What do working-memory tests really measure?

Michael J. Kane, Andrew P. A. Conway, and Randall W. Engle

Abstract: Individuals may differ in the general-attention executive component or in the subdomainspecific "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). Individu-
The age invariance of working memory measures and noninvariance of producing complex syntax

Susan Kemper and Karen A. Kemtes
Department of Psychology and Gerontology Center, University of Kansas, Lawrence, KS 66045, skemper@ukans.edu
www.lsi.ukans.edu/dlang/sklab.htm kkemtes@ukans.edu

Abstract: In challenging current conceptions of the role of working memory in sentence processing, Caplan & Waters consider studies comparing young and older adults on sentence processing. This commentary raises two challenges to Caplan & Waters' conclusions: (i) working memory tasks appear to be age invariant, and (ii) the production of complex syntactic constructions appears not to be age invariant.

Age invariance. C&W question whether the tasks commonly used to measure individual differences in working memory are reliable and valid. An additional concern is that if not considered by C&W, is whether the tasks are age invariant. As Stine (1995) notes, even when younger and older adults are matched based on working memory scores, researchers might inevitably find that older adults have processing difficulties that are seemingly unexplained by their performance on the working memory tasks.

Horn and McArdle (1992) explain that the latent variable approach to constructing measurement permits the test of two common concepts that receive little attention in intergenerational research: construct validity and construct invariance, across samples such as age groups. Horn and McArdle define measurement invariance as factorial invariance: the characteristic that measurements are composed of linear composites, that is, item scores summed to form a total score, is stable across measurement occasions of same or different samples. The authors present four criteria that are sufficient to determine factor invariance: (1) Factor patterns are equivalent across groups, a necessary condition for factor invariance to hold; (2) Factor variances are equivalent across groups; (3) Factor covariances are equivalent across groups; (4) Factor averages of composite scores are equivalent across groups. A fifth criterion they impose is whether cross-products are equivalent across groups. If the answer is "yes," then the answer to questions (1–4) is necessarily "yes": if the answer is "no," invariance is not necessarily satisfied.

To date, few studies have used a latent variable approach to assess the age invariance of working memory measures. As an example of this approach, Kemtes and Kemper (1998) assessed working memory performance for 248 adults (young and old) in four separate studies using a variety of verbal working memory tasks. With the latent variable approach, the authors found that a simple within-studies factor analytic structure with a single latent variable of working memory was reliable across age groups. For three of the studies, this single-latent-factor working memory model was validated across age groups, suggesting that a common construct was measured. Kemtes and Kemper's finding of factor invariance across these different samples strongly supports the conclusions that working memory, as measured by the digit spans, reading span, and listening span, is age invariant.

Age noninvariance. C&W also review the existing literature on the effects of aging on sentence processing and conclude that there are few age differences in on-line sentence processing although post-interpretative processes of question answering and text recall may be affected by aging. However, they overlook an additional body of research on age differences in production. Older adults show a reduction in their production of complex syntactic constructions such as those involving subordinate and embedded clauses (Kemper 1987, 1989; Kemper et al. 1989, 1992, Kynett & Kemper 1986). The age-related decline in syntactic production is somewhat greater for left-branching constructions (e.g., *The girl who runs a nursery school for our church is awfully young*) than for coordinate or right-branching constructions (e.g., *She's awfully young to be running a nursery school for our children*). This asymmetry in the production of right-branching constructions versus left-branching constructions provides strong evidence for the effects of working memory limitations. Left-branching constructions, including center-embedded object relative clauses such as *The dog that the man that the cat bit chased escaped*, are typically considered more complex than right-branching constructions such as *The cat bit the man that chased the dog that escaped* (Gibson et al. 1996; Lewis 1996b).

Cheung and Kemper (1992) investigated the relationship between age, working memory, and production using a number of different measures of linguistic complexity including: mean length of utterance (MLU: Chapman & Miller 1984), developmental sentence scoring (DSS; Lee 1974), developmental level (DLEVEL; Rosenberg & Abbeduto 1987), two alternative ways of measuring Yngve depth (Yngve 1960), and two variants of Frazier's (1985) node count. In addition, propositional density (PDENSITY), based on Knitsch and Keenan's (1973) analyses of text difficulty, was computed in order to assess whether semantic content covaries with grammatical complexity. Cheung and Kemper (1993) applied structural equation modeling to these linguistic complexity metrics using language samples from younger and older adults. The best-fitting model fit the data by specifying two correlated factors, verbal ability and working memory. Age was negatively associated with working memory, leading to a decline in digit span with advancing age, and was somewhat positively associated with verbal ability, reflecting a slight improvement in vocabulary with advancing age. Working memory was related to three syntactic factors: length, measured by MLU; the amount of embedding, measured by MCU, and the type of embedding, measured by DSS and DLEVEL as well as by both Yngve depth metrics and both Frazier counts. Finally, verbal ability predicted another linguistic factor, semantic content, measured by PDENSITY, which was not correlated with the syntactic factors.

Working memory limitations associated with aging affect older adults' production of complex syntactic structures. A key determinant of syntactic complexity, affecting the DSS, DLEVEL, Yngve, and Frazier metrics, is whether embeddings occur in the main-clause subject of left-branching constructions such as "Going to the St. Louis World Fair was a major undertaking" or in the