Title

III. NIH TOOLBOX COGNITION BATTERY (CB): MEASURING EPISODIC MEMORY

Permalink

https://escholarship.org/uc/item/6fq0z800

Journal

Monographs of the Society for Research in Child Development, 78(4)

ISSN

0037-976X

Authors

Bauer, Patricia J
Dikmen, Sureyya S
Heaton, Robert K
et al.

Publication Date

2013-08-01

DOI

10.1111/mono.12033

Peer reviewed

Published in final edited form as:
Monogr Soc Res Child Dev. 2013 August ; 78(4): 34–48. doi:10.1111/mono.12033.

III. NIH TOOLBOX COGNITION BATTERY (CB): MEASURING EPISODIC MEMORY

Patricia J. Bauer, **Sureyya S. Dikmen**, **Robert K. Heaton**, **Dan Mungas**, **Jerry Slotkin**, and **Jennifer L. Beaumont**

Abstract

One of the most significant domains of cognition is episodic memory, which allows for rapid acquisition and long-term storage of new information. For purposes of the NIH Toolbox, we devised a new test of episodic memory. The nonverbal NIH Toolbox Picture Sequence Memory Test (TPSMT) requires participants to reproduce the order of an arbitrarily ordered sequence of pictures presented on a computer. To adjust for ability, sequence length varies from 6 to 15 pictures. Multiple trials are administered to increase reliability. Pediatric data from the validation study revealed the TPSMT to be sensitive to age-related changes. The task also has high test–retest reliability and promising construct validity. Steps to further increase the sensitivity of the instrument to individual and age-related variability are described.

In this chapter, we introduce the NIH Toolbox Picture Sequence Memory Test (TPSMT), a measure developed as a test of episodic memory for ages 3– 85 years. Episodic memory permits rapid learning and retention of new information. It is the basis for formation of memories of the mundane—such as where the car was parked last—to the special events that constitute one's life story or personal past.

Subdomain Definition

Evidence from nonhuman animals, patient populations, and typically developing children and adults makes clear that memory is not a unitary construct. Rather, it is comprised of different systems of information encoding, storage, and retrieval. One major distinction is the dichotomy between maintenance of information over the short term, and over the long term (long-term memory). Short-term or working memory is discussed in Tulsky and colleagues (Chapter 5, this volume); long-term memory is the subject of the current chapter. Within long-term memory there is a distinction between procedural (or implicit, or nondeclarative) memory and declarative (or explicit) memory (e.g., Squire, Knowlton, & Musen, 1993; Tulving, 2000). Procedural, implicit, or nondeclarative memories guide behavior but seemingly require no effort to retrieve and are not accessible to consciousness. Work with animal models and human patients makes clear that the procedural memory system remains relatively intact in aging and is virtually unaffected by neurological conditions and diseases that target the medial temporal lobe structures involved in the declarative memory system (Churchill, Stanis, Press, Kushelev, & Greenough, 2003; Reber, Martinez, & Weintraub, 2003).

In contrast to procedural memory, declarative or explicit memory is effortful and involves conscious recollection of information that is potentially verbally accessible. The declarative memory system is further divided into semantic and episodic memory (e.g., Squire, 2004).

Corresponding author: Patricia J. Bauer, Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA 30322, patricia.bauer@emory.edu.

Semantic memory is specialized for storage of timeless, placeless facts, concepts, and the vocabulary to describe them. Knowledge stored in semantic memory typically is long lasting, which is the basis for reference to semantic memory as “crystallized intelligence.” Semantic memory is resistant to decline with age and to neurological insult. The construct of semantic memory is measured by the Toolbox Picture Vocabulary Test and the Toolbox Oral Reading Recognition Test (see Gershon et al., Chapter 4, this volume). In contrast, episodic memory is specialized for storage of unique events or experiences encoded in a time-specific manner. Episodic memory is fragile and time-limited (though some memories, such as those of special, personally relevant events, may endure over long periods of time), and it is sensitive to decay and interference, as well as to both normal aging and many brain diseases.

Importance During Childhood

Episodic memory allows for rapid, even one-trial, learning of new information and for retention of information for later retrieval. As such, it provides the building blocks for cognitive growth during development and throughout the lifespan. Its importance to mental life is nowhere more in evidence than in the historic case of the patient HM, who as a result of surgery to the medial–temporal lobe, lost the capacity for forming new episodic memories (e.g., Corkin, 2002). As well, the relatively protracted course of development of episodic memory is a major source of one of the most robust phenomena in the memory literature, namely, infantile or childhood amnesia: the relative paucity among adults of memories of unique events for the first 3–4 years of life (see Bauer, 2007, for a review). Thus, this cognitive ability is critical for achieving a concept of self that is continuous over time, for independence, education, and success in personal and professional activities of daily life throughout the lifespan. The centrality of episodic memory explains why it is the most frequently measured form of memory and why it is included in the CB.

The course of development of episodic memory is protracted, with pronounced changes throughout the first two decades of life (see Bauer, 2007; Bauer, Larkina, & Deocampo, 2011, for reviews). In infancy, the ability is measured using nonverbal tasks, such as elicited and deferred imitation (props are used to produce a specific action or sequence of actions that the infant is permitted to imitate either immediately, after a delay, or both; e.g., Bauer & Shore, 1987; Bauer & Mandler, 1989). Use of imitation-based tasks in the first 3 years of life has revealed age-related increases in the length of time over which memory is preserved, in the robustness of memory, and in the reliability with which it is observed (e.g., Bauer, Wenner, Dropik, & Wewerka, 2000). Developmental changes are especially apparent in the ability to remember the temporal order of events (e.g., Bauer et al., 2000), making this aspect of episodic memory a target for the CB.

Episodic memory continues to develop throughout childhood and into adolescence. There are age-related increases in the amount of information that children remember. For example, relative to younger children, older children remember longer lists of items (see, e.g., Bjorklund, Dukes, & Brown, 2009, for a review), thus making list length a prime target for exploitation in tests of episodic memory designed for wide age ranges (see below). With development, children’s memory becomes more deliberate and strategic, with resulting increases in the organization that children impose on to-be-remembered material (e.g., Bjorklund et al., 2009). In addition, children become more aware of their own and others’ memory processes (i.e., increases in metamemory), enabling them to recruit information-processing resources in the service of increased memory demands. As a result, we may expect to see age-related increases in the amount of information children are able to remember, and in their ability to bring organization to it.

The overall importance of episodic memory to typical development is brought into stark relief in pediatric populations in whom the function is impaired. The most striking example is a population who as infants and very young children, sustained damage to the medial–temporal lobe structures that support declarative memory (see below). Without exception, individuals in this population of so-called *developmental amnesics* experienced difficulty learning in school and deficits on episodic memory tasks (Gadian et al., 2000). Importantly, the damage sustained by this population is confined to the neural structures involved in declarative memory in general and episodic memory in particular, described next.

Relations of Domain With Brain Function

Encoding, storage, and retrieval of episodic memories depend on a multi-component network involving the temporal lobe (including hippocampus and surrounding cortices) and other cortical area (including prefrontal cortex and limbic/temporal association areas) (e.g., Eichenbaum & Cohen, 2001; Zola & Squire, 2000). Specifically, the process of encoding of new memories begins as the elements that constitute an event register across primary sensory areas (auditory, somatosensory, visual). Inputs from the primary cortices are projected to unimodal association areas, where they are integrated into whole percepts of what objects sound, feel, and look like. Unimodal association areas in turn project to polymodal prefrontal, posterior, and limbic association cortices where inputs from the different sense modalities are integrated and maintained over brief delays (seconds; e.g., Petrides, 1995). For maintenance beyond the short term, the inputs must be stabilized or consolidated, a task attributed to medial temporal structures, in concert with cortical areas (McGaugh, 2000). Eventually, new traces become stabilized, permitting long-term storage in the neocortex. The prefrontal cortex is implicated in memory retrieval (e.g., Cabeza et al., 2004; Maguire, 2001). Demands on the temporal–cortical network are especially high when tasks require free recall versus recognition, and memory for temporal order information versus for items alone (e.g., Shimamura, Janowsky, & Squire, 1990). As discussed earlier, these conditions are those under which the most pronounced age-related differences are observed, thus informing design of the Toolbox measure for episodic memory.

Each of the brain regions involved in episodic memory, as well as the connections between them, undergoes substantial postnatal developmental change that extends well into the second decade of life (see Bauer, 2008, for a review). Throughout childhood and into adolescence there are gradual increases in hippocampal volume (e.g., Gogtay et al., 2004; Pfluger et al., 1999; Utsunomiya, Takano, Okazaki, & Mistudome, 1999) and in myelination in the hippocampal region (Arnold & Trojanowski, 1996; Benes, Turtle, Khan, & Farol, 1994; Schneider, Il'yasov, Hennig, & Martin, 2004). In the prefrontal cortex, pruning of synapses to adult levels does not begin until late childhood; adult levels are not reached until late adolescence or even early adulthood (Huttenlocher, 1979; Huttenlocher & Dabholkar, 1997). Although there are well-documented reciprocal connections between the hippocampus and frontal lobes, their development has not been fully elucidated (see Barbas, 2000; Fuster, 2002). Finally, it is not until adolescence that neurotransmitters such as acetylcholine reach adult levels (Benes, 2001).

Developmental changes in the neural structures and network that support episodic memory may be expected to have functional consequences. Consistent with this expectation, Sowell, Delis, Stiles, and Jernigan (2001) reported relations between structural changes in the medial–temporal and frontal regions, as measured by MRI, and performance on behavioral tests of memory. Children with structurally more mature medial–temporal lobe regions performed at higher levels on spatial memory tasks. Children with structurally more mature frontal cortices performed at higher levels on verbal and spatial memory tasks. Increases in myelination also are correlated with functional changes (Olesen, Nagy, Westerberg, &

Klingberg, 2003). Developmental studies of the magnitude and patterns of neural activation during episodic memory task performance are rare and have focused on encoding only. Ofen et al. (2007) found that in regions of prefrontal cortex associated with successful encoding, activations increased with age. Although medial–temporal activations were associated with successful encoding, they were not age-related (although see Menon, Boyett-Anderson, & Reiss, 2005).

Among children with known compromise of the medial–temporal structures, there are clear deficits on episodic memory tasks. For example, infants born with low iron stores as a result of failed regulation of maternal glucose during pregnancy (i.e., poorly controlled maternal gestational diabetes) show impaired performance on imitation-based tests of episodic memory from infancy through the preschool years (e.g., DeBoer, Wewerka, Bauer, Georgieff, & Nelson, 2005; Riggins, Miller, Bauer, Georgieff, & Nelson, 2009). Individuals with developmental amnesia provide another example: When tested as adolescents or adults, they show impaired performance on standard tests of episodic memory as well as age-appropriate analogues of imitation-based tasks (Adlam, Vargha-Khadem, Mishkin, & de Haan, 2005). In summary, the neural substrate that supports episodic memory is relatively well understood. Compromised development in the implicated structures, especially in the medial–temporal components of the network, is associated with impaired performance.

Toolbox Picture Sequence Memory Test

To measure episodic memory as part of the CB, we developed a new measure, the NIH TPSMT. The measure is derived from imitation-based tasks (elicited and deferred imitation) developed by Bauer and her colleagues for research with pre- and early-verbal infants and young children (e.g., see Bauer, 2005, 2006, 2007, for descriptions and discussion). For infant and young child populations, the stimuli are three-dimensional props used to produce sequences of action that the infant or child imitates. For the CB, the stimuli are sequences of pictured objects and activities presented on a computer screen.

The TPSMT developed for the CB addresses several needs. First, it provides a measure of episodic memory for children below 5 years of age. Based on the success of imitation-based measures even in infancy, the decision was made to adjust the procedure upward to cover the full age range (3–85 years). Development of the NIH TPSMT also addresses a need for tests that can be readily used with nonEnglish speakers. Finally, multiple alternate forms of the TPSMT can easily be created to reduce practice effects in longitudinal studies. In the validation study described in this chapter, three alternate forms were included. Total administration time is approximately 10 min. The validation study included the full age range of the NIH Toolbox, ages 3–85 years.

METHOD

Participants

The participants in the validation study are described in detail in Weintraub et al. (Chapter 1, this volume; Table 2).

Measure Development

The task stimuli for the NIH TPSMT are sequences of pictured objects and activities presented on a computer screen (see Figure 4). The objects and activities are thematically related but with no inherent order. The general themes are “Working on the farm,” “Playing at the park,” and “Going to the fair.” Although there are no inherent constraints on the order in which the activities in the sequences must occur, the pictures are presented in a specific order that the participant must remember and then reproduce. This requirement is made clear in verbal instructions as well as through practice sequences (see below). The level of difficulty of the task for different age ranges was determined during pilot testing. The number of pictures the participants were required to order ranged from 6-picture sequences to 15-picture sequences.

For the test itself, color-illustrated pictures appear one at a time on the computer monitor in a fixed order. Each picture originally is displayed in the center of the computer screen, as a 3″ × 5″ image. As it appears, a recording briefly describes the content of each picture. The duration of presentation for each picture is 2.2 sec. Once described, the picture reduces in size and is translocated to its position in the sequence (requiring 1.5 sec), making way for the next picture, until all pictures in a sequence have been displayed. After 3 sec during which the entire sequence is shown, the pictures are placed in a random spatial array at the center of the screen. For the validation study described in this report, participants used a touch screen to “move” each picture to its correct location in the sequence. Participants were permitted as much time as necessary to complete their responses. Three trials were administered to improve test score variability and test–retest reliability (Strauss, Sherman, & Spreen, 2006); the same sequence was presented on each trial.

For all ages, practice sequences are administered prior to administration of the first test trial to orient participants to the TPSMT task and to provide experience moving the pictures to the correct position in the sequence. For the validation study, for young children, the first practice sequence involved 3 pictures that followed a logical order. Specifically, the sequence involved putting a cake into an oven, frosting it, and then putting candles on the cake. The logical sequence of pictures was intended to reinforce for children the importance of placing the pictures in the correct temporal order. After the pictures in the practice sequence were displayed, children were given practice ordering the pictures. Errors were corrected. After successful completion of the first practice sequence, children were administered two additional practice trials that were 3 and 4 pictures in length, respectively. The 3- and 4-picture sequences had no inherent constraints on their order. Participants ages 5 years and older received one 3-picture and one 4-picture practice sequence; none of their practice sequences had inherent constraints on their order.

Based on pilot testing the following sequence lengths were administered to the different age groups: For ages 3–4 years: 6 pictures; 5–6 years: 9 pictures; 8 years: 12 pictures; 9–60 years: 15 pictures, 65–85 years: 9 pictures. In the next phase of measure development, we will test possible means of adjusting sequence length on-line, in response to participants' performance (see Discussion Section).

Scoring

The participant's score on the TPSMT is derived from the cumulative number of adjacent pairs of pictures remembered correctly over three learning trials. Adjacent pairs are two adjacent pictures placed in consecutive, ascending order. Thus, pictures placed in the orders 1–2 and 2–3 would receive credit, whereas pictures placed in the orders 1–3 and 3–15 would not receive credit (because they are not consecutive). For example, for a 6-picture sequence, the sequence of placement 1–2–3–5–4–6, would result in a score of 2: one point would be awarded for each of the adjacent pairs 1–2 and 2–3. No points would be earned for the correctly ordered pairs 3–5 and 4–6 because they items are nonadjacent. For each trial, the possible number of adjacent pairs is the number of pictures in the sequence, −1. The total possible number of adjacent pairs is the sum of the adjacent-pairs scores across trials.

Validation Measures

Convergent validity was assessed with correlations between TPSMT and established measures of memory. Evidence of discriminant validity was tested with correlations with a validation measure of a *different* cognitive construct, namely vocabulary. We expected the correlations with other memory measures to be high and with vocabulary to be lower.

Convergent Validity—One impetus for development of the TPSMT was that there did not exist a single measure that could be used across the age span of 3–85 years. In fact, neither was there a single measure that could be used within the target age span of 3–15 years. As a result, for purposes of examination of convergent validity, different validation measures were used for ages 3–6 years and ages 8–15 years. For children 3–6 years of age, we selected the sentence repetition subtest of the Developmental Neuropsychological Assessment, 2nd Edition (NEPSY-II; Korkman, Kirk, & Kemp, 2007). The NEPSY-II Sentence Repetition subtest involves an examiner reading a series of sentences of increasing complexity and length. The participant is required to recall each sentence after it is presented. The measure was selected because of its good psychometric properties within the target age range and because it is relatively brief to administer. The subtest was administered and scored using the standard protocol. For analysis we used the Sentence Repetition Total Score.

For children 8–15 years, two different measures were used, one using verbal and the other visuospatial information: the Rey Auditory Verbal Learning Test (RAVLT) and the Brief Visuospatial Memory Test-Revised (BVMT-R; Benedict, 1997). These measures were selected because they have good psychometric properties within the target age range and combined, they sample multiple modalities of learning. In the RAVLT, an examiner reads aloud a list of 15 words, at the rate of one word per second. The test-taker's task is to repeat as many words as possible, in any order. The test was administered and scored using the standard protocol, with the exception that only three trials, rather than five, were administered. In analysis, we used the total score of the three learning trials. The BVMT-R is designed to measure visuospatial memory. The examinee views six geometric figures on a page after which the figures are removed from view. The examinee then is asked to draw as many of the figures as possible from memory in their correct location. The test was administered and scored using the standard protocol. The score used in analysis was the total score of the three learning trials. Scores on the RAVLT and BVMT-R were strongly correlated with one another, $r(n = 83) = .35$, $p < .001$. Therefore, to simplify data analysis, for the 8 to 15 year olds, we created a combined validation memory score, which was the mean of performance on the two separate measures.

Discriminant Validity Measure—The Peabody Picture Vocabulary Test, 4th Edition (PPVT-IV; Dunn & Dunn, 2007) was used as a discriminant measure at all ages. The PPVT-IV is a test of receptive vocabulary and is often used as a proxy for full scale IQ or general developmental level. On each trial, a set of 4 pictures is provided along with a word describing one of the pictures. The examinee is asked to point to or say the number of the picture that best corresponds to the word. The test was administered and scored using the standard protocol. The PPVT-IV was selected in part because it has good psychometric properties within the target age range. In addition, it was attractive as a measure because it could be administered across the entire age range for the battery (ages 3–85 years), allowing for use of the same metric for all ages.

Data Analysis

As described in Zelazo et al. (Chapter 2, this volume), normalized scaled scores were used for all analyses. Pearson correlation coefficients between age and test performance were calculated to assess the ability of the TPSMT for detecting developmental growth during childhood. Intraclass correlation coefficients (*ICC*) with 95% confidence intervals were calculated to evaluate test–retest reliability. Convergent validity was assessed with correlations between the TPSMT and the NEPSY-II (ages 3–6 years), and the TPSMT and the mean of the RAVLT and BVMT-R (8–15 years); discriminant validity was assessed with correlations between the TPSMT and PPVT-IV scores (all ages).

RESULTS

Age Effects

The NIH TPSMT provides a valid test of age-related differences in learning and episodic memory, as evidenced by strong associations between the TPSMT and age. As depicted in Figure 5, scores on the TPSMT increased with age. Pairwise comparisons between age groups are reported in Appendix A. Across the 3- to 15-year age span, the correlation with age was $r(202) = .78$, $p < .001$. A quadratic model provided the best fit of the data, with $R^2 = .70$. For the 3- to 6-year age group, the correlation was $r(115) = .69$, $p < .001$; within this group, the correlation between age and performance on the sentence repetition subtest of the NEPSY-II was .58. Thus, the TPSMT proved a nominally stronger relation with age than the validation measure. For the 8-to 15-year age group, the correlation with age was $r(85) = .26$, $p = .016$. Within this group, the correlation between age and performance on the average validation measure was $r(85) = .44$, $p < .001$. Thus, for the 8–15 year olds, relative to the validation measure, the TPSMT was not as sensitive to age-related improvements in performance.

Test–Retest Reliability

The test–retest reliability of the TPSMT was excellent in the full sample of children age 3–15 years, as evidenced by a high intraclass correlation: $ICC = .76$ (95% confidence interval: .64–.85; $n = 66$).

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 TPSMT showed a practice effect over an average 2-week test–retest interval: *mean practice effect* = .99, $SD = 1.88$, $t(65) = 4.27$, $p < .0001$.

Construct Validity

Convergent Validity—

For children 3–6 years of age, the TPSMT was moderately correlated with the Sentence Repetition subtest of the NEPSY-II (the validation measure): $r(110) = .50$, $p < .001$. For children 8–15 years of age, the TPSMT was moderately correlated with the mean of the two gold-standard measures of memory (RAVLT and BVMT-R): $r(84) = .47$, $p < .001$. These correlations provide evidence of adequate convergent validity.

Discriminant Validity—

For children 3–6 years of age, the correlation between the TPSMT and the PPVT-IV was higher than expected: $r(112) = .58$, $p < .001$, and nominally higher than the correlation with the convergent validity measure. However, the validation NEPSY-II scores also were correlated with PPVT-IV scores: $r(109) = .67$, $p < .001$. Thus, the problem with discriminant validity does not appear to be unique to the TPSMT in the young age group. In the older age group, the correlation between the TPSMT and the PPVT-IV was low, as expected: $r(84) = .28$, $p = .009$.

Discriminant correlations did not differ significantly from the corresponding convergent correlations.

DISCUSSION

The NIH TPSMT is a new measure of learning and episodic memory developed for use across the age span of 3–85 years. It is based on the nonverbal elicited or deferred imitation task designed for use with preverbal infants and early-verbal children. The TPSMT proved to be a valid assessment of age-related changes in episodic learning and memory, as indexed by a strong correlation with age. In spite of the potential for artifactual relations with age as a result of the confound between age and sequence length, the pattern of performance was consistent with true age effects. That is, because sequence length increased with age, the maximum number of correctly ordered adjacent pairs that the children could receive also increased with age: 3 and 4 year olds were tested on sequence length 6; 5 and 6 year olds were tested on sequence length 9; 8 year olds were tested on sequence length 12; and 9–15 year olds were tested on sequence length 15. As a result, the maximum scores that children could earn over three trials were 15, 24, 33, and 48, respectively. Examination of Figure 5 shows that TPSMT scores increased in a linear fashion from age 3 to 9 years, without evidence of departures from linearity at ages where the sequence length changed (e.g., age 4 years vs. age 5 years). This is supportive of a true age effect. Examination of effect sizes between age groups (available from the authors) also was consistent with a true age effect.

The TPSMT also showed strong test/retest reliability within the sample of children ages 3–15 (*ICC* = .76). The assessment of convergent and discriminant validity also was quite promising, with one notable exception involving the younger age group. For the older age group of 8–15 years, the correlation of the TPSMT with the combined validation measure of memory was moderate and higher than the correlation with the measure of vocabulary, as expected. For the younger age group, however, the TPSMT correlated equally highly with the validation measures of memory and also of vocabulary. The lack of discriminant validity was not a problem unique to the TPSMT, however: the validation measure of memory (Sentence Repetition of the NEPSY-II) was even more strongly correlated with the vocabulary test (i.e., the PPVT-IV) than was the TPSMT. As discussed in more detail in Mungus et al. (Chapter 7, this volume), there is evidence of the gradual differentiation of cognitive skills over the course of development, and it is difficult to detect variance associated with any single domain of function within the 3- to 6-year age range.

The NIH TPSMT shows substantial promise as a means of assessment of learning and episodic memory within the 3- to 15-year age range; as described in Weintraub et al. (2013), it also shows promise for young to older adults. Norming will provide a substantial, empirical resource for normative expectations for performance at different ages and for different demographically defined groups. Several changes will be made to the instrument in the norming phase. First, rather than using a touch-screen, participants will register their responses on the TPSMT using a mouse. The change in administration will be made to lessen the demand for specialized equipment for the CB. Participants who are uncomfortable using a mouse will be encouraged to point to the screen and the test administrator will operate the mouse. As such, the change is expected to have a negligible impact on performance. Second, in an effort to avoid floor effects, the norming study will feature different practice trials that are expected to more effectively convey task requirements. This change is expected to aid the youngest children and the elderly, in particular. Third, also in an effort to further improve the sensitivity of the instrument—this time to avoid ceiling effects—18-step sequences have been developed for potential use with participants who achieve perfect or near-perfect performance on shorter sequence lengths. In the norming study, some participants 8–60 years of age will be tested on a 15-step sequence on Trial 1 and 18-step sequences on Trials 2 and 3, whereas others will experience 15-step sequences on each test trial (the procedure followed in the present research). A similar scheme will be implemented for the shorter sequence lengths as well (for participants younger than 8 years and older than 60 years). The resulting data will be used to create item response functions and to explore the feasibility of computerized adaptive testing (CAT) that would permit selection of sequence length based on a participant's own performance.

In conclusion, NIH TPSMT is a newly developed measure of episodic memory for use across the 3- to 85-year age range. It requires participants to reproduce the order of a sequence of pictures presented on a computer. To adjust for ability, sequence length varies from 6 to 15 pictures. In the next phase of testing of the instrument (i.e., the norming study), data will be collected that may permit construction of a computerized adapative testing (CAT) version of the test. The data reported in this chapter are the results of validation testing with children ages 3–15 years. Within this age range, the TPSMT is sensitive to age-related changes in learning and episodic memory and also has high test–retest reliability and promising construct validity. The TPSMT thus appears to have strong potential for use in cross-sectional and longitudinal research throughout childhood to early adolescence.

References

Adlam ALR, Vargha-Khadem F, Mishkin M, de Haan M. Deferred imitation of action sequences in developmental amnesia. Journal of Cognitive Neuroscience. 2005; 17:240–248. [PubMed: 15811236]

Arnold SE, Trojanowski JQ. Human fetal hippocampal development: I. Cytoarchitecture, myeloarchitecture, and neuronal morphologic features. Journal of Comparative Neurology. 1996; 367:274–292. [PubMed: 8708010]

Barbas H. Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. Brain Research Bulletin. 2000; 52:319–330. [PubMed: 10922509]

Bauer, PJ. New developments in the study of infant memory. In: Teti, DM., editor. Blackwell Handbook of research methods in developmental science. Oxford, UK: Blackwell; 2005. p. 467-488.

Bauer PJ. Constructing a past in infancy: A neuro-developmental account. Trends in Cognitive Sciences. 2006; 10:175–181. [PubMed: 16537115]

Bauer, PJ. Remembering the times of our lives: Memory in infancy and beyond. Mahwah, NJ: Erlbaum; 2007.

Bauer PJ. Toward a neuro-developmental account of the development of declarative memory. Developmental Psychobiology. 2008; 50:19–31. [PubMed: 18085555]

Bauer, PJ.; Larkina, M.; Deocampo, J. Early memory development. In: Goswami, U., editor. The Wiley-Blackwell handbook of childhood cognitive development. 2. Oxford, UK: Wiley-Blackwell; 2011. p. 153-179.

Bauer PJ, Mandler JM. One thing follows another: Effects of temporal structure on one- to two-year-olds' recall of events. Developmental Psychology. 1989; 25:197–206.

Bauer PJ, Shore CM. Making a memorable event: Effects of familiarity and organization on young children's recall of action sequences. Cognitive Development. 1987; 2:327–338.

Bauer PJ, Wenner JA, Dropik PL, Wewerka SS. Parameters of remembering and forgetting in the transition from infancy to early childhood. Monographs of the Society for Research in Child Development. 2000; 65(4 Serial No 263)

Benedict, R. Brief visuospatial memory test-revised professional manual. Odessa, FL: Psychological Assessment Resources; 1997.

Benes, FM. The development of prefrontal cortex: The maturation of neurotransmitter systems and their interaction. In: Nelson, CA.; Luciana, M., editors. Handbook of developmental cognitive neuroscience. Cambridge, MA: MIT Press; 2001. p. 79-92.

Benes FM, Turtle M, Khan Y, Farol P. Myelination of a key relay zone in the hippocampal formation occurs in the human brain during childhood, adolescence, and adulthood. Archives of General Psychiatry. 1994; 51:477–484. [PubMed: 8192550]

Bjorklund, DF.; Dukes, C.; Brown, RD. The development of memory strategies. In: Courage, ML.; Cowan, N., editors. The developoment of memory in infancy and childhood. New York: Taylor & Francis; 2009. p. 145-175.

Cabeza R, Prince SE, Daselaar SM, Greenberg DL, Budde M, Dolcos F, et al. Brain activity during episodic retrieval of autobiographical and laboratory events: An fMRI study using a novel photo paradigm. Journal of Cognitive Neuroscience. 2004; 16:1583–1594. [PubMed: 15622612]

Churchill JD, Stanis JJ, Press C, Kushelev M, Greenough WT. Is procedural memory relatively spared from age effects? Neurobiology of Aging. 2003; 24:883–892. [PubMed: 12927770]

Corkin S. What's new with the amnesic patient H.M.? Nature Reviews. 2002; 3:153–160.

DeBoer T, Wewerka S, Bauer PJ, Georgieff MK, Nelson CA. Explicit memory performance in infants of diabetic mothers at 1 year of age. Developmental Medicine and Child Neurology. 2005; 47:525–531. [PubMed: 16108452]

Dunn, LM.; Dunn, DM. Peabody picture vocabulary test. 4. San Antonio, TX: Pearson; 2007.

Eichenbaum, H.; Cohen, NJ. From conditioning to conscious recollection: Memory systems of the brain. New York: Oxford University Press; 2001.

Fuster JM. Frontal lobe and cognitive development. Journal of Neurocytology. 2002; 31:373–385. [PubMed: 12815254]

Gadian DG, Aicardi J, Watkins KE, Porter DA, Mishkin M, Vargha-Khadem F. Developmental amnesia associated with hear hypoxic-ischaemic injury. Brain. 2000; 123:499–507. [PubMed: 10686173]

Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, et al. Dynamic mapping of human cortical development during childhood through early adulthood. Proceedings of the National Academy of Science United States of America. 2004; 101:8174–8179.

Huttenlocher PR. Synaptic density in human frontal cortex: Developmental changes and effects of aging. Brain Research. 1979; 163:195–205. [PubMed: 427544]

Huttenlocher PR, Dabholkar AS. Regional differences in synaptogenesis in human cerebral cortex. Journal of Comparative Neurology. 1997; 387:167–178. [PubMed: 9336221]

Korkman, M.; Kork, U.; Kemp, S. NEPSY-II. 2. San Antonio, TX: Harcourt Assessment; 2007.

Maguire EA. Neuroimaging studies of autobiographical event memory. Philosophical Transactions of the Royal Society of London. 2001; 356:1441–1451. [PubMed: 11571035]

McGaugh JL. Memory—A century of consolidation. Science. 2000; 287:248–251. [PubMed: 10634773]

Menon V, Boyett-Anderson JM, Reiss AL. Maturation of medial temporal lobe response and connectivity during memory encoding. Cognitive Brain Research. 2005; 25:379–385. [PubMed: 16122916]

Ofen N, Kao YC, Sokol-Hessner P, Kim H, Whitfield-Gabrieli S, Gabrieli JDE. Development of the declarative memory system in the human brain. Nature Neuroscience. 2007; 10:1198–1205.

Olesen PJ, Nagy Z, Westerberg H, Klingberg T. Combined analysis of DTI and fMRI data reveals a joint maturation of white and grey matter in a fronto-parietal network. Cognitive Brain Research. 2003; 18:48–57. [PubMed: 14659496]

Petrides M. Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in monkeys. The Journal of Neuroscience. 1995; 15:359–375. [PubMed: 7823141]

Pfluger T, Weil S, Wies S, Vollmar C, Heiss D, Egger J, et al. Normative volumetric data of the developing hippocampus in children based on magnetic resonance imaging. Epilepsia. 1999; 40:414–423. [PubMed: 10219266]

Reber PJ, Martinez LA, Weintraub S. Artificial grammar learning in Alzheimer's disease. Cognitive Affective Behavioral Neuroscience. 2003; 3:145–153.

Riggins T, Miller NC, Bauer PJ, Georgieff MK, Nelson CA. Consequences of low neonatal iron status due to maternal diabetes mellitus on explicit memory performance in childhood. Developmental Neuropsychology. 2009; 34:762–779. [PubMed: 20183732]

Schneider JFL, Il'yasov KA, Hennig J, Martin E. Fast quantitative difusion-tensor imaging of cerebral white matter from the neonatal period to adolescence. Neuroradiology. 2004; 46:258–266. [PubMed: 14999435]

Shimamura AP, Janowsky JS, Squire LR. Memory for the temporal order of events in patients with frontal lobe lesions and amnesic patients. Neuropsychologia. 1990; 28:803–813. [PubMed: 2247207]

Sowell ER, Delis D, Stiles J, Jernigan TL. Improved memory functioning and frontal lobe maturation between childhood and adolescence: A structural MRI study. Journal of International Neuropsychological Society. 2001; 7:312–322.

Squire LR. Memory systems of the brain: A brief history and current perspective. Neurobiology of Learning and Memory. 2004; 82:171–177. [PubMed: 15464402]

Squire LR, Knowlton B, Musen G. The structure and organization of memory. Annual Review of Psychology. 1993; 44:453–495.

Strauss, E.; Sherman, EMS.; Spreen, O. A compendium of neuropsychological tests: Administration, norms, and commentary. New York: Oxford University Press; 2006.

Tulving, E. Concepts of memory. In: Tulving, E.; Craik, FIM., editors. The Oxford handbook of memory. New York: Oxford University Press; 2000. p. 33-42.

Utsunomiya H, Takano K, Okazaki M, Mistudome A. Development of the temporal lobe in infants and children: Analysis by MR-based volumetry. American Journal of Neuroradiology. 1999; 20:717–723. [PubMed: 10319988]

Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Bauer PJ, et al. NIH Toolbox for the assessment of behavioral and neurological function: Cognition domain instruments. Neurology. 2013; 80:S54–S64. [PubMed: 23479546]

Zola, SM.; Squire, LR. The medial temporal lobe and the hippocampus. In: Tulving, E.; Craik, FIM., editors. The Oxford handbook of memory. New York: Oxford University Press; 2000. p. 485-500.

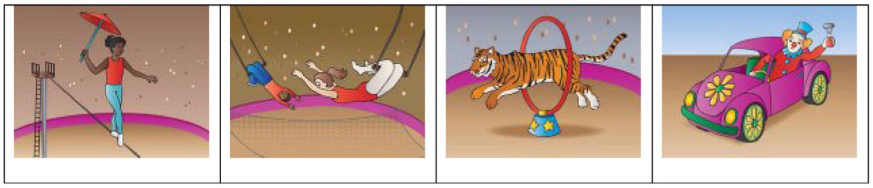

Figure 4.

Four-step practice sequence with “circus” theme: Walk a tightrope, swing on the trapeze, jump through the hoop, and drive the funny car.

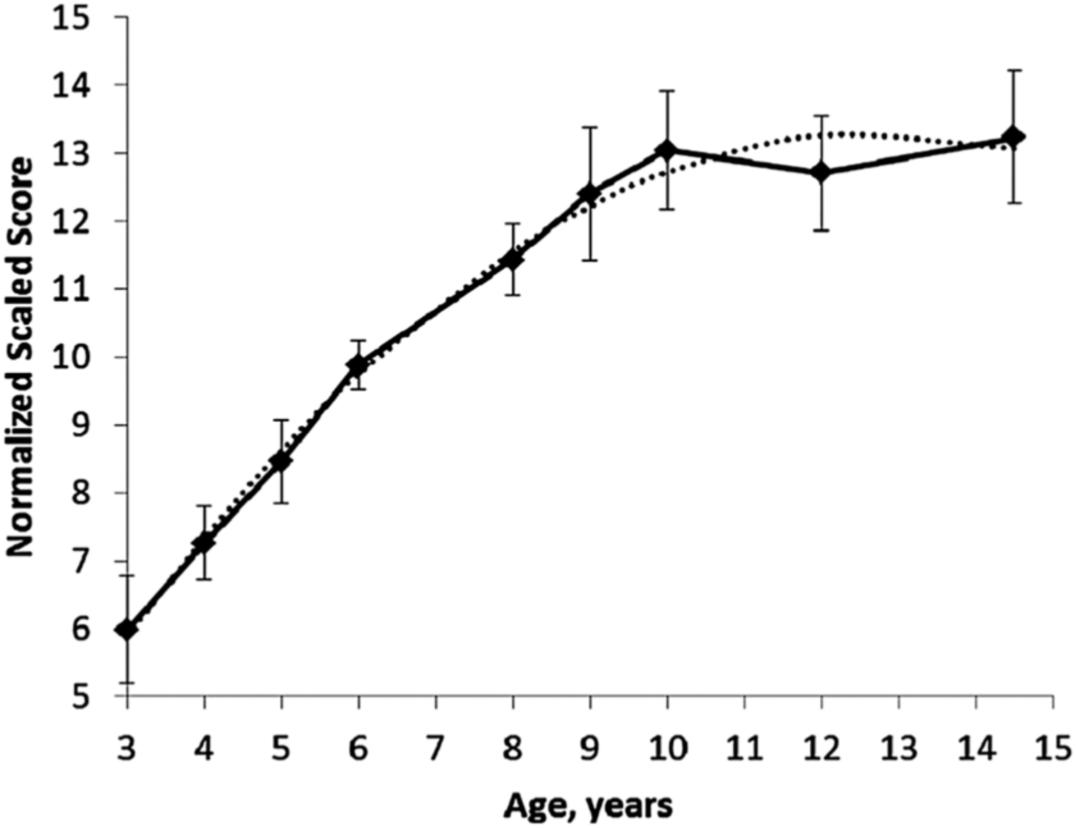

Figure 5.
Normalized scaled scores on the Toolbox Picture Sequence Memory Test across age groups. Error bars are ±2 standard errors. Best-fitting polynomial curve is also shown (see text).