TEST-RETEST RELIABILITY OF WORKING MEMORY SPAN TASKS: OPERATION SPAN, READING SPAN AND AUDITORY DIGIT SPAN

Chethana P

Registration No: 18AUD009

A Dissertation Submitted in Part Fulfillment for the Degree of Masters of Science

(Audiology)

University of Mysore

Mysuru

ALL INDIA INSTITUTE OF SPEECH AND HEARING,

MANASAGANGOTHRI, MYSORE- 570006

JULY, 2020

CERTIFICATE

This is to certify that this dissertation entitled **"Test-retest reliability of working memory span tasks: operation span, reading span and auditory digit span"** is a bonafide work submitted in part fulfilment for the degree of Master of Science (Audiology) of the student Registration No: 18AUD009. This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other university for the award of any diploma or degree.

Mysuru
July, 2020

Dr. M. Pushpavathi
Director
All India Institute of Speech and Hearing
Manasagangothri, Mysuru-570006

CERTIFICATE

This is to certify that this dissertation entitled **"Test-retest reliability of working memory span tasks: operation span, reading span and auditory digit span"** has been prepared under my supervision and guidance. It is also certified that this dissertation has not been submitted earlier to any other university for the award of any diploma or degree.

Mysuru

July, 2020

Dr. Ajith Kumar U

Guide

Professor in Audiology

Department of Audiology

All India Institute of Speech and Hearing

Manasagangothri, Mysuru- 570006

DECLARATION

This is to declare that this Master's dissertation entitled **"Test-retest reliability of working memory span tasks: operation span, reading span and auditory digit span"** is the result of my own study under the guidance of Dr. Ajith Kumar U, Professor in Audiology, Department of Audiology, All India Institute of Speech and Hearing Mysuru and has not been submitted in any other University for the award of any Diploma or Degree.

Mysuru | Register No.
July 2020 | **18AUD009**

ACKNOWLEDGEMENT

First and foremost, I would be grateful for almighty God for being with me like always.

Daddy, maa, and madh my constants, stand with me and behind me like a backbone throughout my life journey. There are no words to express for the love and blessings that you shower on me.

I would express my heartfelt gratitude to my guide, Dr. Ajith Kumar U. Sir, thanks for being so awesome and supporting me continuously throughout my academic learning. Though there were so many ups and downs, you have always made things better than how it was.

I would like to thank all the participants who have taken time off their busy schedules and helped me complete my work.

Prashanth sir, thanks for being so approachable and helping me handle the software independently.

Anup sir, thanks for all your inputs that helped me make my dissertation a better one.

Sreedevi mam, Vikas sir, Abhishek Sir, and Prajwal, my Gymkhana team, it was a beautiful experience to work with all of you. It is one to be remembered and cherished forever.

Thanks to all the faculties who guided and helped me. Teaching and non-teaching staff for all your support, seniors and juniors for all the love and encouragement that you gave throughout my six years of AIISH journey.

Special mention to my bunch of friends who always kept me lively Anki, Archu U, Appu, D/o Swamy, Jha, Kranthi, Kutta, Lavanya, Madhu, Mesim, Prassuu, Prawin, Riddhi, Sabin, Sru, Sumantha, and Thomas.

Deka, Akhil sir, and Danush sir, you were the constant seniors who kept me motivated and helped me with all sorts of issues.

Ajji, Sunanty, Mammu, and Jyo, thanks for making this 6-year journey of my life in Mysuru even more memorable from then till now.

Thanks to the sweetest bunch of juniors Chandu, Teju, Yeshu, Apurva, Jerusha, and Kruthika, for having a partnership in the shuttle matches. It was one of the best parts of the lockdown series.

My Bachelors' batch mates (as we did not have any batch name) thanks to all for the best memories that we made.

'Brainiacs'! I found so many new brains in these two years. Indeed it was a short span for all of us, but never short memories that I have bagged from you all. It was a fab time that I worked with all of you. A big thank you.

Mr. Gaaru and Suuu, how can I not mention both, my two best buddies who have so much strength to tolerate me and keep encouraging for whatever I do! It is a bliss to have you both.

I would also thank all those who have directly or indirectly helped me complete this dissertation.

Last but not least, Mysuru, the city that taught me life lessons, from the day when my journey started to this place and till now. I have no words to express what the city has made me today. The locations, food, people, weather all have been so perfect. Mysuru is not a city for me but an 'Emotion' – Agape. Thank you.

Table of Contents

List of Tables

List of Figures

Abstract

The present study aimed to investigate the test-retest reliability of the working memory span task: Operation span, Reading span and auditory digit span task in 50 participants. To assess the test- retest reliability of working memory within a day and across days. The test was conducted five times wherein test 1(T1) and test 2 (T2) was done within a day i.e., intra-session while test 3 (T3), test 4 (T4) and test 5 (T5) was done across the days (inter-session) with a gap of at least two days among each testing sessions. The forward and backward digit span tasks had range from acceptable to good internal consistency while the operation and reading span task internal consistency ranged from unacceptable to acceptable. Since the internal consistency for auditory digit span ranged from acceptable to good can be further used for clinical purpose.

Chapter I

Introduction

The term ‘working memory’ was initially devised in 1960 by Miller, Galanter, and Pribram in the book ‘Plans and the Structure of Behaviour’; later developed by (Atkinson & Shiffrin, 1968) publishing a paper and finally (Baddeley & Hitch, 1974) implemented this as the title for a Multi-component model. In layman's term, working memory can also be defined as a temporary storage system that helps to store and manage the information to perform a higher complex task such as thinking, reasoning and to improve intellectual ability.

The psychophysiological construct of Working Memory (WM), as described by Baddeley’s Multi-component model (Alan D Baddeley, 1992, 2000) postulates four main components – 'the central executive, the phonological loop, the visuospatial sketchpad, and the (newly added) episodic buffer.' The most critical component, the Central Executive, controls the overall allocation of attention to the task. The Phonological Loop and the Visuo-Spatial Sketchpad components serve as temporary storage units for speech-related and visual and spatial information. The Episodic Buffer acts as a 'binding agent' for the often multimodal information from the working memory systems different sources. The phonological loop element is where the rehearsal of the phrases to be remembered is most important; according to Baddeley's multi-component system (Baddeley, 1992, 2000a). The phonological loop component comprises of two sub-components – the phonological store and the articulatory rehearsal mechanism. The phonological store is the time-limited part of the working memory system as it is observed to undergo significant decays post the storage time of two seconds (Henry, 2012). However, by using ‘articulatory’ or ‘verbal’ rehearsal mechanisms before the decay time, the two-seconds of stored information re-enters the phonological loop, and the memory trace is refreshed. Hence, the overall ‘span’ of a listener’s memory is determined by the rate at which the trace fades and the rehearsal rate within the span of fading. Typically, working memory is measured by the use of 'working memory span tasks.' Some of the most commonly used working memory span tasks are forward digit and backward digit tasks, reading span, and operation span tasks. Among these digit span (forward and backward) are considered as simple tasks. These tasks primarily tap information storage and rehearsal. Whereas, operation and reading span tasks not only require information storage and rehearsal as "simple" measures of Short Term Memory Capacity, such as digit span or word span, but also on the simultaneous processing of additional information (Daneman & Carpenter, 1980; Case et al. 1982; Turner & Engle, 1989). Such working memory span tasks involves the introduction of target stimuli to be remembered, such as numbers or phrases, with the presentation of a challenging, secondary processing task, such as understanding phrases, or checking equations.

1.1 Need for the study

The Smriti- Shravan (2014) is an indigenously developed software package to assess the working memory span in Indian languages. It uses adaptive procedures to measure digit span tasks. Reading span and operation span tasks are measured as per the guidelines of Kane et al (2004). It is important to evaluate the reliability of working memory spans before they are used clinically. Furthermore, working memory spans are also used to track the effect of training regimes. In these context it is important to assess the reliability of these measures. Therefore, the current study aims to assess the reliability of four working memory spans: forward digit span, backward digit span, reading span and operation span using indigenously developed Smirti-Shravan software package.

1.2 Aim of the study

The aim of the present study is to find out reliability of working memory span testing.

1.3 Objectives of the study

1. To assess the test- retest reliability of working memory within a day.
2. To assess the test- retest reliability of working memory across days.

Chapter II

Review of Literature

Working memory is one of the most vital components that are used by everyone in their daily life. It helps us to perform complex tasks such as understanding a situation, reasoning for what could be done, and learning new things.

Working memory is also that sort of memory that could store less information, which could be easily assessable for a short duration. Originally the term “Working memory” arose from the study of computers. This term was referred to as the structure that was set up with their program to hold information temporarily and to execute the procedure that included such as solving geometric evidence (Newell & Simon, 1956). Similarly, the concept of temporary storage in humans was given by Miller, Galanter, and Pribram (1960) in their classic book ‘*Plans and the Structure of Behavior’,* which was used in routine life like solving problems. Also on animals wherein the animals had to hold information across several trials within the same (Olton, 1979).

The working memory was considered as the part of the mind which followed three patterns. First, it operates for success in life; second, it completes our goals and third complete our sub-goals. So, all these three are useful to store information and to execute these planned actions. For example, the goal of an individual is to achieve one's career, for that, they have a sub-goal to get an academic degree, for that they have to go to class, before that they have to get dressed so on, failing to perform these activities or forgetting the information would lead to errors (Adams et al., 2018).

2.1 Modal model

This model is an influential model that was a turning point in the evolution of theories of memory and threw a light on how data is processed in separate memory types. Atkinson & Shiffrin (1968) developed this model, as shown in figure 2.1 which included the concepts of the memory system, where the components were divided into structural elements and processes that control memory storage and, later, retrieve. The controlled memory process involves the flow of information, storage, retrieval, and decision-making, which also revolves around the concept of stores. These include three leading stores: sensory store, short-term store, and long-term store.

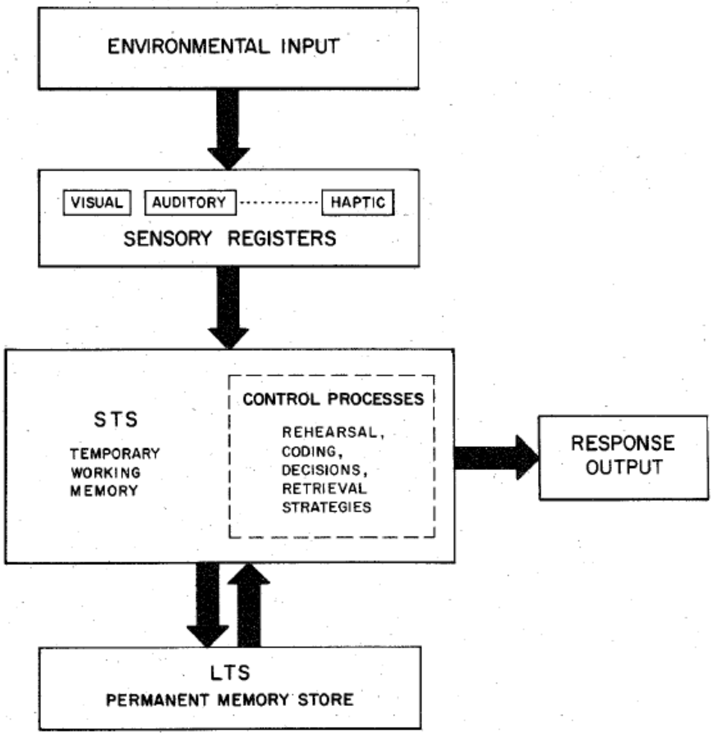
The information enters the sensory store or sensory registers from the environment. The sensory store is considered as inactive (i.e., subconscious) as it is the first phase of memory. This sort of memory retains data from less than a second to several seconds for a short period. The trace disappears very fast while this sort of memory has a huge capacity. The information entering the sensory store can be in three different modes: auditory (echoic) memory, the information may last between two and three seconds, but after 300 to 500ms, the trace begins to decrease (Pisoni, 1975; Cowan, 1984). Visual (iconic) memory when the stimulus is through visual form, this sort of memory does not last more than half a second (not more than 300ms) (Marzi et al., 1979). Similarly, if the data is tactile (haptic), this sort of memory is short-lived (< 2s) and has a duration and decay similar to iconic visual memory (Shih et al., 2009).

If the information is worth remembering, then it will be passed on to a temporary short-term store (STS). Short- term store is also known as the main or active memory. Generally, the information is stored for at least 15-18 sec (Peterson & Peterson, 1959). The short- term store contains some control process that is used in situations such as rehearsals where an individual has to remember and rehearse phone number until it has been written, to remember the number in the number plate of a vehicle which made and the act of hit-and-run or to rehearse one-time password until it has been entered the respective place and so on. Another component of the control process includes coding; this is done to enhance the retrieval and to remember the content using additional information; for example, to remember the spectrum's colors, we use mnemonic VIBGYOR where starting word indicates the color.

Similarly, the mnemonic ‘My Very Educated Mother Just Served Us Noodles,' where the starting letter of each word represents the planet's names. A few control processes also include making decisions, organizing, retrieval strategies, and problem-solving techniques. The information from STS may get decayed within 15 seconds (Atkinson & Shiffrin, 1968) if not rehearsed. If rehearsed, the data gets transferred to the next store, i.e., long- term store (LTS) here; the data almost stays permanently.

Figure 2.1

Represents the Working of Modal Model (Atkinson & Shiffrin, 1971)



The studies done by neuropsychological patients took the concept of the two-component model. Individuals whose medial temporal lobe was damaged had the grossly impaired capacity to learn new things while their short-term memory tasks were not affected (Baddeley & Warrington, 1970; Milner, 1966). Similarly, Shallice & Warrington (1970) studied working memory on individuals with conduction aphasia. It was thought that they had a specific deficit, which will affect short-term memory. However, there was a contradictory statement stating that if short-term memory system functions like working memory, those individuals should have problems with long-term memory and other complex cognitive tasks, but these tasks were not affected.

This model represented the working memory as a single mechanism to store the information temporarily. It included a few simple tasks in which a list of words will be presented, and one has to repeat it in verbatim. The most extended list that could be recalled and repeated correctly is the memory span. This model also gave importance to the information between the stores, which enriched the focus more on the short-term store. They distinguished the terms of short-term storage in which one has to recall the list of items without manipulating while working memory is defined when the information has to manipulate the stored information. For example, in a grocery list, when one has to remember and repeat back, that is called short-store storage. While the same list has to be repeated in a different order such as fruits and vegetables first, dairy products next and so on this would be considered as a test of working memory (Adams et al., 2018).

Short-term memory can maintain a limited quantity of data in the order of seconds. 7±2 components (Olichney & Hillert, 2004) (Miller, 1956) and 4 ± 1, on average (Cowan, 2001), can be memorized in chunks, and it can be used from the short term memory. Short-term memory should be differentiated from working memory, referring to structures and procedures used to store and manipulate data temporarily. Long-term memory, on the other hand, can retain an unlimited quantity of data.

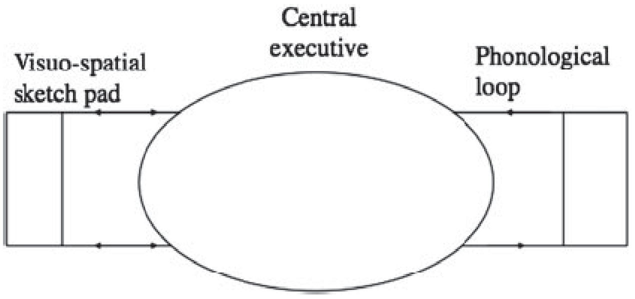
This model was criticized by Baddeley (1968). First, it was assumed that the information from short-term memory would surely get transferred to long-term memory, while the nature of the process was more crucial, more profound, and elaborated (Craik & Lockhart, 1972). Second, neuropsychological evidence showed that there was an inconsistency in the availability of information to long-term memory from short- term stores as individuals who had an absence of recency in the free recall should have had deficits in the short-term store that would have impaired the long-term memory also. However, this was not the situation. Third, the short-term store is assumed to be working memory, which plays an essential role in patients with significant mental deficits in cognition. Although, they had no problem with that, i.e., they were able to perform their tasks, which they were working earlier.

Similarly, the psychophysiological construct of working memory by Baddeley & Hitch (1974) defined working memory as a multi-component system to store information as it is not processed in a single process. Instead, it is broken into multiple boxes for representation. Baddeley (1966), conducted a test on recall of sequencing of five phonologically similar (man, mat, map, can, cat) versus different words such as (pit, day, cow, pen, sup). The study also included sequencing of semantically similar words (huge, big, wide, large, and tall) versus dissimilar (wet, soft, old, late, good). The test resulted by depicting a considerable effect of phonological similarity that is 80% sequencing was correct for dissimilar words, meanwhile 10% for similar words. The semantic similarity also showed small but significant scores, i.e., 71% versus 65%. When the same demonstration was, and the pattern was reversed when it was for long-term memory. Finally, it was concluded that there are two storage systems; short-term memory for phonological and long-term memory is semantically based. They also widely found based on studies that verbal long-term memory and language understanding, the working memory is not a single unitary store, instead of a three-component system.

The components include verbal (phonological store), visual and spatial information (visual-spatial store), and central controller processor (central executive), where each represented the part of the memory system.

Figure 2.2

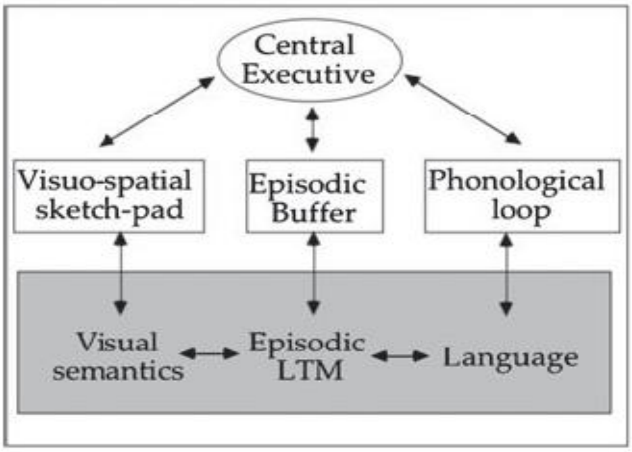
Represents the Model Working Memory Proposed in 1974



The central executive, the most significant component. The Phonological Loop, and the Visuo-Spatial Sketchpad components as shown in figure 2.2, serve as temporary storage units for speech-related and visual and spatial information. There was another component named ‘Episodic buffer’ (newly added), which acts as a ‘binding agent’ of the entire store for multimodal information within the working memory system as shown in figure 2.3. It holds the semantic information temporarily with the association of different kinds of information (Baddeley, 2000). Also, it helps to preserve the central executive's information by paying attention to help control cognition.

Figure 2.3

Represents the Later Development of the Multi-Component Model



The central executive, the most significant component, regulates the general attention distribution and problem-solving mission (Baddeley, 1998). It coordinates between the subcomponents processes, i.e., it supervises & coordinates the information retrieved from the two subsystems and further helps - Reasoning, comprehension of language, transfer of information to long-term memory through rehearsal, chunking and recovery.

The phonological loop element is where the rehearsal of the phrases to be remembered is most important; according to Baddeley's multi-component system (Baddeley, 1992, 2000). The phonological loop component is comprised of two sub-components – the phonological store and the articulatory rehearsal mechanism. The phonological store is the time-limited part of the working memory system as it is observed to undergo significant decays post the storage time of two seconds (Henry, 2012). However, by using ‘articulatory’ or ‘verbal’ rehearsal mechanisms before the decay time, the two-seconds of stored information re-enters the phonological loop, and the memory trace is refreshed. Hence, the overall ‘span’ of a listener’s memory is determined by the rate at which the trace fades and the rate of rehearsal within the span of fading.

The auditory data is encoded, rehearsed, and retained. It involves auditory data and vibrant observable phonetic indications from the speaker's face, recorded in a phonological form (Gathercole et al., 2008). When there is no sub-vocal rehearsal that enables the depiction of memory is declining. The visuospatial sketchpad is defined as the retention of the visual or spatial information over a short time.

Typically, working memory can be measured and assessed by the use of 'working memory span tasks.' Those include auditory digit span tasks, reading span, and operation span tasks. These tasks primarily tap information storage and rehearsal. Whereas, operation and reading span tasks require not only information storage and rehearsal (as do “simple” measures of Short-term Memory Capacity, such as digit span or word span), but also the simultaneous processing of additional information (Turner & Engle, 1989; Daneman & Carpenter, 1980).Such working memory span tasks involves the introduction of target stimuli to be remembered, such as numbers or phrases, with the presentation of a challenging, secondary processing task, such as understanding phrases, or checking equations.

`Chapter III

Method

Fifty adults in the age group of 18-35 (mean age = 21.22 years, SD = 3.151 years, Females = 25) years participated in the study. All the participants were native speakers of Kannada and were able to read and write Kannada. Before the commencement of the test, all the participants were provided with a consent form. Through administering a structured interview, it was determined that none of the participants had any complaint or history of otological disorders, neurological disorders, noise exposure, ototoxicity, or ear infection. The detailed audiological assessment was performed on all participants before recruiting them for the study.

The audiological evaluation consisted of otoscopy, otoacoustic emissions, pure-tone audiometry Tympanometry, and measurement of ipsilateral and contralateral acoustic reflex thresholds. All these participants had normal hearing sensitivity (less than 15 dB HL) at octave frequencies between 250 Hz and 8000 Hz for air conduction and between 250 Hz and 4000 Hz for bone conduction. All participants had “A” Type tympanogram with static compliance between 0.3 to 1.5cc and peak pressure between +60 and -100 daPa (Margolis & Heller, 1987)and normal ipsilateral as well as contralateral acoustic reflexes at 500, 1000, 2000 and 4000 Hz frequencies.

3.1 Test Environment

All audiological assessments were carried out in a sound-treated room with ambient noise levels within the permissible limits as per ANSI (ANSI B1. 1999). Working memory tests were carried out in a quiet room with minimal audio-visual distractions.

3.2 Procedure

3.2.1 Working Memory Assessment

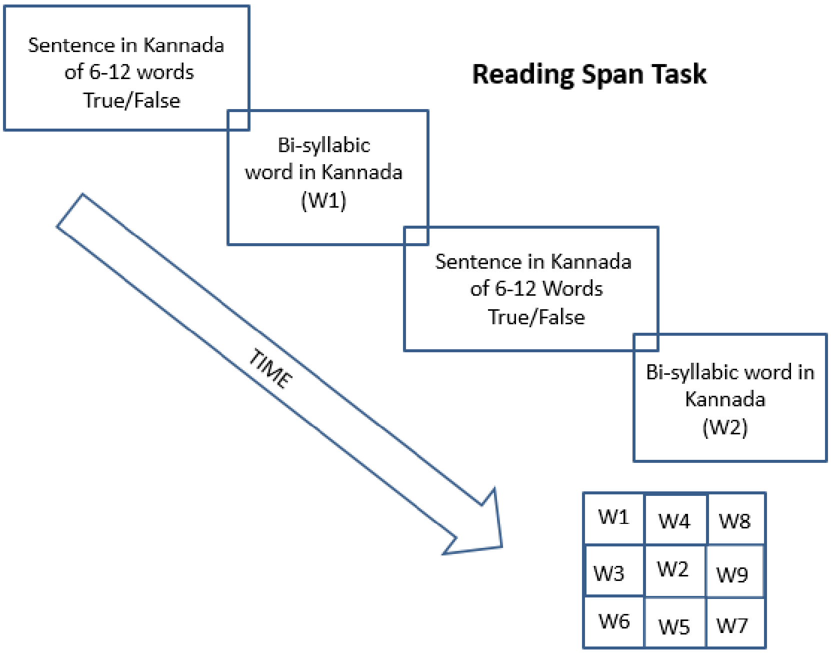
All the tasks for assessing working memory were conducted using Smriti- Shravan 3.0 software (Kumar & Sandeep, 2013). The assessment included reading span task, operation span task, and auditory digit span.

3.2.1.1 Reading Span Task. The participant's ability to remember the target stimuli, which interleaves with a secondary processing task, was evaluated. The secondary processing task was to judge the semantic/pragmatic correctness of a sentence Stimulus for the reading span task had been developed following the guidelines(Kane et al., 2004). The test was administered, which consisted a sentence to judge and a bi-syllable word to be remembered (e.g./rākṣasara//pūje//māḍuvudarinda//manasige//Śānti//hāgu/ /sukha//prāptiyāguttade/ followed by bi-syllabic word / Kāge/), the difficulty level of the targeted words was randomized such that the numbers of elements were unpredictable at the outset of an item as depicted in figure 3.1.

Guidelines recommended by (Conway et al., 1942) and(Kane et al., 2004)were followed during the scoring. A score of 1 was assigned for every word correctly recalled. At the end of each trial, the entire target words were shown along with the non-target words. Here participants were supposed to recall and choose all the target words in order, which was depicted in each trial in a correct sequence.

Figure 3.1

A Sample Representation of Stimuli in Reading Span Task

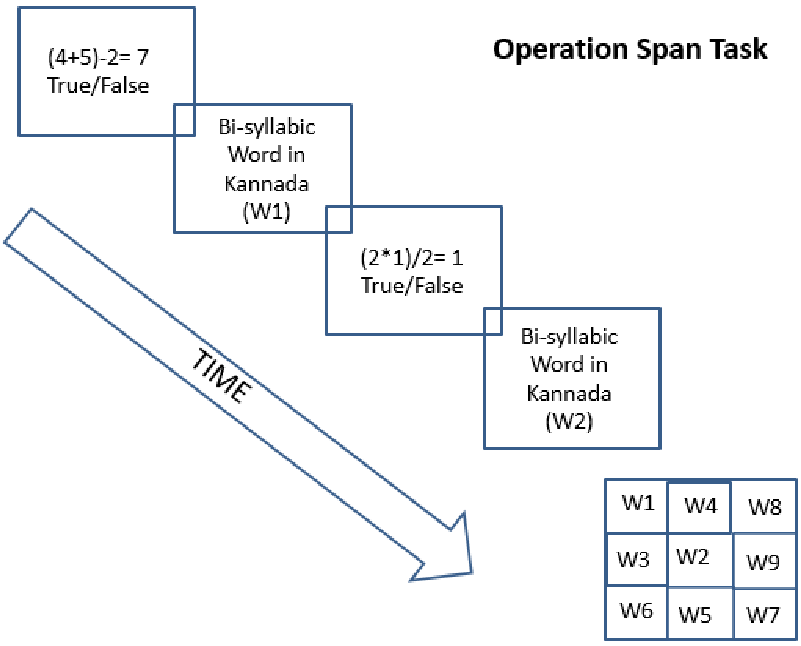


3.2.1.2 Operation Span Task. In this task, the participant's ability to remember the target stimuli was assessed. The stimulus was presented using the Smrithi Shravan 3.0 software, whichwas presented along with a secondary task. Here the secondary task was a distracting stimulus that involves solving an arithmetic problem, which was followed by a bi-syllabic Kannada target word which was recalled (e.g., is (7-4)*4=12 ---- true or false? ---- /mara/). The participant was instructed to solve the arithmetic problem and to judge whether the arithmetic problem is true or false and then remember the target word. Similarly, a series of arithmetic problems and target words difficulties were randomized such that the numbers of elements were unpredictable at the outset of an item as depicted in figure 3.2.

Guidelines recommended by (Conway et al., 1942) were followed during the scoring. A score of 1 was assigned for every word correctly recalled. At the end of each trial, the entire target words were shown along with the non-target words. Here the participant had to recall and choose all the target words in an order which were depicted in each trial in a correct sequence. The procedure and scoring were adapted from versions of the operation span task by (Kane et al., 2004).

Figure 3.2

A Sample Representation of Stimuli in Operation Span Task



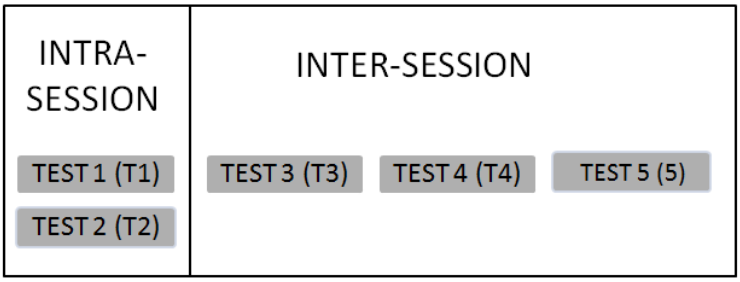
3.2.1.3 Auditory Digit Span Test. This is one of the tasks that assess working memory through auditory sequencing of numbers and auditory digit span. In this auditory digits were randomly presented, with an increasing level of difficultly. Each time the participant responds correctly, the length of the digits is increased by 1. If the response is incorrect, the length of the digit was shortened by one digit. An inter-stimulus interval of 1 second.

The auditory digit span is broadly categorized into a forward and backward span. Here, the digits group was presented in random order with an increasing number of difficulty, and participants are told to repeat the numbers in the same for forwarding span and backward order for backward span by typing the numbers. The stimuli consist of numerals from one to nine except seven and two. In forwarding span test (e.g., if the test stimulus is 'three, two, six, eight,' the response expected was 'three, two, six, eight') and in backward span test (e.g. 'four, nine, six, two,' the expected response was 'two, six, nine, four'). Working memory capacity is calculated as the total number of digits that the person can successfully recall in auditory number sequencing and digit span tests.

The test was conducted five times to check test-retest reliability, wherein test 1(T1) and test 2 (T2) was done within a day i.e., intra-session while test 3 (T3), test 4 (T4), test 5 (T5) was done across the days with a gap of at least two days among each testing sessions as in figure 3.3.

Figure 3.3

Sample Representing the Intra and Intersession Testing



3.3 Data Analysis and Statistical Analysis

Data were analyzed separately to assess intra-sessionand inter-session reliability. For statistical analysis, the following parameters were considered:

1. Reliability coefficients

Cronbach's alpha (reliability coefficient) was used to assess internal consistency between intra-session and intersession, and intra-class correlation coefficients (ICC) were also assessed to determine the reliability of some of the working memory span tasks. In both tests a value of 1 suggested absolute reliability.

2. Standard error of measurement (SEM)

The standard deviation of the measuring errors is known as SEM. The measurement was made using the following equation:

$$SEM = SD * (\sqrt{1 - \alpha})$$

Where SD is the standard deviation of the observed values α is the reliability coefficient. SEM was used to calculate work memory span tasks at 95 % confidence intervals. For each of the 1.96 SEM working memory scores the confidence interval has been calculated.

3. Smallest detectable difference (SDD)

The smallest detectable difference is the minimum difference in the session's mean scores. $SDD=1.96 * SEM * \sqrt{2}$

Chapter IV

Results

The present study aimed to assess the test-retest reliability of some of the working memory span tasks namely: Auditory digit span – Forward and Backward; Operation span task and Reading span tasks done within a day and across days.

4.1.1 Auditory Forward Digit Span Task

The intra-session reliability (Test 1 [T1] and Test 2 [T2], testing done within a day) and inter-session reliability (Test 3 [T3], Test 4 [T4] and Test 5 [T5], testing was done across days with a gap of at least two days among each testing sessions.

Figure 4.1 shows the mean along with one standard deviation for forward span scores within session (Test 1 and Test 2) and across sessions (Test 3, Test 4 and Test 5). Figure 4.2 and figure 4.3 indicate the forward span obtained by individual participants within and across different sessions. From the Figures 4.1-4.3 it can be inferred that forward span scores did not vary much within and across sessions.

Figure 4.1

Mean and Standard Deviation for Auditory Forward Digit Span Task within a Day (Test 1 and 2) and across Days (T3, T4, and T5). Error Bars Indicate One Standard Deviation.

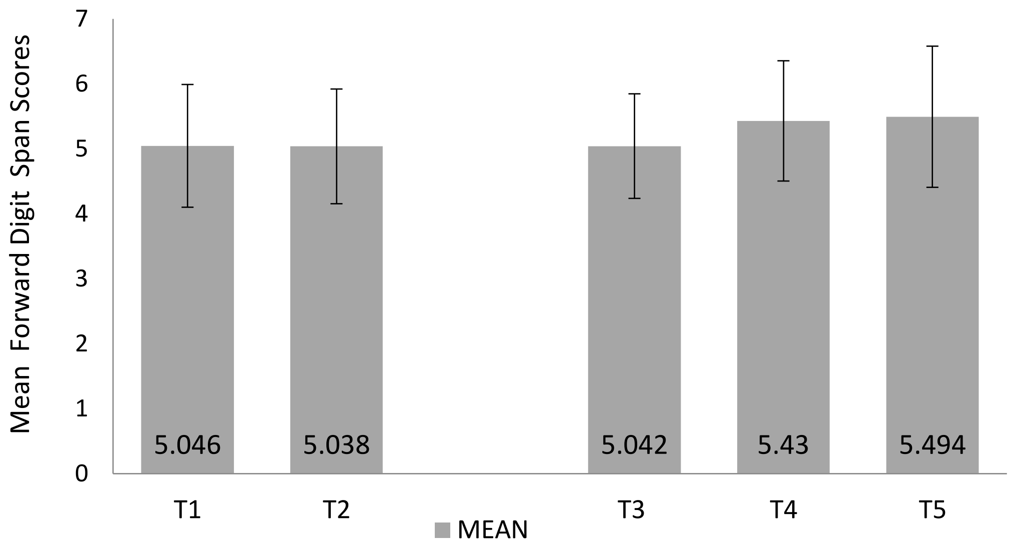


Table 4.1

Reliability Measures for Auditory Forward Digit Span Task for Intra and Inter-session.

	T1 (Test 1)	T2 (Test 2)	T3 (Test 3)	T4 (Test 4)	T5 (Test 5)
Cronbach' s Alpha	0.66			0.792	
Single Measure ICC	0.493			0.559	
SEM	0.55	0.514	0.367	0.422	0.495
SDD	1.535	1.434	1.024	1.178	1.381

Table 4.1 showed Cronbach's Alpha, ICC, SEMs, and SDDs of the maximum level scores obtained for intra-session (within a day) and inter-session (across days). Cronbach's α of 0.66for intra-session showed acceptable reliability. In contrast, for inter-session, the Cronbach's $\alpha$of 0.792 showed good reliability, ICC coefficients of 0.493for intra-session showed poor reliability and 0.559for inter-sessions moderate reliability. SEM ranged from 0.367 to 0.55, and SDD ranged from 1.024 to 1.535.

Figure 4.2

Auditory Forward Digit Span Scores within a Day (Test 1 and 2). Where X-axis Represents Participants, Y-axis Represents Auditory Forward Digit Span Scores

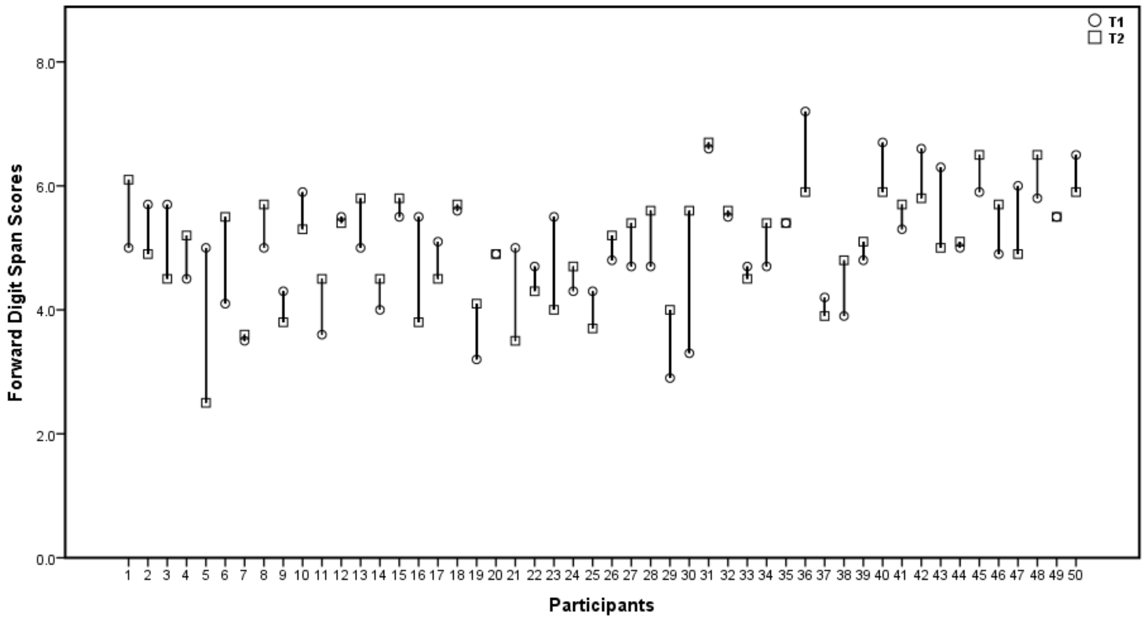
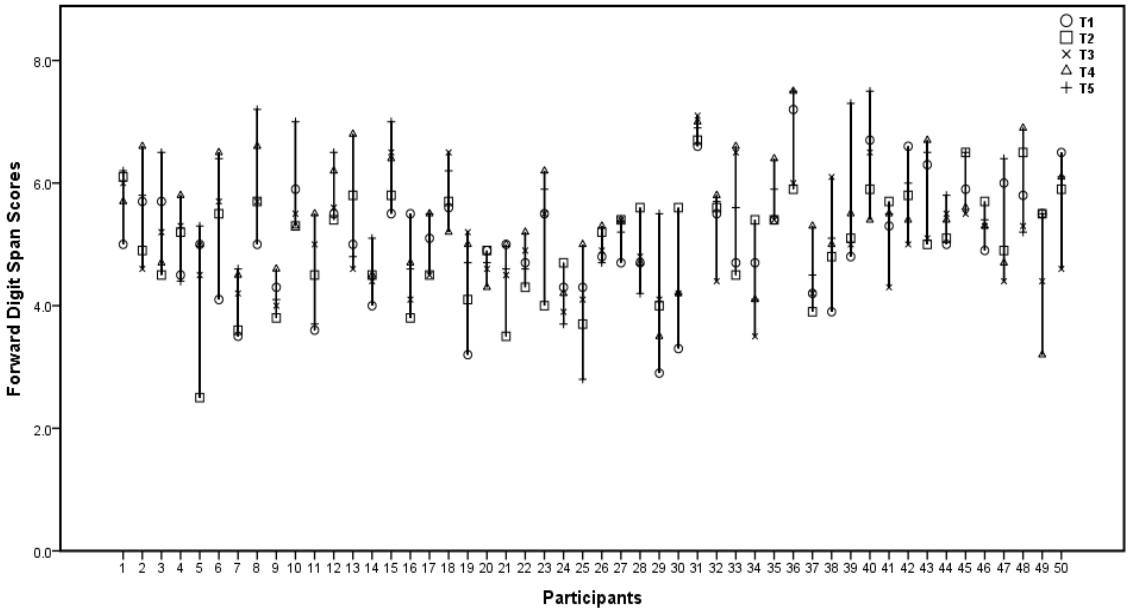


Figure 4.3

Auditory Forward Span Scores by Individual Participants within and across Sessions (Test 1,2,3,4 and 5)



4.1.2. Auditory Backward Digit Span Task

Figure 4.4 shows the mean along with one standard deviation for backward span scores within session (Test 1 and Test 2) and across sessions (Test 3, Test 4 and Test 5). Figure 4.5 and 4.6 indicate the backward span obtained by individual participants within and across different sessions. From the Figures 4.4- 4.6 it can be inferred that forward span scores did not vary much within and across sessions.

Figure 4.4

Mean and Standard Deviation for Auditory Backward Digit Span Tasks for Intra-session (Test 1 and 2) and Intersession (Test 3, 4, and 5). Error Bars Indicate One Standard Deviation.

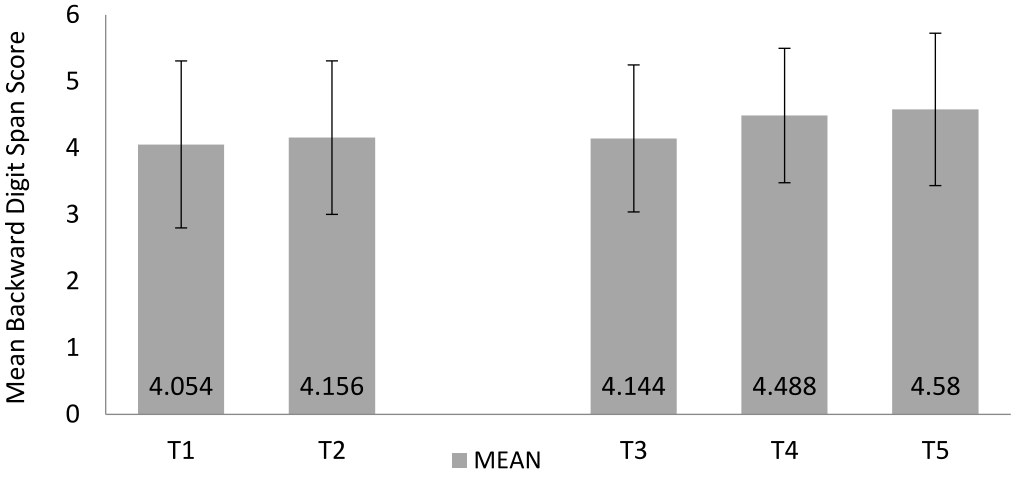


Table 4.2

Reliability Measures for Auditory Backward Digit Span Tasks between Intra and Inter-Session for Maximum Level Scores

	T1 (Test 1)	T2 (Test 2)	T3 (Test 3)	T4 (Test 4)	T5 (Test 5)
Cronbach's Alpha		0.822		0.8	
Single Measure ICC		0.697		0.571	
SEM	0.527	0.485	0.485	0.444	0.503
SDD	1.456	1.34	1.34	1.227	1.39

Table 4.2 showed Cronbach's Alpha, ICC, SEMs, and SDDs of the maximum level scores obtained for intra-session (within a day) and inter-session (across days). Cronbach's α of 0.822 for intra-session and 0.8for inter-session showed good reliability. ICC coefficients of 0.697for intra-session indicated moderate reliability, and 0.571for inter-sessions showed poor reliability. SEM ranged from 0.444 to 0.527, and SDD ranged from1.227 to 1.456.

Figure 4.5

Auditory Backward Digit Span Scores for Intra-session (Test 1 and 2). Where X-axis Represents Participants, Y-axis Represents Auditory Backward Digit Span Scores

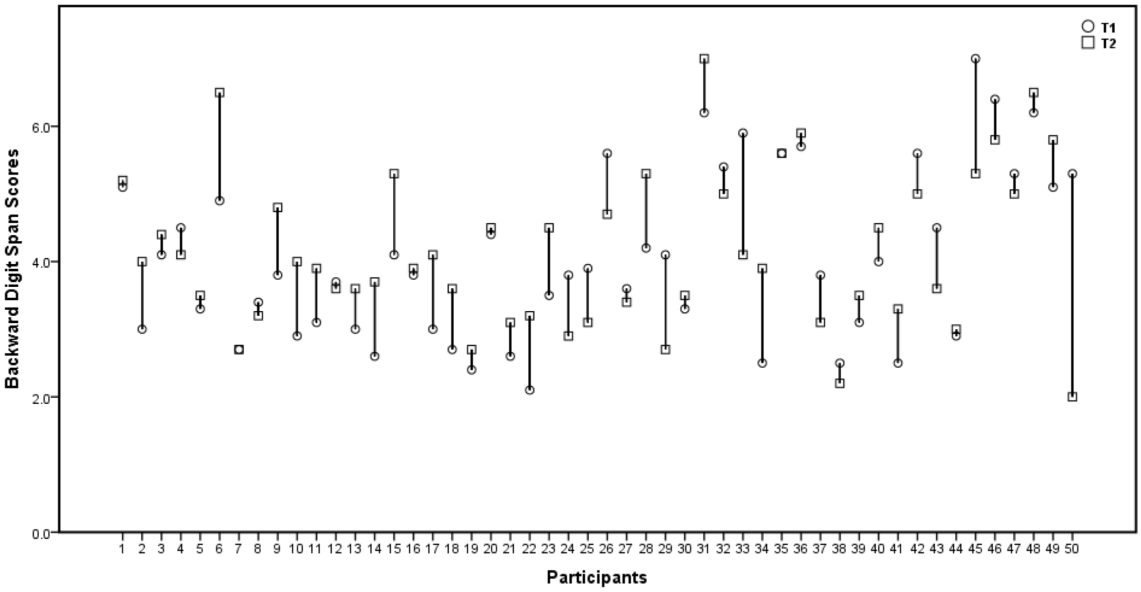
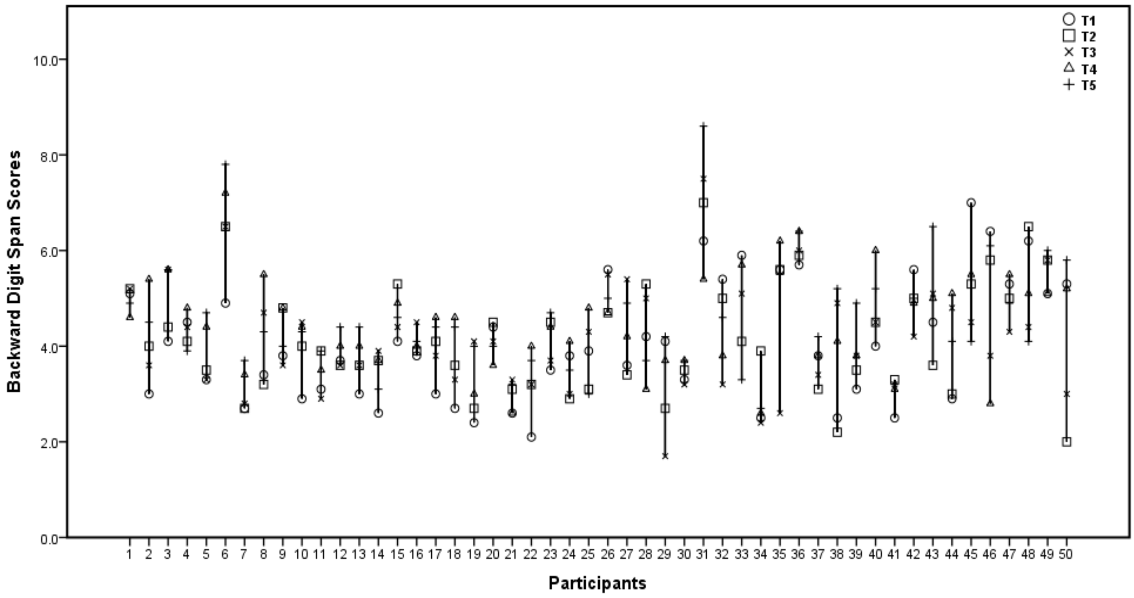


Figure 4.6

Auditory Backward Digit Span Scores by Individual Participants within and across Sessions (Test 1,2,3,4 and 5).



4.1.3 Operation Span Task

The results obtained for operation span would indicate the number of stimuli presented for each trial and their respective responses. Two scores were taken into consideration, i.e., item score and accuracy score. The item score was considered to be 1 is all the item has been repeated irrespective of the sequence. If one item is also missed, the score will be 0. Similarly, the accuracy score will be 1 if all the items are repeated in the preferred order, i.e., in the order in which the item has been presented. The score will be 0, even if one of the items has been missed out. Based on the scores Partial Credit Scores Weighted (PCSW) is calculated automatically Conway et al., (2005).

Figure 4.7 shows the mean along with one standard deviation obtained for operation span task within session (Test 1 and Test 2) and across sessions

(Test 3, Test 4 and Test 5). Figure 4.8 and figure 4.9 is the graphical representation of the participants for partial credit score/ weighted scores within and across the sessions. From figure 4.7 to 4.9 it can be inferred there was no much variation across the sessions.

Figure 4.7

Mean and Standard Deviation for PCSW for Operation Span Tasks for Intra and Intersessions

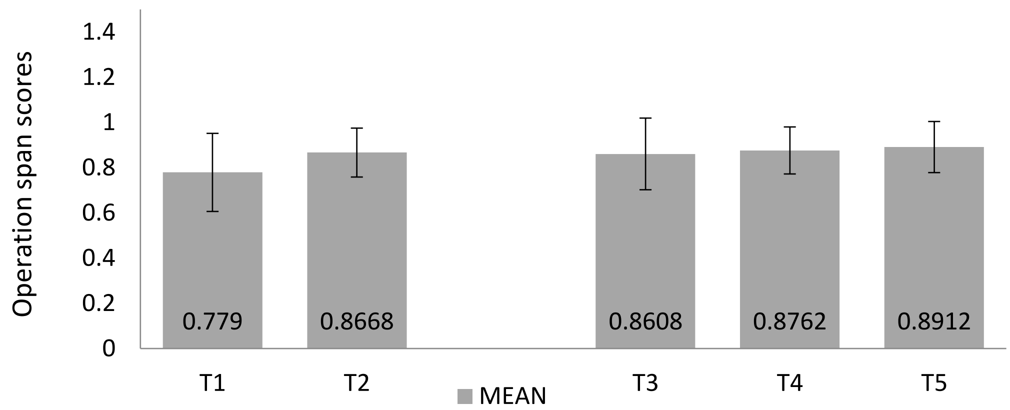


Table 4.3

Reliability Measures for Operation span Task for Intra and Intersessions

	T1 (Test 1)	T2 (Test 2)	T3 (Test 3)	T4 (Test 4)	T5 (Test 5)
Cronbach's Alpha	0.644			0.463	
Single Measure ICC	0.174			0.223	
SEM	0.102	0.064	0.107	0.07	0.076
SDD	0.281	0.176	0.295	0.193	0.21

Table 4.3 showed Cronbach's Alpha, ICC, SEMs, and SDDs of the maximum level scores obtained for intra (within a day T1 and T2) and inter-session (across days T3, T4 and T5). The Cronbach's α of 0.644 for intra-session showed acceptable reliability, and 0.463 for intersessions showed unacceptable reliability across the testing, ICC coefficients of0.174 for intra-sessions and 0.223 for inter-sessions showed poor reliability. SEM ranged from 0.064 to 0.102, and SDD ranged from0.176 to 0.295.

Figure 4.8

PCSW scores for Operation Span Task within a day (Test 1 and 2). Where X-Axis Represents Participants, Y- Axis Represents Operation Span Scores

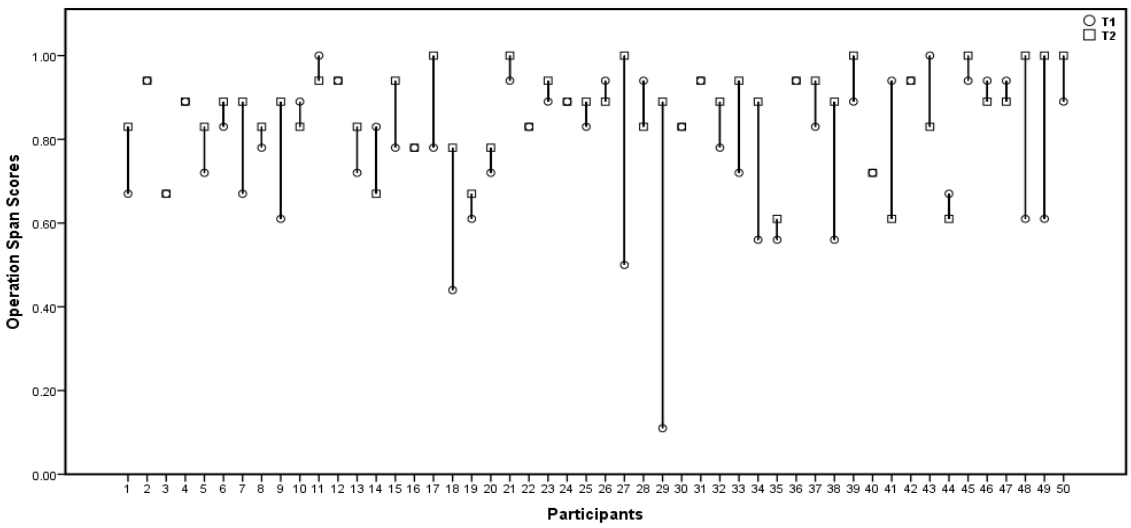
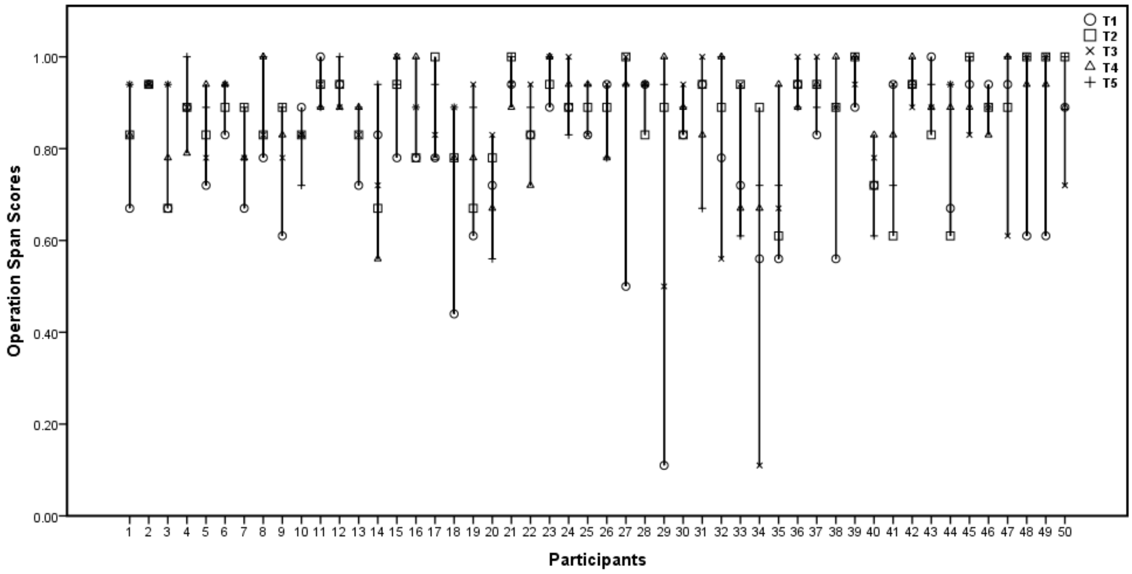


Figure 4.9

PCSW Scores across the Participants for Operation Span Task within and across the Sessions (Test 1,2,3,4 and 5)



4.1.4 Reading Span Task

The results obtained for reading span would indicate the number of stimuli presented for each trial and their respective responses. Two scores were taken into consideration, i.e., item score and accuracy score. The item score was considered to be 1 is all the item has been repeated irrespective of the sequence. If one item is also missed, the score will be 0. Similarly, the accuracy score will be 1 if all the items are repeated in the preferred order, i.e., in the order in which the item has been presented. The score will be 0, even if one of the items has been missed out. Based on the scores Partial Credit Scores Weighted (PCSW) is calculated automatically Conway et al., (2005).

Figure 4.10 shows the mean along with one standard deviation obtained for reading span task within session (Test 1 and Test 2) and across sessions (Test 3, Test 4 and Test 5). Figure 4.11and figure 4.12 is the graphical representation of the participants for partial credit score/ weighted scores within and across the sessions. From figure 4.10 to 4.12 it can be inferred there was no much variation across the sessions.

Figure 4.10

Mean and Standard Deviation for PCSW for Reading Span Tasks between Intra and Intersessions

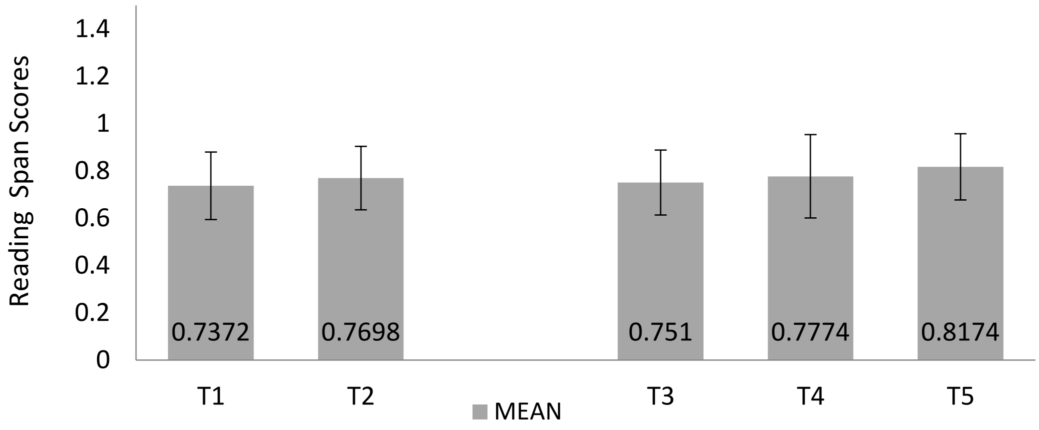


Table 4.4

Reliability Measures for Reading Tasks for Intra and Intersessions

	T1 (Test 1)	T2 (Test 2)	T3 (Test 3)	T4 (Test 4)	T5 (Test 5)
Cronbach's Alpha	0.644			0.604	
Single Measure ICC	0.475			0.337	
SEM	0.085	0.08	0.086	0.11	0.088
SDD	0.234	0.221	0.237	0.303	0.243

Table 4.4 showed Cronbach's Alpha, ICC, SEMs, and SDDs of the maximum level scores obtained for intra (within a day T1 and T2) and inter-session (across days T3, T4 and T5). The Cronbach's α of 0.644 for intra-session and 0.604for intersessions indicated acceptable reliability across the testing, ICC coefficients of0.475for intra- sessions and 0.337for inter-sessions indicated poor reliability. SEM ranged from 0.08 to 0.11, and SDD ranged from0.221 to 0.303.

Figure 4.11

PCSW Scores for Reading Span Task within a day (Test 1 and 2). Where X-axis Represents Participants, Y- axis Represents Reading Span Scores.

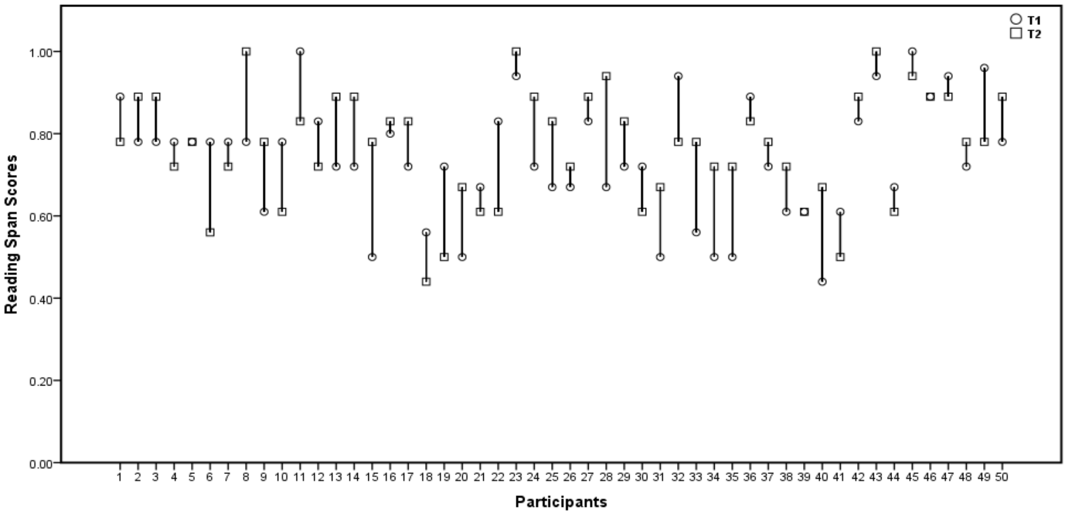
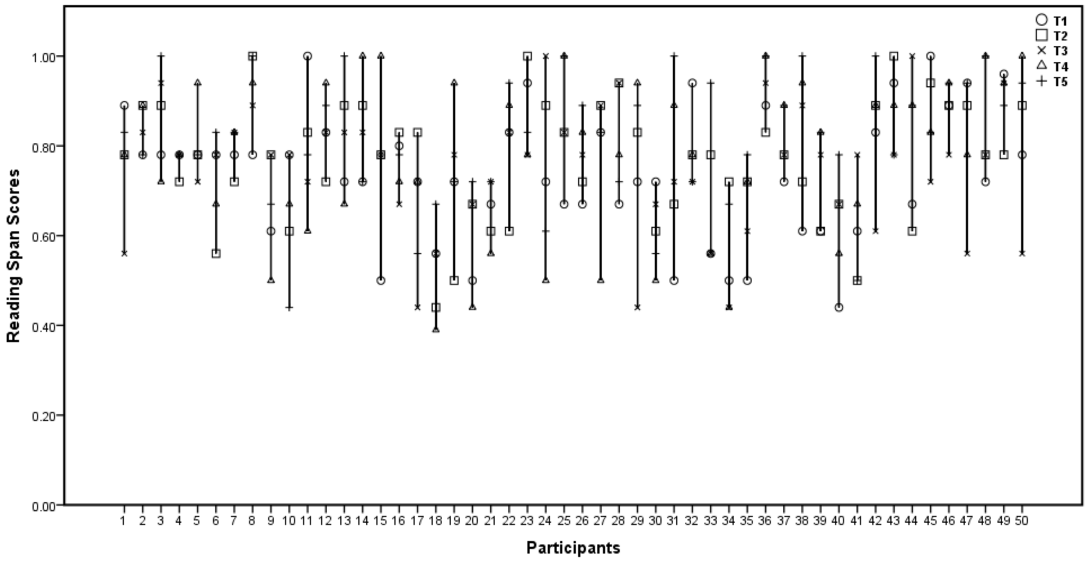


Figure 4.12

PCSW scores for reading span tasks across the participants for intra and intersessions (Test 1,2,3,4 and 5).



Chapter V

Discussion

The study's main aim was to find the test-retest reliability of the working memory span tasks. The present study investigated the working memory span tasks: Auditory digit span tasks – Forward and Backward; Operation span task and Reading span task. The test-retest reliability was assessed by conducting the study intra and inter-session, i.e., within a day and across days.

5.1 Working memory measures

The working memory span results were obtained by conducting the statistical analysis Cronbach's Alpha, Intraclass correlation, Standard error of measurement, and smallest detectable difference.

5.1.1 Auditory Digit Span Task

The auditory digit span task included both forward and backward span. The mean scores obtained for auditory forward digit span for intra-session and inter-sessions had no much variation within and across the session. Cronbach’s alpha also indicated an acceptable and good internal consistency within and across sessions, respectively. The intraclass correlation coefficient (ICC) indicated from poor to moderate reliability. Similar scores were also found in auditory backward digit span tasks. While the Cronbach's alpha for auditory backward digit span had good reliability, the ICC co-efficient had moderate reliability.

A study by Woods et al. (2011) assessed the improvement in forward and backward digit span and reported significant improvement with test-retest reliability of the digit span test. A study was conducted to assess the working memory measures across different age and reliability of the test. Seven different working memory span was measured, the backward digit span was one among them. All of the tests conducted, including backward digit span, had adequate internal consistency on the performance of task twice (Waters & Caplan, 2003).

A study was conducted to evaluate the paced serial addition test (PASAT) to assess the cognitive abilities with traumatic brain injury. The working memory digit span test was used as one of the most important tests to evaluate cognitive function. The study also founded that an individual whose working memory was impaired showed a remarkable sign of aphasia. The digit span test also had strong test reliability and established the same test for clinical and research purposes (Nikravesh et al., 2017).

5.1.2 Operation and Reading Span Task

The partial credit score/ weighted was the score obtained for operation and reading span task. The mean scores did not vary much within and across the sessions. The Cronbach's alpha for intra-session was acceptable and unacceptable reliability for inter-session. ICC co-efficient measured had poor reliability for both sessions, respectively. Whereas, for the reading span task had acceptable reliability for both intra and inter-session. ICC co-efficient measured for both the sessions had poor reliability.

Unsworth et al. (2005) conducted a study on the operation span task using the automatic version and inferred that it showed good reliability. Turner and Engle (1989) also conducted (Turner & Engle, 1989)a similar task worked on operation span and finally arrived at the results that the task was moderately reliable. The explanation given about the reliability by the authors were the difference in the task to be recalled, and the mode of responding to the stimuli, i.e., (Unsworth et al., 2005) used letter and the participants had to the respond by choosing the items from the given pool items scoring happened automatically. While, (Turner & Engle, 1989) used words the participants had to respond to the task by writing down the items that they have recalled, later the experimenter had to evaluate it. So, it can be seen that for the individual who had to write the response, and the individual who used the automated version, the response rate was faster. So the test reliability was better for the automated version of the operation span task. However, according to the present study, through an automatic version used for testing, the reliability scores obtained were poor.

A study by Klein et al.(1999) worked on the operation span task on individuals three times with a gap of 3 weeks for the second session and gap of 6- 7 weeks for the third session. It was inferred that the internal consistency and reliability were high. The possible explanation would be that the individual would have got much familiar with the testing, and the scores were higher for the most prolonged interval. The same was not found in our study. For further clarification, in the present study, the testing has to be done on a larger population.

A study by Compostela (2008) worked on the automated external validation of the operation span task done for individuals and in groups (test being administered for many individuals at once). It can be inferred from the study that on comparing both groups, internal consistency for operation span task conducted for individuals was high. In the present study, though the testing was conducted individually, the internal consistency obtained was poor.

Similarly, a study by Redick et al. (2012) which measured the working memory capacity with automated Complex Span Tasks, included both operation span task and reading span task. Both the working memory span task had good internal consistency.

Chapter VI

Summary and Conclusion

The present study aimed to assess the test-retest reliability of some of the working memory span tasks, namely: Auditory digit span task – forward and backward, operation span task and reading span task, which was conducted five times to check test-retest reliability, wherein test 1(T1) and test 2 (T2) was done within a day, i.e., intra-session while test 3 (T3), test 4 (T4), test 5 (T5) was done across the days with a gap of at least two days among each testing sessions. Fifty adults in the age group of 18-35 (mean age = 21.22 years, SD = 3.151 years) years were taken to fulfil the study. All the participants were native speakers of Kannada, were also able to read and write Kannada. The data obtained were tabulated and analyzed with statistical analysis using software packages for statistical analysis (SPSS, Version 21.0). The statistical analysis, namely Cronbach's Alpha, intraclass correlation coefficient, standard error of measurement, and smallest detectable difference, were used.

On finding obtained from the results it indicated the auditory digit span test had acceptable to good reliability for forward span and good reliability for the backward span. Simultaneously, the reliable measures obtained for operation span ranged from acceptable to unacceptable for intra and inter-sessions, respectively, while reading span had acceptable reliability. On keeping in thought, the auditory digit span task can be utilized to test the clinical population. In contrast, further research needs to be done for operation and reading span tasks to implement on the clinical purpose.

References

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of Working Memory: Differences in Definition, Degree of Modularity, Role of Attention, and Purpose. *Language, Speech, and Hearing Services in Schools*, *49*(3), 340–355. https://doi.org/10.1044/2018_lshss-17-0114

Atkinson, R., & Shiffrin, R. (1968). Human Memory: A Proposed System and its Control Processes. En K. Spence & J. Spence (Eds.). *The Psychology of Learning and Motivation*, pp.89-195.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, *225*(2), 82-91.

Baddeley, A. (2000). *The episodic buffer : a new component of working memory ? 4*(11), 417–423.

Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *The Quarterly journal of experimental psychology*, *18*(4), 302-309. https://doi.org/10.1080/14640746608400047

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly journal of experimental psychology*, *18*(4), 362-365. https://doi.org/10.1080%2F14640746608400055

Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *The Quarterly Journal of Experimental Psychology*, *20*(3), 249-264, https://doi.org/10.1080/14640746808400159

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long-and short-term memory. *Journal of verbal learning and verbal behavior*, *9*(2), 176-189. https://doi.org/10.1016/S0022-5371(70)80048-2

Baddeley, A. D., & Hitch, G. (1974). Working Memory. *Psychology of Learning and Motivation*, *8*, 47–89. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, Alan D. (1992). Working memory. *Science*, *255*(5044), 556–559.

Baddeley, Alan D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D. (2017). The central executive: A concept and some misconceptions. *Exploring Working Memory: Selected Works of Alan Baddeley*, 247–252. https://doi.org/10.4324/9781315111261

Compostela, S. De. (2008). *External validation of the computerized , group administrable adaptation of the " operation span task "*. *40*(1), 46–54.

Cowan, N. (1984). On Short and Long Audiology Stores. In *Psychonomic Bulletin* (Vol. 96, Issue 2, pp. 341–370).

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786. https://doi.org/10.3758/BF03196772

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Memory and Language*, *19*(4), 450.

Henry, L. A., Messer, D. J., & Nash, G. (2012). Phonological and visuospatial short-term memory in children with specific language impairment. *Journal of Cognitive Education and Psychology*, *11*(1), 45-56. https://doi.org/10.1891/1945-8959.11.1.4

Kane, M. J., Tuholski, S. W., Hambrick, D. Z., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Klein, K., Fss, W. H., Carolina, N., & Carolina, N. (1999). *The reliability and stability of the Turner and Engle working memory task*. *31*(3), 429–432.

Kumar, A. U., & Sandeep, M. (2013). Auditory cognitive training module. ARF funded departmental project submitted to All India Institute of Speech and Hearing, Mysore.

Margolis, R. H., & Heller, J. W. (1987). Screening tympanometry: Criteria for medical referral: Original papers. *International Journal of Audiology*, *26*(4), 197–208. https://doi.org/10.1080/00206098709081549

Marzi, C. A., di Stefano, M., Tassinari, G., & Crea, F. (1979). Iconic storage in the two hemispheres. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(1), 31–41. https://doi.org/10.1037/0096-1523.5.1.31

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81. https://psycnet.apa.org/doi/10.1037/h0043158

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). Plans and the structure of behaviour. https://psycnet.apa.org/doi/10.1037/10039-000

Milner, B. (1966). Amnesia following operation on the temporal lobes. *Amnesia*.

Nikravesh, M., Jafari, Z., Mehrpour, M., Kazemi, R., & Shavaki, Y. A. (2017). *The paced auditory serial addition test for working memory assessment : Psychometric properties*. *2017*.

Olichney, J. M., & Hillert, D. G. (2004). Clinical applications of cognitive event-related potentials in Alzheimer's disease. *Physical Medicine and Rehabilitation Clinics of North America*, *15*(1), 205–233. https://doi.org/10.1016/S1047-9651(03)00103-7

Olton, D. S. (1979). Mazes, maps, and memory. *American Psychologist*, *34*(7), 583–596. https://doi.org/10.1037/0003-066X.34.7.583

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*(3), 193–198. https://doi.org/10.1037/h0049234

Pisoni, D. B. (1975). Auditory short-term memory and vowel perception. *Memory &*

Cognition, *3*(1), 7-18. https://doi.org/10.3758/BF03198202

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). *Measuring Working Memory Capacity With Automated Complex Span Tasks*. https://doi.org/10.1027/1015-5759/a000123

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: a neuropsychological study. *The Quarterly Journal of Experimental Psychology*, *22*(2), 261–273. https://doi.org/10.1080/00335557043000203

Shih, R., Dubrowski, A., & Carnahan, H. (2009). Evidence for haptic memory. *Proceedings - 3rd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, World Haptics 2009*, *January*, 145–149. https://doi.org/10.1109/WHC.2009.4810867

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5

Unsworth, N., Heitz, R. P., & Engle, R. W. (2005). *An automated version of the operation span task*. *37*(3), 498–505.

Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, and Computers*, *35*(4), 550–564. https://doi.org/10.3758/BF03195534

Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Poliva, O., Hink, R. F., & Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal of Clinical and Experimental Neuropsychology*, *33*(1), 101–111. https://doi.org/10.1080/13803395.2010.493149