimulus was on the same or a different sensory dimension. A multi-variate analysis was then performed in which the relevant factors were age, intelligence, and speed of processing, measured by using the mean reaction time from the condition in which the irrelevant stimulus was presented on the unattended dimension. Age continued to have a significant impact on time to respond while ignoring a stimulus in the same dimension, even when the influence of IQ differences and speed of processing were partialled out.

It is clearly important to extend and replicate these findings before drawing firm conclusions about the cognitive effects of ageing. However, if they do prove replicable, this pattern of results appears to suggest that the effects of ageing go beyond the hypothesized simple slowing in general speed of information processing or, indeed, decline in fluid intelligence, because neither of these can account for the remaining age difference in the capacity to ignore irrelevant stimuli in the target dimension. On the other hand, this result would seem to be consistent with the proposal by Hasher and Zacks (1988) that age limits the capacity for utilizing inhibition to sharpen attentional focus and limit distraction. Note, however, that our results do not support the idea of a general reduction in inhibition, since the differential effect does not occur when the irrelevant stimuli are presented on a different sensory dimension.

The nature of the inhibition process is at this point purely speculative, but if one uses its physiological counterpart as a model, then one might imagine something like an attentional focus on the features specifying the target, represented by a bell-shaped distribution, with the target at its centre. The nearer a stimulus is in characteristics to the focal-point, the more attention it will receive, and the more rapidly it will be processed. Stimuli falling outside this distribution will be ignored, and those falling on the border will require processing before rejection.

A simple assumption that age leads to a less highly peaked attentional focus might offer a simple account of our results. For older subjects, with a broader distribution of attention, the irrelevant stimulus will demand and receive more processing than for the middle-aged group. In the case of stimuli on a completely different dimension, then, such stimuli will fall outside the focus of attention for both groups.

It is, however, important not to place too much reliance on these preliminary results. At this stage, their principal implication is in the support they lend to the view that the capacity for focused selective attention provides a promising further component of any complete specification of the central executive.

## Activation of Long-term Memory

One feature that must surely be possessed by the central executive but which has been totally ignored by my own work is the capacity for the temporary activation of long-term memory. Consider the case of KJ, a highly intelligent patient with a dense but pure amnesic syndrome (Wilson & Baddeley, 1988). When presented with a paragraph relating a brief story such as that involved in the logical memory subtest of the Wechsler Memory Scale, his immediate recall was above average (12 idea units), although half an hour later, his recall score was zero. He had not only forgotten the story but even forgotten that he had been told a story. KJ's good initial recall is, of course, by no means typical, because many amnesic patients perform extremely badly even on immediate test. The question remains, however, as to how KJ achieves his excellent initial performance.

My own assumption is that, as Johnson-Laird (1983) suggests, the process of comprehension involves setting up some form of mental model, and that this, in turn, demands working memory capacity. It is conceivable that this process operates entirely within the two slave systems, but it seems unlikely that they would be capable of reflecting the semantic complexity, and capacity to utilize earlier learning that seems to lie at the root of comprehension. A more plausible assumption might be to assume that such models represent the temporary activation of components of long-term memory. Such a view has recently been developed by Ericsson and Kintsch (1995) who illustrate this aspect of WM extensively using examples from prose comprehension and from the performance of mnemonic experts.

The idea that working memory might represent the selective activation of representations in long-term memory is not, of course, a novel one, and in North America at least, could probably be regarded as the modal view (e.g. Crowder, 1993; Roediger, 1993). At one level the view is comparatively uncontroversial. Given that memory span for nonwords that resemble English is higher than that for those differing from English in their phonotactic structure (Adams & Gathercole, 1995), it is clear that even the phonological loop is not a *tabula rasa*, but, rather, a system that has developed on the basis of the phonological experience of the rememberer. The phonological store, therefore, depends on activation of a system that itself reflects long-term memory. The problem with the view that STM simply represents currently activated LTM, however, is that such a view is so general as to be theoretically sterile, unless an attempt is made to specify in detail the processes involved. At this point, it is likely to become a matter of theoretical taste whether one chooses to emphasize this single very general feature of dependence on some aspect of LTM or to concentrate on those characteristics that differentiate different memory subsystems. In general, I myself have been more interested in the differences, for example, between the characteristics of the phonological loop and the visuo-spatial sketchpad, or in contrasting both of these with the rich multimodality that characterizes coding in semantic or episodic memory.

Returning to the role of LTM in WM, one way of conceptualizing a hypothetical general retrieval system might be through the concept of a central executive. Such a system should be able to encode and retrieve information both from the slave systems and from temporarily activated components of long-term memory. The Baddeley and Hitch working memory model has already considered one rather specialist form of retrieval, that underlying the recency effect, where an implicit priming mechanism is assumed to be usable across both slave systems and long-term memory (Baddeley & Hitch, 1993). However, we have almost completely neglected the possible role of the central executive in setting up, maintaining, and retrieving temporary representations in long-term memory.

Fortunately, others have been less negligent. In recent years the dominant theme in North American research on working memory has been that stemming from Daneman and Carpenter's (1980) development of the measure they term *working memory span*. This combines the simultaneous requirement to process and store information, and, in its most typical form, involves presenting the subject with a series of sentences, each of which must be processed, and the last word stored. At the end of the sequence, the subject is required to recall the terminal words of each sentence. The maximum number of sentences that can be processed while retaining the final word is the subject's working memory span, and it typically ranges between two and five. The measure has proved very useful in studying the hypothetical role of working memory in a range of complex cognitive tasks, extending from measures of reading comprehension to predicting performance on computer programming courses (Just & Carpenter, 1994; Shute, 1991). When scores on a number of such tasks are combined, the resultant measure correlates very highly with performance on standard intelligence tests (Kyllonen & Christal, 1990), suggesting that it is indeed measuring some capacity of general importance to cognitive functioning. Exactly what this capacity reflects is, however, considerably more controversial, as reflected for example in the contributions to the present volume of Lehto and of Waters and Caplan.

Some of the most careful analytic work in this area has been carried out by Engle, whose recent chapter (in press) gives an excellent overview of his extended research programme. Early work was concerned with the issue of whether the working memory system assumed to underlie the Daneman–Carpenter task was limited to language processing or reflected a more broadly based limited-capacity system. In demonstrating equivalent working memory span effects when arithmetic operations were used instead of sentence verification, Turner and Engle (1989) provided evidence for a more general limited-capacity system—evidence that is, of course, also consistent with our own random generation studies and with Shute's (1991) work on individual differences in working memory span. Subsequent work has been concerned to specify in greater detail the nature of the limitation in capacity that gives rise to individual differences in working memory span.

Cantor and Engle (1993) suggested that working memory might reflect the temporary activation of areas of long-term memory, with high-span subjects being able to activate more extensive regions of long-term memory. Following the assumptions and techniques developed by Anderson (1974), they investigated this hypothesis by using the *fan effect*. This term refers to the fact that time to verify a statement that has previously been presented will be greater if the subject or object of that statement has also been linked to a range of other statements. Hence, a subject who has mastered the set of statements,

*The vicar is in the canoe*
*The vicar has red hair*
*The sailor is in the supermarket*
*The vicar is in Scotland*

will take less time to verify that the sailor (one proposition) is in the supermarket than that the vicar (four propositions) is in the canoe. Anderson explains this pattern of results by assuming that a limited amount of activation automatically spreads from one unit of the sentence to its associated features. As the amount of activation is limited and the vicar is associated with four different features, each link will be weaker than the equivalent link between sailor and supermarket. Cantor and Engle propose that high-working-memory-span subjects simply have more activation available, and, in accordance with this view, they demonstrate that the slope relating set size to verification time is steeper for subjects with a low working memory span, as measured by a variant of the Daneman and Carpenter task.

A later study by Rosen and Engle (cited by Engle, in press) studies the capacity to generate items from a semantic category such as animal names, demonstrating that performance is substantially higher in high-working-memory-span subjects. A third series of experiments explores the relation between working memory span, Anderson's fan effect, and the closely related demonstration by Sternberg (1966) of a linear relationship between the time it takes to decide whether a probe item comes from a set that has just been presented, and set size. Sternberg himself attributed this effect to an internal memory scanning mechanism, but the phenomenon is also open to a wide range of alternative interpretations, of which Anderson's model provides one.

An ingenious series of experiments by Conway and Engle (1994) is based on the link between the Sternberg and fan effects. Subjects are first of all taught groups of two, four, six, or eight letters, to a point at which, when asked for that group, they can provide the constituent letters perfectly. When given the group and a particular letter, subjects are thus able to indicate whether or not that letter falls within that group. As predicted, reaction time increases linearly with number of items within the probed group, with the slope being steeper for subjects with low working memory span. By specifying the set first (e.g. the four-letter set) but delaying the presentation of the probe letter, Conway and Engle are able to separate out the time it takes to access a given set, from the time to check it for the presence of the probe letter. They find that working memory span does not influence the time to access the set, only the time to verify the presence of the probe. They conclude from this that the former retrieval process is relatively automatic and does not depend on limited-capacity working memory, whereas the latter involves an active search process that depends upon the limited-capacity system available.

The data described so far are captured well by a model that assumes that individual differences in working memory reflect differences between subjects in the amount of activation available. However, despite this supportive evidence, Engle reports two further observations that cause him to abandon this hypothesis. The first of these concerns the detailed analysis of the performance of high- and low-span subjects on the various tasks. Although the effects I have described all hold for high-span subjects, many of them do not do so for subjects with low spans (Engle, in press). For example, combining the category-generation task with an attention-demanding concurrent activity reduces the performance of high-span subjects but has no effect on those with low spans. Having subjects learn a subset of categorized items, which they are instructed to ensure are *not* included in their category generation, has no effect on low-span subjects but severely inhibits the performance of those with high spans. Somewhat surprisingly, a marked effect occurs even when the items to be excluded come from a category (e.g. *furniture*) that is totally unrelated to the generation category (e.g. *animals*). Such findings have important implications for the general utilization of the working memory span measure, because they suggest that rather than providing a measure of some continuously varying capacity across subjects, the measures differentiate between subjects who are using different strategies. Although the differences in strategies may well result from underlying differences in some form of processing capacity, this needs to be tapped more directly if the measure is to be used as anything more than a convenient way of dichotomizing groups into good and poor performers.

The second observation is even more problematic for the simple excitation hypothesis. It resulted from Engle's concern that his experiments based on the Sternberg paradigm had required subjects to learn sets of items that were not mutually exclusive—hence the letter *K* could appear in both Set Size 2 and Set Size 6. In two final experiments, this confounding was avoided, with no item occurring in more than one set. Under these circumstances, the difference in slope between high- and low-working-memory-span subjects disappeared, and the linear relationship between set size broke down around length 8. It therefore appears to be the case that a crucial feature of the previous results was the overlap of items across categories. Engle abandons his earlier model in favour of something that more closely resembles the Baddeley and Hitch concept of a central executive, but an executive that is limited principally in its capacity to inhibit irrelevant information. As Engle points out, such a view is very close to that of Hasher and Zacks; it is, of course, consistent with our own much less extensive findings from the previously described age and selective attention studies.

The presence of individual differences in inhibitory capacity does not, of course, rule out the possibility that excitatory processes also differ across individuals. However, Engle's results suggest at the very least that we need to look very carefully at claims for such differences. Engle's results ask some searching questions about three of the major phenomena of cognitive psychology, namely the Sternberg effect, the fan effect, and the Daneman–Carpenter measure of working memory span. As such, they have substantial implications for understanding the role of working memory in retrieval from long-term memory.

## Conclusion: Should We Sack the Homunculus?

To what extent is it still useful to talk about the central executive as if it were a unitary system? My own view is that it continues to be a useful concept that is able to focus on both the attentional characteristics of working memory and its more traditional links with short-term visual and verbal memory. In short, I think the homunculus can be useful, given two provisos: (1) the continued recognition that it constitutes a way of labelling the problem, not an adequate explanation; (2) a continued attempt to understand the component processes that are necessary for executive control, gradually stripping away the various functions we previously attributed to our homunculus, until eventually it can be declared redundant. Whether we will then be left with a single coordinated system that serves multiple functions, a true executive, or a cluster of largely autonomous control processes—an executive committee—remains to be seen.

## REFERENCES


Adams, A.M., & Gathercole, S.E. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech and Hearing Research*, *38*, 403–414.

Alderman, N. (in preparation). *Maximising the learning potential of brain injured patients*.

Allport, A., Styles, E.A., & Hsieh, S. (1994). Shifting international set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.

Anderson, J.R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*, 451–474.

Baddeley, A.D. (1966). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology*, *18*, 119–129.

Baddeley, A.D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A.D., Bressi, S., Della Sala, S., Logie, R., & Spinnler, H. (1991). The decline of working memory in Alzheimer's Disease: A longitudinal study. *Brain*, *114*, 2521–2542.

Baddeley, A.D., Emslie, H., Kolodny, J., & Duncan, J. (in preparation). Random generation and the executive control of working memory.

Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation, Vol. 8* (pp. 47–89). New York: Academic Press.

Baddeley, A.D., & Hitch, G.J. (1993). The recency effect: Implicit learning with explicit retrieval? *Memory and Cognition*, *21*, 146–155.

Baddeley, A.D., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, *113*, 518–540.

Baddeley, A.D., Logie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and working memory. *Quarterly Journal of Experimental Psychology*, *38A*, 603–618.

Baddeley, A.D., & Wilson, B. (1986). Amnesia, autobiographical memory and confabulation. In D.C. Rubin (Ed.), *Autobiographical memory* (pp. 225–252). New York: Cambridge University Press.

Baddeley, A.D., & Wilson, B. (1988). Frontal amnesia and the dysexecutive syndrome. *Brain and Cognition*, *7*, 212–230.

Burgess, P.W., & Shallice, T. (in press). Fractionnement du syndrome frontale. *Revue de Neuropsychologie*.

Cantor, J., & Engle, R.W. (1993). Working memory capacity as long-term memory activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *5*, 1101–1114.

Cattell, R.B., & Cattell, A.K.S. (1960). *Handbook for the individual or group Culture Fair Intelligence Test*. Champaign, IL: Testing Inc.

Conway, A.R.A., & Engle, R.W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*, 354–373.

Craik, F.I.M. (1995). *Should PET change our views on the frontal lobes and memory?* Presented at Memory Disorders Research Society Meeting, King's College, August.

Craik, F.I.M., Anderson, N., Kerr, S.A., & Li, K.H. (1995). Memory changes in normal ageing. In A.D. Baddeley, B.A. Wilson, & F.N. Watts (Eds.), *Handbook of memory disorders*. Chichester: John Wiley.

Crowder, R.G. (1993). Systems and principles in memory theory: Another critique of pure memory. In A.F. Collins, S.E. Gathercole, M.A. Conway, & P.E. Morris (Eds.), *Theories of memory* (pp. 139–161). Hove: Lawrence Erlbaum Associates, Ltd.

Dalrymple-Alford, J.C., Kalders, A.S., Jones, R.D., & Watson, R.W. (1994). A central executive deficit in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 57*, 360–367.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Della Sala, S., Baddeley, A., Papagno, C., & Spinnler, H. (in press). Dual-task paradigm: A means to examine central executive. *Annals of the New York Academy.*

Della Sala, S., Baddeley, A., Papagno, C., & Spinnler, H. (in preparation). *Dual task performance in dysexecutive and non-dysexecutive patients with a focal frontal lesion.*

Della Sala, S., Gray, C., Spinnler, H., & Trivelli, C. (in preparation). *The riddle of executive testing.*

Dienes, Z., Broadbent, D.E., & Berry, D.C. (1991). Implicit and explicit knowlege bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 875–887.

Duncan, J. (1986). Disorganisation of behaviour after frontal lobe damage. *Cognitive Neuropsychology, 3*, 271–290.

Duncan, J. (1993). Selection of input and goal in the control of behaviour. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness and control* (pp. 171–187). Oxford: Clarendon Press.

Duncan, J., Johnson, R., Swales, M., & Freer, C. (in preparation). *Frontal lobe deficits after head injury: Unity and diversity of function.*

Engle, R.W. (in press). Working memory and retrieval: An inhibition-resource approach. In J.T.E. Richardson (Ed.), *Working memory in human cognition.* New York: Oxford University Press.

Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Gathercole, S.E., & Baddeley, A.D. (1994). *Working Memory and Language.* Hove: Lawrence Erlbaum Associates, Ltd.

Gick, M.L., Craik, F.I.M., & Morris, R.G. (1988). Task complexity and age differences in working memory. *Memory and Cognition, 16*, 353–361.

Greene, J., Baddeley, A.D., & Hodges, J. (in press). Autobiographical memory and executive function in early dementia of Alzheimer type. *Neuropsychologia.*

Hasher, L., & Zacks, R.T. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), *The psychology of learning and motivation, Vol. 22* (pp. 193–225). New York: Academic Press.

Heim, A. (1975). *AH2 and AH3.* Windsor, Berks: NFER Publishing.

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

Just, M.A., & Carpenter, P.A. (1994). Unpublished paper presented at the Working Memory Conference, Cambridge.

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity. *Intelligence, 14*, 389–433.

Lezak, M.D. (1983). *Neuropsychological assessment.* New York: Oxford University Press.

Logie, R.H. (1995). *Visuo-spatial working memory.* Hove: Lawrence Erlbaum Associates, Ltd.

Morris, R.G. (1986). Short-term forgetting in senile dementia of the Alzheimer's type. *Cognitive Neuropsychology, 3*, 77–97.

Morris, R.G., Gick, M.L., & Craik, F.I.M. (1988). Processing resources and age differences in working memory. *Memory and Cognition, 16*, 362–366.

Nelson, H.E. (1976). A modified card-sorting task sensitive to frontal lobe defects. *Cortex, 12*, 313–324.

Norman, D.A., & Shallice, T. (1980). *Attention to action: Willed and automatic control of behavior.* University of California at San Diego, CHIP Report 99.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia, 12*, 323–330.

Raaijmakers, J.G.W., & Shiffrin, R.M. (1981). Search of associative memory. *Psychological Review, 88*, 93–134.

Rabbitt, P.M.A. (1983). How can we tell whether human performance is related to chronological age? In D. Samuel, S. Alegeri, S. Gershon, V.E. Grimm, & G. Teffano (Eds.), *Aging of the Brain* (pp. 9–18). New York: Raven Press.

Robbins, T.W., Anderson, E.J., Barker, D.R., Bradley, A.C., Fearneyhough, C., Gillespie, P.H., Henson, R., Hudson, S.R., & Baddeley, A.D. (in press). Working memory in chess. *Memory and Cognition*.

Robertson, I.H., Ward, T., & Ridgeway, V. (1994). *The Test of Everyday Attention*. Flempton: Thames Valley Test Company.

Roediger, H.L. (1993). Learning and memory: Progress and challenge. In D.E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 509–528). Cambridge, MA: MIT Press.

Rylander, G. (1939). Personality changes after operation on the frontal lobes. *Acta Psychiatrica Neurologica* (Supplement No. 30).

Salthouse, T.A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, *2*, 179–183.

Salthouse, T.A. (1992). *Mechanisms of age-cognition relations in adulthood*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transaction of the Royal Society London, B*, *298*, 199–209.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.

Shallice, T., & Burgess, P.W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*, 727–741.

Shallice, T., & Burgess, P. (1993). Supervisory control of action and thought selection. In A. Baddeley & L. Weiskrantz (Eds.) *Attention: Selection, awareness and control* (pp. 171–187). Oxford: Clarendon Press.

Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*, 127–190.

Shute, V.J. (1991). Who is likely to acquire programming skills? *Journal of Educational Computing Research*, *7*, 1–24.

Spinnler, H., Della Sala, S., Bandera, R., & Baddeley, A.D. (1988). Dementia, ageing and the structure of human memory. *Cognitive Neuropsychology*, *5*, 193–211.

Sternberg, S. (1966). High speed scanning in human memory. *Science*, *153*, 652–654.

Thurstone, L. (1938). Primary mental abilities. *Psychometric Monograph*, *1*.

Turner, M.L., & Engle, R.W. (1989). Is working-memory capacity task-dependent? *Journal of Memory and Language*, *28*, 127–154.

Welford, A.T. (1958). *Ageing and human skill*. London: Oxford University Press.

Wilkins, A.J., Shallice, T., & McCarthy, R. (1987). Frontal lesions and sustained attention. *Neuropsychologia*, *25*, 359–365.

Wilson, B.A., & Baddeley, A.D. (1988). Semantic, episodic and autobiographical memory in a post-meningitic amnesic patient. *Brain and Cognition*, *8*, 31–46.

Yntema, D.B., & Mueser, G.E. (1960). Remembering the present states of a number of variables. *Journal of Experimental Psychology*, *60*, 18–22.

*Manuscript received 11 November 1994*
*Accepted revision received 16 August 1995*