



Published in final edited form as:

Monogr Soc Res Child Dev. 2013 August ; 78(4): 70–87. doi:10.1111/mono.12035.

NIH Toolbox Cognitive Function Battery (NIHTB-CFB): Measuring Working Memory

David S. Tulsky¹, Noelle Carlozzi¹, Nic Chevalier², Kim Espy², Jennifer Beaumont³, and Dan Mungas⁴

¹Department of Physical Medicine and Rehabilitation, University of Michigan Medical School, Ann Arbor, MI

²Department of Psychology, University of Nebraska-Lincoln, Lincoln, NE

³Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL

⁴Department of Neurology, University of California Davis Health System, Davis, CA

Abstract

This chapter focuses on the Toolbox List Sorting Working Memory Test, which was developed to assess processing speed within the NIH Toolbox Cognitive Function Battery (NIHTB-CFB). This test was adapted from the Spanish and English Neuropsychological Assessment Scales (SENAS) List Sorting task, which is an auditory working memory sequencing test that had been validated and normed in samples of older adults (Crane, Narasimhalu, Gibbonset al., 2008; Mungas, Reed, Tomaszewski, Farias, & DeCarli, 2005). We describe the development of the Toolbox List Sorting Working Memory Test, highlighting its utility in children. In addition, we examine descriptive data, test-retest reliability, and convergent and discriminant validity. Results indicated that List Sorting performance was positively correlated with age; performance on the task improved throughout childhood and early adolescence. Further, test-retest reliability was very good and there was support for both convergent and discriminant validity. These data suggest that the NIH Toolbox List Sorting Working Memory Test is reliable and shows evidence of construct validity.

In this chapter we discuss the development of the Toolbox List Sorting Working Memory Test, a new measure of working memory.

The NIH Toolbox (NIHTB) is designed to be comprised of sensitive measures to evaluate cognitive, emotional, sensory and motor functioning across the lifespan (ages 3 to 85 years). In particular, the NIHTB-Cognitive Function Battery (CFB) was designed to evaluate executive function (including both cognitive flexibility and inhibitory control), episodic memory, language (including both vocabulary comprehension and reading decoding), working memory, processing speed, and sustained attention, in less than 30 minutes. In this chapter, we discuss the development of the Toolbox List Sorting Working Memory Test, a new measure of working memory. We place a special emphasis on the relevance of this task

for children, as well as on its convergent and discriminant validity. Below, we identify and define the main constructs of working memory and provide a review of literature on working memory development during childhood.

Subdomain Definition

Working memory is probably one of the most widely studied constructs in psychology due to its prominent role in complex cognitive tasks and daily activities (e.g., mental arithmetic and reading). It is a capacity-limited system devoted to holding information in mind over brief periods of time, typically while manipulating it for ongoing activity. Given the importance of this construct, a wide number of tasks have been designed to assess it. In young children, working memory tasks tend to focus on actively maintaining information over brief intervals in the face of interference. From preschool age on, most tasks either require retaining and reorganizing items before recalling them (e.g., backward digit span task), or completing some processing activity in between presentations of the to-be-recalled items (e.g., listening span task). Such tasks tap into both information processing and storage, and yield a working memory span measure that corresponds to the maximal amount of accurately recalled information.

The cognitive structure of working memory is actively debated. Baddeley's (Baddeley & Hitch, 1974) tripartite model is one of the most influential models of working memory. It identifies two domain-specific components, the *phonological loop* and *visuospatial sketchpad* (devoted to temporary storage of verbal and visuospatial information, respectively), and a domain-general component, the *central executive* (responsible for filtering out irrelevant information, integrating the information held in the other two components and in long-term memory and supervising its processing). A fourth component—the *episodic buffer*—was subsequently added as the storage locus for integrated information (Baddeley, 2000). The central executive also is a main component of Cowan's (2005) model which posits working memory as the activated element of long-term memory, wherein the most strongly activated information—the *focus of attention*—receives direct attention from the central executive. Because resources of the central executive are limited, the degree of control dilutes as the amount of information increases in the focus of attention (whose capacity is limited to 3–4 items). Building on the proposal of two levels of activation in working memory, Unsworth and Engle (2007) argued that, when information no longer receives attention, it leaves the focus of attention and must subsequently be retrieved from the activated long-term memory using context cues. Both active maintenance in the focus of attention and retrieval processes from activated long term memory allegedly tax the central executive and contribute to working memory capacity. Despite structural differences across theoretical models (Cowan, 2005; Engle, Kane, & Tuholsky, 1999; Miyake & Shah, 1999), all models identify temporary storage and control components, as illustrated by Engle et al.'s (1999) formula, working memory = short-term memory + controlled attention.

Working memory encompasses short-term memory (i.e., temporary storage of information, irrespective of processing demands) and shares its properties. In particular, information in working memory is short-lived and susceptible to the interference created by goal-irrelevant information unless it is shielded and/or actively maintained through attention control. Given

this prominent role of attention control, working memory also is intermingled with the construct of executive functions, that is, the set of cognitive processes that support goal-oriented thought and action. Executive processes are considered to comprise inhibiting irrelevant information, switching task sets, updating working memory content, and maintaining information in an active state (e.g., Garon, Bryson, & Smith, 2008; Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000 ; Munakata, 2001). Therefore, executive processes can be viewed as the processes by which the central executive component operates. In the developmental literature, working memory is often conceived as a subset of executive functions, but in such case, working memory generally refers to the active maintenance process (instead of the whole construct of working memory). Consistently, tasks assessing working memory and executive functions load onto a single latent factor in children under 7 years of age, lending support to substantial shared variation between the two constructs (Shing, Lindenberger, Diamond, Li, & Davidson, 2010; Wiebe, Espy, & Charak, 2008; Wiebe et al., 2011).

Importance During Childhood

Working memory development has a tremendous impact on children's cognition as it is associated with academic achievement, including mathematic skills and reading skills (e.g., Bull & Scerif, 2001; Nevo, & Breznits, 2010). Temporary maintenance and manipulation of information is required during learning episodes in the classroom as these often require remembering lengthy instructions. Thus, low working memory capacity puts children at risk for poor academic progress (Alloway & Gathercole, 2006; Alloway, Gathercole, Kirkwood, & Elliott, 2009). In addition, working memory deficits, especially in verbal short-term memory, have been associated with reading difficulties and developmental dyslexia (e.g., Gathercole, Alloway, Willis, & Adams, 2006; Smith-Spark & Fisk, 2007). Working memory deficits also have been reported in a variety of developmental disorders including Attention Deficit/Hyperactivity Disorder (e.g., Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005), specific language impairment (e.g., Briscoe & Rankin, 2009), and autism (e.g., Williams, Goldstein, & Minshew, 2006), as well as medical conditions (e.g., childhood cancer, Dennis et al., 1991), prematurity (e.g., Vicari et al., 2004), and traumatic brain injury (e.g., Levin et al., 2004). Although individual differences in working memory capacity are relatively stable over time, recent evidence suggests that working memory capacity can be effectively improved through intervention programs during childhood (Holmes, Gathercole, & Dunning, 2009; Thorell, Lindqvist, Nutley, Bohlin, & Klingberg, 2009). Such evidence opens up avenues to remediate low working memory capacity in children and enhance academic outcome, and therefore reinforces the need for precise identification of at-risk children through adequate assessment of working memory during childhood.

Relations of Domain with Brain Function

In adults and older children, working memory task performance is sustained by a distributed fronto-parietal network that includes parietal cortex (including the intraparietal sulcus, inferior and posterior parietal cortex), dorso-lateral prefrontal cortex, and ventro-lateral prefrontal cortex (Bunge, Klingberg, Jacobsen, & Gabrieli, 2000; Kwon, Reiss, & Menon, 2002; Wager & Smith, 2003). Some studies additionally report involvement of the anterior

cingulate cortex, basal ganglia (especially, the striatum), medial temporal cortex, and cerebellum (Chein, Moore, & Conway, 2011; McNab & Klingberg, 2008; O'Hare, Lu, Houston, Bookheimer, & Sowell, 2008; Osaka et al., 2004). Consistent with the role of prefrontal regions in executive functions (e.g., Casey, Galvan, & Hare, 2005), evidence suggests that these regions and basal ganglia act as a selective gating mechanism that controls the information accessing working memory and maintained in parietal regions (McNab & Klingberg, 2008; Postle, 2006). Further evidence comes from findings that the nature of the task items influence activation in the parietal lobe (left-lateralized for verbal items, right-lateralized for visuospatial items; Thomason et al., 2008) but not in prefrontal regions (Wager & Smith, 2003). Although prefrontal regions may support domain-general processes, dorsal prefrontal regions may be especially involved in information manipulation whereas ventral prefrontal regions may more strongly relate to active maintenance of the information stored in posterior regions (D'Esposito, Postle, Ballard, & Lease, 1999; Wager & Smith, 2003).

Working memory capacity develops on a protracted course. As early as 6 months of age, infants are able to retain information over brief intervals in spite of distraction (Reznick, Morrow, Goldman, & Snyder, 2004). Working-memory capacity then steadily improves through late adolescence (Gathercole, Pickering, Ambridge, & Wearing, 2004; McAuley & White, 2011). On average, a preschooler's working memory span triples by early adulthood (Dempster, 1981), although at each age span length varies as a function of context-specific demands (Conlin, Gathercole, & Adams, 2005). The tripartite structure of working memory (verbal and visuospatial short-term stores along with a control entity) is observable from age 4 years on, suggesting that little structural change occurs after that age, although the relation between visuospatial information storage and executive control seems stronger between ages 4 and 6 years than later in childhood (Alloway, Gathercole, & Pickering, 2006; Gathercole et al., 2004). However, visuospatial information may heavily draw upon executive control in adulthood as well (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). Consistent with evidence of constant working memory structure across ages, children recruit the same fronto-parietal network while performing working memory tasks as do adults (e.g., Nelson, Monk, Lin, Carver, Thomas, & Truwit, 2000). With age, however, activation in these regions becomes stronger and more focal, and as children get older, working memory involves other regions such as the cerebellum. This is also common in adults as working-memory demands increase (Geier et al., 2009; Kwon et al., 2002; Thomason et al., 2008).

Developmental change in working memory is driven by age-related increase in both temporary storage capacity and control efficiency (see Cowan, 2010; Gathercole et al., 2004), the latter probably being related to improvement in executive function (e.g., Best, Miller, & Jones, 2009; Carlson, 2005). In addition to such quantitative improvements, working memory development also results from change in strategy use over age (e.g., Camos & Barrouillet, in press). Rehearsal of verbal information is an especially efficient strategy whose corresponding neural circuit (including Broca's area, premotor cortex, and inferior parietal areas) is part of the fronto-parietal network associated with working memory (e.g., Kwon et al., 2002). Yet, only around age 7 years do children start verbally rehearsing information spontaneously (see Gathercole & Hitch, 1993). Similarly, with age,

children become increasingly prone to recode visuospatial information into a phonological format so that it can be rehearsed more easily (Hitch & Halliday, 1983).

Processing speed is another important factor to consider. Indeed, increase in processing speed has been shown to account for up to 75% of the variance in working memory improvement with age (Fry & Hale, 2000; Kail & Hall, 2001; McAuley & White, 2011; Nettlebeck & Burns, 2010). According to cascade theory (Fry & Hale, 1996), age-related improvement in processing speed drives changes in working memory which, in turn, lead to increasing fluid intelligence (see also Case, 1987). There are at least two potential reasons why working memory span increases with processing speed. First, faster processing speed may speed up the rate of verbal rehearsal, hence improving information maintenance (Fry & Hall, 2000). Second, higher processing speed may accelerate information manipulation and therefore leave extra time when attention can be allocated to information maintenance (Barrouillet et al., 2004; Towse et al., 1998). Prominent as the role of processing speed may be, part of working memory change occurs independently of this skill (McAuley & White, 2011).

NIHTB-CFB Measurement

Development of the Toolbox List Sorting Working Memory Test

The Toolbox List Sorting Working Memory Test is a sequencing task requiring children and adults to sort information and sequence it. Items are presented both visually and auditorily. The participants (either children or adults) are presented with a series of illustrated pictures, each depicting an item (e.g., an animal) on the computer, along with their auditory names. Participants are instructed to remember the stimuli but to repeat them verbally to the examiner in order of size, from smallest to largest. The number of objects in a series increase on successive items thereby taxing the working memory system when longer sequences need to be remembered. Furthermore, the task starts with a “1-list” version where the children have to sequence one type of stimuli (e.g., “animals” or “food”) according to size order and then switch to a “2-list” version where two types of stimuli have to be sequenced, each in size order. In the 2-list version, the working memory load is increased substantially as the stimuli are presented from two categories (animals and food) and the participant has to track and organize stimuli from both categories and report by size the items from one category (i.e., animals) and then the other category by size (i.e., food). It is this “dual” tracking and processing information that increases the working memory load of the task. See Figure 1 for an example of 1-List and 2-List items.

Other sequencing tasks include the Letter Number Sequencing Test (Gold, Carpenter, Randolph, Goldberg, & Weinberger, 1997) that was incorporated into the Wechsler Intelligence Scales (WAIS-III, WAIS-IV, and WISC-IV; Tulsky, Saklofske, & Zhu, 2003; The Psychological Corporation, 1997; Wechsler, 1997; Wechsler, 2003, Wechsler, 2008) and the Spanish and English Neuropsychological Assessment Scales (SENAS) Working Memory task (Mungas, Reed, Tomaszewski Farias & DeCarli, 2005; Crane, Narasimhalu, Gibbonset al., 2008). The Toolbox List Sorting Working Memory Test that we have developed is modeled after the SENAS, which is an auditory working memory test; stimuli are presented both visually (object) and auditorily (corresponding word) in an attempt to

make the task easier and more relevant to children. The earliest versions of the Toolbox List Sorting Working Memory Test provided children with multiple opportunities for practice (prior to the administration of actual test items). Children were given two practice items involve sequencing of toy animals, followed by computerized items presenting all items on the screen simultaneously, followed by practice items involving individual administration of each item in the series.

Preparing the Toolbox List Sorting Working Memory Test for Children

Most working memory tests are developed using one modality, either visual or auditory, and the tests are often used to evaluate cognition the specific subsystems of this cognitive process (e.g., visuospatial sketchpad, phenomenological loop. In the case of the Toolbox, the distinction between visual and auditory working memory was less important and the test development team had visual images drawn to accompany auditory presentation of the stimuli. The logic was that the visual images would enhance the usability of the task for children. Three preliminary studies were conducted to adapt the List Sorting task for pediatric use and ensure that the task is relevant to children.

For the first pilot study, ten 3-year-olds were recruited from a local nursery school. The goal of this study was to examine task feasibility of list sorting. Children were shown the visual pictures of the objects and asked to repeat what they had seen except reorder them according to size, smallest to largest. The children received three practice items before the test began. We examined the range of scores on test items for our participants to determine test feasibility. We also examined performance on practice items to determine if children as young as 3 years of age were able to understand the basic tenets of this task; children were given up to three training trials on each item, prior to moving on to the next item. Three-year-olds scores ranged from 0–5 on the 1-list, and 0–3 on the 2-list. Further, all children were able to complete all of the practice items within the three training trials. These results suggested that children as young as 3 years of age were able to understand the basic concepts of the task. In general, children were able to sequence items in size order. From these data, we revised the task, increasing the number of items on both the 1-list and 2-list to increase the range of scores.

In the second pilot study, we administered the revised version of List Sorting to 47 children ages 3 to 6 years. Twenty-two participants completed a retest within two weeks. We examined the range, mean, standard deviation, and test-retest reliability to determine task feasibility. Results indicated that children as young as 3 years of age could complete the initial items on the Toolbox List Sorting Working Memory Test. For the 1-list total score, the range of scores for 3- to 6-year-olds was 0–16 with an average score of 6.67 ($SD = 4.42$). Performance was lower for the 2-list component of the test as the range was 0–10 with an average score of 3.4 ($SD = 2.33$) which is expected given that the 2-list task is more challenging. Test-retest reliability coefficients were $r = .85$ and $r = .86$ respectively for both the 1-list and 2-list task, indicating that the performance was highly reliable. These pilot data indicated that children as young as 3 years of age were able to sequence items in size order. The results also helped us modify the List Sorting task further as we added some easier

items to further increase the range of scores. We also combined scores on the 1- and 2-list tasks to increase the range and variability of scores on this task.

In the third pilot study, we administered the Toolbox List Sorting Working Memory Test to three groups of children ages 3 to 4 years ($n = 35$), 5 to 7 ($n = 26$), and 8 to 14 years ($n = 28$). The goal was to fully examine the descriptive statistics of the Toolbox List Sorting Working Memory Test across different age bands along with a closer examination of the test-retest reliability of the test. Our results indicated that cognitive abilities across childhood improve. Performance on the combined 1-list and 2-list score in the 3- to 4-year-old children ranged from 5–28 with an average score of 18.3 ($SD = 6.5$). Performance in the 5- to 7-year-old group ranged from 9–37 with a mean of 28.4 ($SD = 5.5$). Finally, performance in the 8- to 14-year-old group ranged from 32–48 with a mean of 40.6 ($SD = 4.5$). Test-retest reliability was highest in the youngest group ($r = .90$), $r = .79$ in the 5- to 7-year-olds, and $r = .74$ in the 8- to 14-year-old group. The results again indicated that most children were able to understand the basic concepts of List Sorting. Further, test-retest reliability was best for the youngest ages.

The testing also allowed us to examine item difficulty and remove redundancy from the test so that we could streamline administration, shorten the length of the task, and drop poor performing items. The findings provided justification for removing a portion of practice items, and creating discontinuation rules for children that were unable to answer all practice items correctly (they would not be administered the test). The validation version of this test was prepared and the next section discusses the results of the validation study. The validation study included the full age range of the NIHTB, ages 3 to 85 years. This monograph focuses on the results of this study for children and young adolescents ages 3 to 15 years. The validation data from the adult and elderly populations will be published in a separate series of papers so that each population can be addressed in greater depth.

Method

Participants

The participants in the validation phase are described in detail in Weintraub et al. (Chapter 1, this volume). Briefly, the sample was 208 children ages 3 to 15 years (120 children ages 3 to 6 years and 88 children ages 8 to 15 years). Sample recruitment was distributed across age, gender, race, and highest parent education strata. A subset of 66 participants (approximately 32%) completed a retest 7 to 21 days later to assess test-retest reliability. Nine children (all age 6 and younger) did not successfully complete the task for reasons such as lack of attention or alertness or general noncompliance.

Measures

Participants were tested with the Toolbox List Sorting Working Memory Test as well as several additional tasks to provide convergent (the Developmental Neuropsychological Assessment, 2nd Edition Sentence Repetition subtest; the Wechsler Intelligence Scale for Children, Fourth Edition Letter-Number Sequencing subtest) and discriminant validity (the

Peabody Picture Vocabulary Test, 4th Edition, the Delis–Kaplan Executive Function System Color-Word Test, and the Wisconsin Card Sorting Test-64 Card Version).

The Toolbox List Sorting Working Memory Test—In this task, a list of stimuli is presented both visually (picture) and auditorily (recording of a one-word description of the stimulus) on a computer monitor, one at a time at a rate of 2 sec per stimulus, and participants are required to repeat all of the stimuli back to the examiner in order of increasing real-world size, from smallest to largest. On practice trials, participants are required to reorder and repeat the items in a 2-item list (e.g., List: pumpkin, lemon; Correct answer: “Lemon, pumpkin”), followed by a 3-item list.

In the first phase of the test (i.e., the 1-List phase), participants are first shown a list with 2 items drawn from a single category (i.e., food). If participants are correct on this 2-item list, the number of items in the list presented on the next trial increases by one item, up to a total of 7 items per list (i.e., list length ranges from a 2-item list to a 7-item list, for a total of six levels of list length). If participants err on a trial at a given list length, they receive another trial with the same number of items in the list; if they err on that trial, this phase of the test is discontinued. That is, the 1-List phase of the test is discontinued when two trials of the same list length are failed.

Following the 1-List phase, all participants proceed to the second phase of the test (the 2-List phase), in which they see lists of items drawn from two different categories (i.e., food and animals). Participants are instructed to reorder and repeat the stimuli first from one category, then the other, in order of size within each category. Lists in the 2-List phase start with a 2-item list and increase in number of items in the same way as in the 1-List phase (i.e., from a 2-item list to a 7-item list, for a total of six levels of list length). For both phases, for each list length, participants receive a score of 2 points if they are correct on the first trial. A second trial at a given list length is only administered when participants fail the first trial. Participants receive a score of 1 point only for a given list length if they fail the first trial at that list length but pass the second trial. Test scores consist of combined total trials correct on the 1-List and 2-List phases of the task. The test takes approximately 10 min to administer.

Convergent validity—The Developmental Neuropsychological Assessment, 2nd Edition Sentence Repetition (NEPSY-II Sentence Repetition; Korkman, Kirk, & Kemp, 2007) involves an examiner reading a series of sentences of increasing complexity and length. The participant is required to recall each sentence after each is presented. Participants ages 3 to 6 years completed this measure. For analysis we used the Sentence Repetition Total Score.

For children ages 8 to 15 years, the measure of convergent validity was the Wechsler Intelligence Scale for Children, 4th Edition (WISC-IV) Letter-Number Sequencing (Wechsler, 2008). In this test, participants are presented with a mixed list of numbers and letters, and their task is to repeat the list by saying the numbers first in ascending order and then the letters in alphabetical order. Scores reflect the number of correct responses (maximum 30 points), with higher scores indicating better performance.

Discriminant validity—The Peabody Picture Vocabulary Test, 4th Edition (PPVT-4; Dunn & Dunn, 2007) provides a measure of expressive vocabulary and word retrieval. Examinees are asked to identify which of four pictures reflects a specific word. Scores are based on the number correct (maximum 228). Participants ages 3 to 15 years completed this measure

The Delis–Kaplan Executive Function System Color-Word test (D-KEFS Color-Word; Delis, Kaplan, & Kramer, 2001) is based on the Stroop procedure and taps the participants’ ability to inhibit overlearned verbal responses. Specifically, the participant is timed during his or her (1) naming of color patches; (2) reading basic color words printed in black ink; and (3) naming the color of the ink in which color words are printed. In the last condition, the colors of the ink and the printed color words differ from each other. Participants ages 8–15 years completed this measure. For this study, we examined scores on the Color-Word interference score.

The Wisconsin Card Sorting Test-64 Card Version (WCST-64; Kongs, Thompson, Iverson, & Heaton, 2000) is a shortened, 64-card version of the Wisconsin Card Sorting Test, which assesses the ability to shift sets using visual stimuli that are easily verbally mediated. It requires participants to sort pictured cards into piles according to changing rules. Successful completion of the test relies on having a number of intact cognitive functions including attention, working memory, and visual processing. Participants ages 8–15 years completed this measure. We examined perseverative errors for this study.

Data Analysis

This study examines associations of the Toolbox List Sorting Working Memory Test scores with age, test-retest reliability, and convergent and discriminant validity. Pearson correlation coefficients between age and Toolbox List Sorting Working Memory Test performance were calculated to describe the developmental-related associations for each measure. Intraclass correlation coefficients (ICC) were calculated to evaluate test-retest reliability and convergent and discriminant validity. Convergent validity was assessed with correlations between the Toolbox List Sorting Working Memory Test and a well-established “gold standard” measures of the same construct (i.e., NEPSY-II Sentence Completion and WISC-IV Letter Number Sequencing); evidence of discriminant validity consisted of lower correlations with selected gold standard measures of a *different* cognitive construct: receptive vocabulary (PPVT-4) and executive function (DKEFS Color-Word and WCST)

Results

Age Effects

Figure 2 provides a graphic presentation of performance on the List Sorting from ages 3–15 years. Across the ages, age was related to performance on the test ($df=199$, $r = .77$, $p<0.0001$), and a quadratic model provided the best fit of the data, with $R^2 = .66$. Positive associations between age and test performance were also seen for ages 3–6 years ($df=111$, $r = .52$, $p<0.0001$) and for ages 8–15 years ($df=86$, $r = .33$, $p=0.002$). Pairwise comparisons between age groups are reported in Appendix A.

Test-Retest Reliability

Overall *ICC* for the Toolbox List Sorting Working Memory Test (ages 3–15 years) was .86 (95% *CI* = .78, .91), which was higher than the test-retest reliability for NEPSY Sentence Completion for ages 3 to 6 years (*ICC* = .80, 95% *CI* = .65, .89), and WISC-IV Letter Number Sequencing for ages 8 to 15 years. (*ICC* = .80, 95% *CI* = .52, .87).

Effect of Repeated Testing

Practice effects were computed as the difference between test and retest normalized scaled scores, with significance of the effect being tested with *t* tests for dependent means. For the total child group (ages 3–15 years, *n* = 66), the Toolbox List Sorting Working Memory Test showed no practice effect over an average 2-week test–retest interval: mean practice effect = 0.15, *SD* = 1.68, *t*(65) = .72, *p* = .48.

Construct Validity

Table 1 shows results for convergent and discriminant validity. Correlations for convergent validity were .57 for both ages 3 to 6 years and 8 to 15 years (all *p* < .0001), suggesting that the List Sorting, NEPSY-II Sentence Completion (3–6 years), and WISC-IV Letter Number Sequencing (8–15 years) tasks tap a similar construct, an indication of the concurrent validity with similar working memory tasks. The correlations with the discriminant validity measure (the PPVT-4/receptive vocabulary) were *r* = .63 (*df*=110, *p*<0.0001) in the younger children and *r* = .45 (*df*=85, *p*<0.0001) in the older children indicating an overlap with this construct. Discriminant correlations did not differ significantly from the corresponding convergent correlations.

Discussion

In this chapter we described the development and validation of a new measure of working memory, List Sorting, for the NIH Toolbox, with specific emphasis on how this measure was developed and adapted for use in a pediatric population. Stimuli were prepared in an auditory and visual modality so that the List Sorting Test would yield a general working memory score (independent of modality) and would not distinguish between specific structural working memory components like the phonological loop or visuospatial sketchpad.

As noted above, most children were able to understand the basic sequencing tenets of List Sorting. List Sorting performance was correlated with age. That is, performance on the task improved throughout childhood and early adolescence. This is consistent with the anticipated developmental trajectory of working memory (Dempster, 1981; Gathercole, Pickering, Ambridge, & Wearing, 2004; McAuley & White, 2011). Further, test-retest reliability was relatively high when computed across the entire 3- to 15-year age range.

In addition to reliability, the Toolbox List Sorting Working Memory Tests also showed adequate convergent validity. Correlation coefficients were in the moderately high range when compared with criteria measures purported to measure working memory. Further, correlations with receptive vocabulary were moderate for the 3- to 6-year-old group showing

that, for this younger group, the working memory task is correlated with verbal functioning. This likely represents general intelligence in the youngest children where specific domains of cognitive functioning are less defined (see Mungas et al., Chapter 7, this volume). The moderate correlations with traditional measures of executive function, as well as verbal functioning, in the 8–15 year olds also demonstrate that the working memory task is correlated with measures of general functioning, however, the correlations are somewhat lower than that with other working memory tasks (which suggest that the Toolbox List Sorting Working Memory Test has adequate convergent and discriminant validity).

In addition to our findings, it is also important to review some of the limitations of this study. First, the small sample sizes utilized within age bands make it difficult to evaluate test-retest correlations within age subgroups. This aspect of evaluation of the instrument will be remedied in the next phase of the task. Further, because stimuli for the List Sorting task utilized both auditory and visual images, we are not able to distinguish between the specific components of the phonological loop or visuospatial sketchpad (Baddely & Hitch, 1974).

Regardless of these limitations, the Toolbox List Sorting Working memory Test presents a number of strengths. First, it is child-friendly and engaging, and it is short and easy to administer. Further, it is reliable and has demonstrated high content validity. It also requires size-order sequencing of both 1- and 2-category trials, which ensures that processing demands remain challenging throughout the lifespan (i.e., into adulthood). This is especially important given that other common measures of working memory (e.g., backward digit span tasks) is not as challenging for adults as it is for children. This fact implies that such tasks are appropriate measures of working memory in childhood, but not in adulthood (where the task reflects short-term memory rather than working memory; St Clair-Thompson, 2010).

Following the norming of the Toolbox List Sorting Working Memory Test future studies will be employed to examine the sensitivity of List Sorting to neurological insult (e.g., traumatic brain injury). Ultimately, the Toolbox List Sorting Working Memory Test promises to provide a measure of working memory that is useful over the lifespan.

References

- Alloway TP, Gathercole SE. How does working memory work in the classroom? *Educational Research and Reviews*. 2006; 1:134–139.
- Alloway TP, Gathercole SE, Kirkwood H, Elliott J. The cognitive and behavioural characteristics of children with low working memory. *Child Development*. 2009; 80:606–621. [PubMed: 19467014]
- Alloway TP, Gathercole SE, Pickering SJ. Verbal and visuospatial short-term memory and working memory in children: are they separable? *Child Development*. 2006; 77:1698–1716. [PubMed: 17107455]
- Baddeley AD. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*. 2000; 4:417–423. [PubMed: 11058819]
- Baddeley, AD.; Hitch, GJ. Working memory. In: Bower, GH., editor. *The psychology of learning and motivation*. Vol. 8. New York, NY: Academic Press; 1974. p. 147-189.
- Barrouillet P, Bernardin S, Camos V. Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*. 2004; 133:83–100. [PubMed: 14979753]
- Best JR, Miller PH, Jones LL. Executive functions after age 5: Changes and correlates. *Developmental Review*. 2009; 29:180–200. [PubMed: 20161467]

- Briscoe J, Rankin PM. Exploration of a 'double-jeopardy' hypothesis within working memory profiles for children with specific language impairment. *International Journal of Language and Communication Disorders*. 2009; 44:236–250. [PubMed: 18821159]
- Bull R, Scerif G. Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Developmental Neuropsychology*. 2001; 19:273–293. [PubMed: 11758669]
- Bunge SA, Klingberg T, Jacobsen RB, Gabrieli JDE. A resource model of the neural basis of executive working memory. *Proceedings of the National Academy of Science*. 2000; 97:3573–3578.
- Camos V, Barrouillet P. Developmental change in working memory strategies: From passive maintenance to active refreshing. *Developmental Psychology*. in press
- Carlson SM. Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*. 2005; 28:595–616. [PubMed: 16144429]
- Case R. The structure and process of intellectual development. *International Journal of Psychology*. 1987; 22:571–607.
- Casey BJ, Galvan A, Hare TA. Changes in cerebral functional organization during cognitive development. *Current Opinion in Neurobiology*. 2005; 15:239–244. [PubMed: 15831409]
- Chein JM, Moore AB, Conway ARA. Domain-general mechanisms of complex working memory span. *NeuroImage*. 2011; 54:550–559. [PubMed: 20691275]
- Conlin JA, Gathercole SE, Adams JW. Children's working memory: Investigating performance limitations in complex span tasks. *Journal of Experimental Child Psychology*. 2005; 90:303–317. [PubMed: 15777923]
- Cowan, N. Working memory capacity. New-York (NY: Psychology Press; 2005.
- Cowan N. Multiple concurrent thoughts: The meaning and developmental neuropsychology of working memory. *Developmental Neuropsychology*. 2010; 35:447–474. [PubMed: 20721769]
- Crane PK, Narasimhalu K, Gibbons LE, Pedraza O, Mehta KM, Tang Y, et al. Composite scores for executive function items: demographic heterogeneity and relationships with quantitative magnetic resonance imaging. *Journal of the International Neuropsychological Society*. 2008; 14(5):746–759. [PubMed: 18764970]
- Crowe SF. Does the letter number sequencing task measure anything more than digit span? *Assessment*. 2000; 7(2):113–117. [PubMed: 10868248]
- Dempster FN. Memory span: sources of individual and developmental differences. *Psychological Bulletin*. 1981; 89:63–100.
- Dennis M, Spiegler BJ, Hoffman HJ, Hendrick EB, Humphrey RP, Becker LE. Brain tumors in children and adolescents—I. Effects on working, associative and serial-order memory of IQ, age at tumor onset and age of tumor. *Neuropsychologia*. 1991; 29:813–827. [PubMed: 1944880]
- D'Esposito M, Postle BR, Ballard D, Lease J. Maintenance versus manipulation of information held in working memory: An event-related fMRI study. *Brain and Cognition*. 1999; 41:66–86. [PubMed: 10536086]
- Engle, RW.; Kane, MJ.; Tuholski, SW. Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In: Miyake, A.; Shah, P., editors. *Models of working memory*. Cambridge: Cambridge University Press; 1999. p. 102-134.
- Fry AF, Hale S. Processing speed, working memory, and fluid intelligence: Evidence for developmental cascade. *Psychological Science*. 1996; 7:237–241.
- Fry AF, Hale S. Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*. 2000; 54:1–34. [PubMed: 11035218]
- Garon N, Bryson SE, Smith IM. Executive functions in preschoolers: A review using an integrative framework. *Psychological Bulletin*. 2008; 134:31–60. [PubMed: 18193994]
- Gathercole SE, Alloway TP, Willis C, Adams A-M. Working memory in children with reading disabilities. *Journal of Experimental Child Psychology*. 2006; 93:265–281. [PubMed: 16293261]
- Gathercole, SE.; Hitch, GJ. Developmental changes in short-term memory: A revised working memory perspective. In: Collins, A.; Gathercole, SE.; Conway, MA.; Morris, PE., editors. *IT>Theories of memory*. Hove, England: Erlbaum; 1993. p. 189-210.

- Gathercole SE, Pickering SJ, Ambridge B, Wearing H. The structure of working memory from 4 to 15 years of age. *Developmental Psychology*. 2004; 40:177–190. [PubMed: 14979759]
- Geier CF, Garver K, Terwilliger R, Luna B. Development of working memory maintenance. *Journal of Neurophysiology*. 2009; 101:84–99. [PubMed: 18971297]
- Gold JM, Carpenter C, Randolph C, Goldberg TE, Weinberger DR. Auditory Working Memory and Wisconsin card sorting test performance in schizophrenia. *Archives of General Psychiatry*. 1997; 54:159–165. [PubMed: 9040284]
- Haut MW, Kuwabara H, Leach S, Arias RG. Neural activation during performance of letter number sequencing. *Applied Neuropsychology*. 2000; 7(4):237–242. [PubMed: 11296686]
- Hawkins KA. Indicators of brain dysfunction derived from graphic representations of the WAIS-III/WMS-III technical manual samples: A preliminary approach to clinical utility. *The Clinical Neuropsychologist*. 1998; 12:535–551.
- Hitch GJ, Halliday MS. Working memory in children. *Philosophical Transactions of the Royal Society of London, Series B*. 1983; 302:324–340.
- Holmes J, Gathercole SE, Dunning DL. Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*. 2009; 12:F9–F15. [PubMed: 19635074]
- Kail R, Hall LK. Distinguishing short-term memory from working memory. *Memory & Cognition*. 2001; 29:1–9. [PubMed: 11277452]
- Kwon H, Reiss AL, Menon V. Neural basis of protracted developmental changes in visuo-spatial working memory. *Proceedings of the National Academy of Science*. 2002; 99:13336–13341.
- Levin HS, Hanten G, Zhang L, Swank PR, Ewing-Cobbs L, Dennis M, Barnes MA, Max J, Schachar R, Chapman SB, Hunter JV. Changes in working memory after traumatic brain injury in children. *Neuropsychology*. 2004; 18:240–247. [PubMed: 15099146]
- Martin TA, Donders J, Thompson E. Potential of and problems with new measures of psychometric intelligence after traumatic brain injury. *Rehabilitation Psychology*. 2000; 45(4):402–408.
- McAuley T, White DA. A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology*. 2011; 108:453–468. [PubMed: 20888572]
- McNab F, Klingberg T. Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*. 2008; 11:103–107.
- Miyake A, Friedman NP, Rettinger DA, Shah P, Hegarty M. How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*. 2001; 130:621–640. [PubMed: 11757872]
- Miyake A, Friedman NP, Emerson AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*. 2000; 41:49–100. [PubMed: 10945922]
- Miyake, A.; Shah, P., editors. *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press; 1999.
- Munakata Y. Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*. 2001; 5:309–315. [PubMed: 11425620]
- Mungas D, Reed BR, Tomaszewski Farias S, DeCarli C. Criterion-Referenced Validity of a Neuropsychological Test Battery: Equivalent Performance in Elderly Hispanics and Non-Hispanic Whites. *Journal of the International Neuropsychological Society*. 2005; 11:620–630. [PubMed: 16212690]
- Nelson CA, Monk CS, Lin J, Carver LJ, Thomas KM, Truwit CL. Functional neuroanatomy of spatial working memory in children. *Developmental Psychology*. 2000; 36:109–116. [PubMed: 10645748]
- Nettelbeck T, Burns NR. Processing speed, working memory and reasoning ability from childhood to old age. *Personality and Individual Differences*. 2010; 48:379–384.
- Nevo E, Breznitz Z. Assessment of working memory components at 6 years of age as predictors of reading achievements a year later. *Journal of Experimental Child Psychology*. 2011; 109:73–90. [PubMed: 21115182]

- O'Hare E, Lu LH, Houston SM, Bookheimer SY, Sowell ER. Neurodevelopmental changes in verbal working memory load-dependency: An fMRI investigation. *NeuroImage*. 2008; 42:1678–1785. [PubMed: 18586110]
- Osaka N, Osaka M, Kondo H, Morishita M, Fukuyama H, Shibasaki H. The neural basis of executive function in working memory: an fMRI study based on individual differences. *NeuroImage*. 2004; 21:623–631. [PubMed: 14980565]
- Pickering, SJ.; Gathercole, SE. *Working Memory Test Battery for Children*. London: Psychological Corporation; 2001.
- Postle BR. Working memory as an emergent property of the mind and brain. *Neuroscience*. 2006; 139:23–38. [PubMed: 16324795]
- The Psychological Corporation. *WAIS–III—WMS–III technical manual*. San Antonio: Author; 1997.
- Tulsky, DS.; Saklofske, DH.; Zhu, J. Revising a standard: An evaluation of the origin and development of the WAIS-III. In: Tulsky, DS.; Saklofske, DH.; Chelune, GJ.; Heaton, RK.; Ivnik, R.J.; Bornstein, R.; Prifitera, A.; Ledbetter, MF., editors. *Clinical Interpretation of the WAIS–III and WMS–III*. San Diego: Academic Press; 2003. p. 43-92.
- Reznick JS, Morrow JD, Goldman BD, Snyder J. The onset of working memory in infants. *Infancy*. 2004; 6:145–154.
- Shing LY, Lindenberger U, Diamond A, Li S-C, Davidson MC. Memory maintenance and inhibitory control differentiate early childhood to adolescence. *Developmental Neuropsychology*. 2010; 35:679–697. [PubMed: 21038160]
- Smith-Spark JH, Fisk JE. Working memory functioning in developmental dyslexia. *Memory*. 2007; 15:34–56. [PubMed: 17479923]
- St Clair-Thompson HL. Backwards digit recall: A measure of short-term memory or working memory? *European Journal of Cognitive Psychology*. 2010; 22:286–296.
- Thomason ME, Race E, Burrows B, Whitfield-Gabrieli S, Glover GH, Gabrieli JDE. Development of spatial and verbal working memory capacity in the human brain. *Journal of Cognitive Neuroscience*. 2008; 21:316–332. [PubMed: 18510448]
- Thorell LB, Lindqvist S, Nutley SB, Bohlin G, Klingberg T. Training and transfer effects of executive functions in preschool children. *Developmental Science*. 2009; 12:106–113. [PubMed: 19120418]
- Towse JN, Hitch GJ, Hutton U. A reevaluation of working memory capacity in children. *Journal of Memory and Language*. 1998; 39:195–217.
- Unsworth N, Engle RW. The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*. 2007; 114:104–132. [PubMed: 17227183]
- Vicari S, Caravale B, Carlesimo GA, Cascadi AM, Allemand F. Spatial working memory deficits in children at ages 3–4 who were low birth weight, preterm infants. *Neuropsychology*. 2004; 18:673–678.
- Wager TD, Smith EE. Neuroimaging studies of working memory: A meta-analysis. (2003). *Cognitive, Affective, & Behavioral Neuroscience*. 2003; 3:255–274.
- Wechsler, D. *WAIS-III administration and scoring manual*. San Antonio: The Psychological Corporation; 1997.
- Wechsler, D. *WISC-IV administration and scoring manual*. San Antonio: PsychCorp; 2003.
- Wechsler, D. *WAIS-IV administration and scoring manual*. San Antonio: Pearson; 2008.
- Wiebe SA, Espy KA, Charak D. Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*. 2008; 44:575–587. [PubMed: 18331145]
- Wiebe SA, Sheffield T, Nelson JM, Clark CAC, Chevalier N, Espy KA. The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*. 2011; 108:436–452. [PubMed: 20884004]
- Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of Attention-Deficit/Hyperactivity Disorder: A meta-analytical review. *Biological Psychiatry*. 2005; 57:1336–1346. [PubMed: 15950006]

Williams DL, Goldstein G, Minshew NJ. The profile of memory function in children with autism. *Neuropsychology*. 2006; 20:21–29. [PubMed: 16460219]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

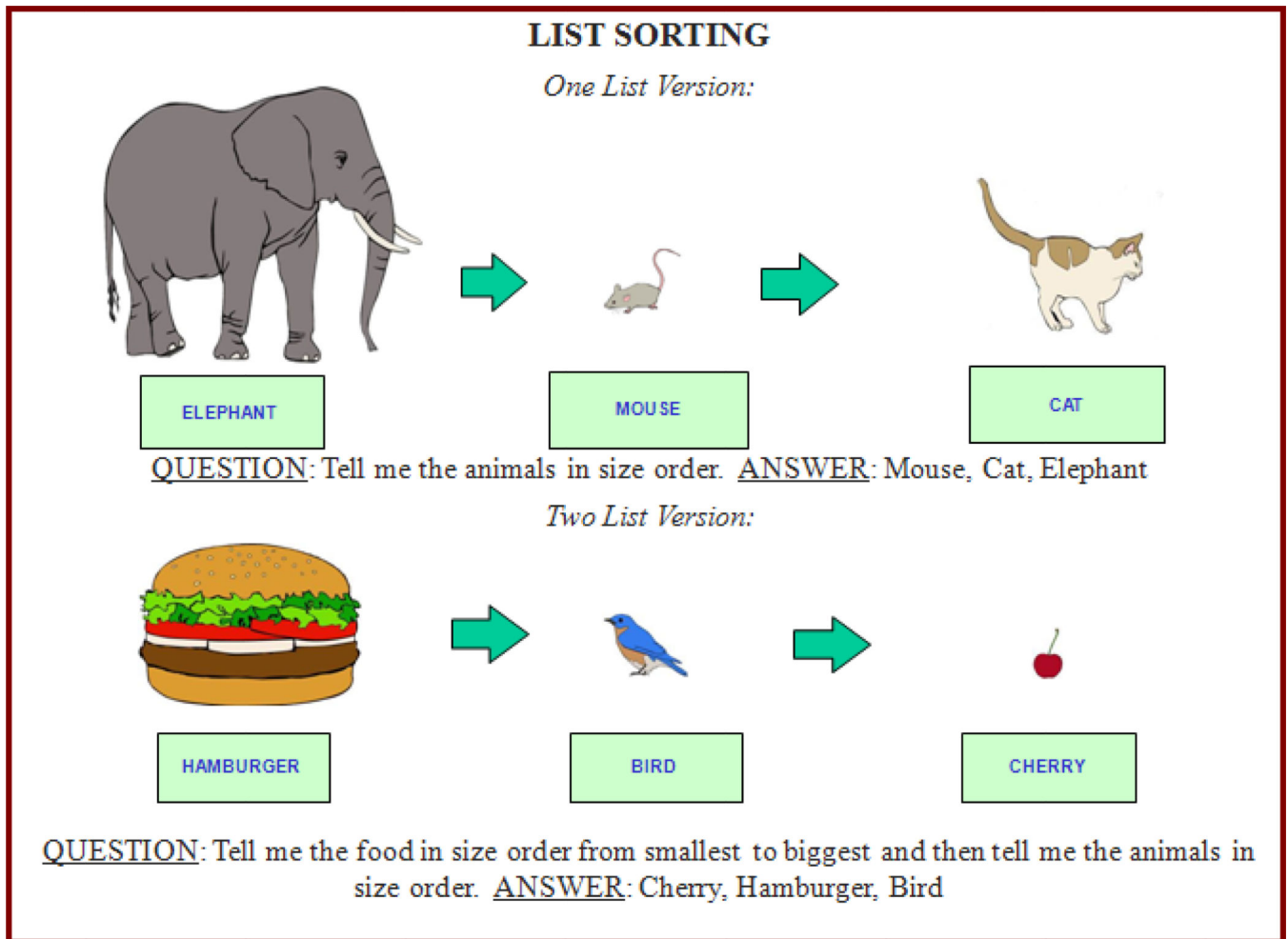


Figure 1. Examples of One-List and Two-List List Sorting Task

Legend: 1-List List Sorting requires participants to sequence items according to a single category, whereas 2-List List Sorting requires sequencing that involves an alternation between two different categories.

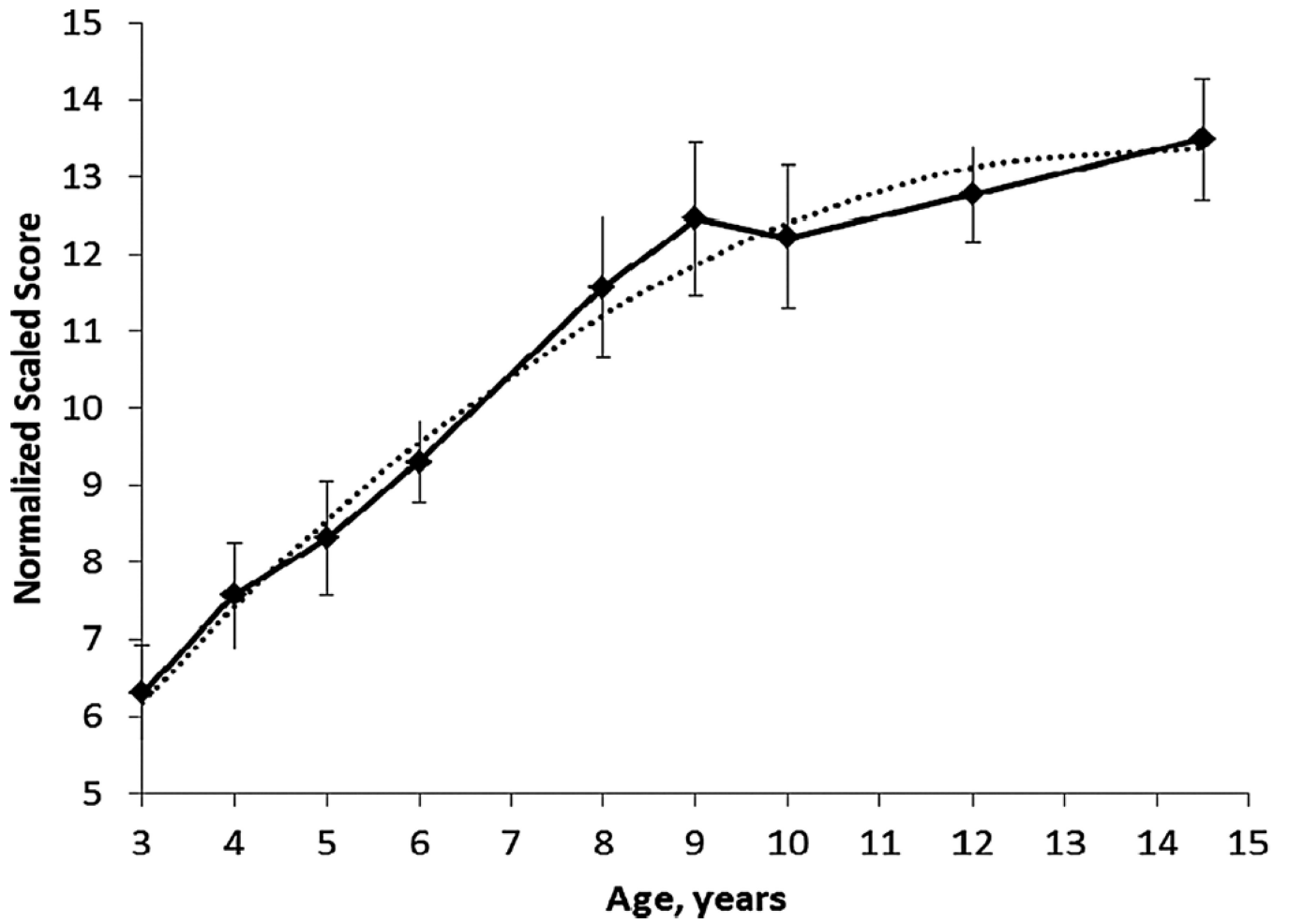


Figure 2. Toolbox List Sorting Working Memory Test scores by age (means \pm 2 standard errors).

Table 1

Convergent & Discriminant Validity of Toolbox List Sorting Working Memory Test

| | Ages 3 to 6 years | | Age 8 to 15 years | |
|---------------------------------------|-------------------|----------|-------------------|----------|
| | <i>df</i> | <i>r</i> | <i>df</i> | <i>r</i> |
| Convergent Validity Measures | | | | |
| NEPSY-II Sentence Completion | 107 | .57 | - | - |
| WISC-IV Letter-Number Sequencing | - | - | 83 | .57 |
| Discriminant Validity Measures | | | | |
| PPVT-4 | 110 | .63 | 85 | .45 |
| D-KEFS Color-Word Interference | - | - | 84 | .45 |
| WCST-64 Perseverative Errors | - | - | 85 | .42 |

Note: *r* = Pearson's Correlation Coefficient; WPPSI = Wechsler Preschool and Primary Scale of Intelligence – 3rd Edition; PPVT-4 = Peabody Picture Vocabulary Test – 4th Edition; DKEFS = Delis-Kaplan Executive Function System; WCST-64 = Wisconsin Card Sorting Test-64 Card Version; Unadjusted scaled scores

All *p* .0001

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript