

What do working-memory tests really measure?

Michael J. Kane,

Andrew R. A. Conway,

and Randall W. Engle

Department of Psychology, Georgia State University, Atlanta, GA 30303-3083;

Department of Psychology, University of Illinois at Chicago, Chicago, IL 30332-0170; and

School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332-0170. mkane@gsu.edu aconway@uic.edu re23@prism.gatech.edu

Abstract: Individuals may differ in the general-attention executive component or in the subordinate domain-specific “slave” components of working memory. Tasks requiring sustained memory representations across attention shifts are reliable, valid indices of executive abilities. Measures emphasizing specific processing skills may increase reliability within restricted samples but will not reflect the attention component responsible for the broad predictive validity of span tasks.

We thank Caplan & Waters (C&W) for pointing out the weaknesses in the view that language interpretation and parsing are limited by general, domain-free working memory resources (see, e.g., Just & Carpenter 1992). Based largely on their earlier work (Caplan & Waters 1990), Engle and Conway (1998) concluded

that there is little evidence to support the claim that the phonological loop or the central executive plays an important role in the early stages of language processing. We have argued that the working memory system is theoretically equivalent to a view of human intelligence proposed by Cattell (1963) and Horn (1980), which distinguishes between a domain-free general fluid intelligence and domain-specific elements of crystallized intelligence, which are peculiar to knowledge, skill, and talent.

Our view, derivative of Baddeley and Hitch (1974), sees working memory as a system of (1) activated traces, which are the output of domain-specific processors, and (2) domain-free controlled attention (Engle et al., in press b; Kane & Engle 1998). Individuals may vary in the quality and skill of their domain-specific processors. They may also vary in their general fluid ability to bring controlled attention to bear on managing information in the temporary storage buffers and on resolving conflict at output (cf. Shallice & Burgess 1993). Thus, variability on working memory tasks can reflect individual differences in the domain-specific processors, the domain-free controlled attention, or both.

We argue that the general capability to sustain an active memory representation in the face of attention shifts or distraction is what drives the relationship between working memory span measures and diverse higher order capabilities. C&W complain that

the processing component of working memory span tasks “introduces a dual-task element into these tasks and allows them to be heavily affected by the capacity for divided attention.” We argue, however, that the divided-attention component is the strength of working memory span tasks. The divided-attention component is critical to the broad predictive validity of span tasks (and, therefore, of fluid intelligence tests) across task and stimulus domains. All studies that use individual or group differences to make inferences about working memory involve the measurement of constructs and the partitioning of variance in those constructs. If the range of general ability in a sample of subjects is sufficiently broad, then measures of working memory capacity, such as the reading span task (Daneman & Carpenter 1980) and the operation span task (Turner & Engle 1989), will be reliable. Furthermore, they will reflect the ability of both the domain-free central executive and the domain-specific processors necessary to perform the tasks because no working memory task is a pure test of either the central executive or the slave systems (Engle et al., in press a). If, however, a sample’s range of general ability is narrow, as it often is at prestigious universities, there will be little variability in the general fluid ability construct. Therefore, the variability in performance for the restricted sample will reflect primarily differences in the skill of the domain-specific processors necessary to perform

the task.

To the extent that the “processing” component is entered into the working memory span score as C&W suggest, reliability may increase. However, the construct measured by that score is more likely to reflect the domain-specific processors required for that task and not the general controlled attention–central executive construct. In fact, our own work (e.g., Conway & Engle 1996; Conway et al. 1998) suggests that processing skills within span tasks can actually suppress the relationship between working memory scores and higher order cognition. Thus, the “innovation” proposed by C&W may increase reliability in a restricted sample, but it will be a good test only if the goal is to measure specific processing skills. It will not be a good test if the goal is to measure general domain-free controlled attention capacity. If the goal is to measure the domain-free executive component of working memory, then we recommend administering a battery of different working memory tasks that share the dual-task quality mentioned above but differ in the domain-specific processes required to perform the task.

C&W questioned the reliability of the reading span task, but, as we pointed out above, low reliability can result from sampling a restricted range of individuals. Kitty Klein and William Fiss of North Carolina State University recently conducted an extensive

analysis of the operation span task (Turner & Engle 1989). They tested 33 subjects at three different times, the second time 3 weeks after the first and the third time 6–7 weeks after the second. The corrected reliability was .88 and the stability scores ranged from .76 to .92.

As for the “fractionation” of working memory, we would argue that C&W’s measures reflect quite specific skills at parsing and processing syntactically complex sentences. These processes occur at a pre-interpretative level and thus do not generally make demands on the central executive. The two different components they propose reflect the outputs from specific processors that do their work with little need for limited-capacity, controlled attention. For example, adding a secondary working memory load to a

Commentary/Caplan & Waters: Working memory and sentence comprehension

BEHAVIORAL AND BRAIN SCIENCES (1999) 22:1 101 syntactic processing task has no detrimental effect. Because syn-

tactic processing may proceed unhindered by a load, it may be performed relatively automatically, without much controlled attention. However, secondary load tasks do interfere with syntactic processing when the load-task stimuli are interleaved with the syntax task, demonstrating that, when syntactic information must be sustained across an attention shift, it suffers significantly.

Rather than assign syntactic processing to a specialized component of working memory, then, we suggest that it operates inde-

pendently of the central executive. That is, working memory capacity is needed only under attention-demanding circumstances, and, insofar as syntactic processing appears to be immune to divided-attention conditions, it likely occurs relatively automatically. Caplan and Waters (1990) argued that the phonological loop may be required in some syntactic parsing circumstances, such as when subjects are “garden-pathed” or when many words must be maintained in active memory. Why not use that interpretation for the present work?