QLANTIC JOURNAL OF SOCIAL SCIENCES
OPEN ACCESS
Pages: 22 – 31
ISSN (Online): 2791-0202
DOI: 10.55737/qjss.vi-i.25272
Vol. 6 | No. 1 | Winter 2025
Research Article
Qlantic Journal of Social Sciences (QJSS)

Experimental Validation and Reliability Study of the N-Back Working Memory Task

Ather Mujitaba [1] Hamza Majeed [2] Ayesha Asif [3] Kanza Shamas [4] Sawera Mehmood [5]

Check for updates

Abstract: *The n-back is a well-known working memory task that has not been experimentally validated among the Pakistani population. Therefore, in light of this gap, this study aimed to experimentally validate and assess the reliability of working memory tasks (alphabet and shapes tasks) in young adults. We employed repeated-measures true experimental design for evaluating the study's main objective, which included 40 participants between the age ranges of 18-25 years, recruited through a purposive sampling technique. The selected participants were randomly allocated into two groups using the fishbowl method to avoid selection bias. Both groups performed alphabet and shape n-back tasks designed on PsychoPy software package 24.4. The study's finding revealed no meaningful difference between both groups' performance in the alphabet and shapes tasks, as both groups' correct and incorrect scores did not vary statistically. In contrast, within-group performance suggested that both groups' performance was distinctively better on the alphabet task than on the shapes task. Thus, we have concluded that the newly developed n-back task is a valid and reliable task for the assessment of working memory.*

Key Words: N-Back, Alphabet and Shapes Task, Working Memory, Young Adults

Introduction

Working memory (WM) is an intricate system that allows us to store and process incoming information despite distractions and interruptions (Diamond, 2013). This cognitive function is necessary for complex reasoning but has limited capability (Ke et al., 2019). WM is critical for adjusting to a constantly changing environment and indicates an individual's ability to store and use knowledge in short-term memory (Baddeley & Hitch, 1974). It facilitates storing and updating relevant information, which is critical for goal-directed actions. Earlier theories, particularly those based on Baddeley's (1986) framework, described WM as having at least two subordinate systems: the phonological loop, which manages auditory input and verbal processing and the visuospatial sketchpad, which deals with visual and spatial data. The central executive controls attention control, decision-making, and problem-resolution tasks. Furthermore, this concept implies that working memory is more than a passive storage system; it actively analyzes and manipulates data to facilitate learning and problem-solving. The behavioural and neurophysiological aspects of working memory (WM), as well as its theoretical constructs, have been the subject of numerous studies that have used both single-session methodologies (Scharinger et al., 2017) and repeated measure approaches (Jaeggi et al., 2014).

The n-back task (Owen et al., 2005) is the most widely used way to measure working memory (WM). It is a challenging task that requires individuals to manage inhibition and interference (Oberauer, 2005; Kane et al., 2007) as well as store, maintain, and manipulate information (Chen et al., 2008; Jaeggi et al., 2008). The n-back task has been used in a variety of contexts, including single-session behavioural studies

[1] Lecturer, Department of Psychology, GIFT University, Gujranwala, Punjab, Pakistan. ✉ ather.mujitaba@gift.edu.pk
[2] Department of Psychology, GIFT University, Gujranwala, Punjab, Pakistan. ✉ 191520163@gift.edu.pk
[3] Department of Psychology, GIFT University, Gujranwala, Punjab, Pakistan. ✉ 191520183@gift.edu.pk
[4] Department of Psychology, GIFT University, Gujranwala, Punjab, Pakistan. ✉ 201520138@gift.edu.pk
[5] Department of Psychology, GIFT University, Gujranwala, Punjab, Pakistan. ✉ 191520164@gift.edu.pk

- **Corresponding Author:** Ather Mujitaba (✉ather.mujitaba@gift.edu.pk)
- **To Cite:** Mujitaba, A., Majeed, H., Asif, A., Shamas, K., & Mahmood, S. (2025). Experimental Validation and Reliability Study of the N-Back Working Memory Task. *Qlantic Journal of Social Sciences*, *6*(1), 22-31. https://doi.org/10.55737/qjss.vi-i.25272

(Brouwer et al., 2012) and neurophysiological research (Scharinger et al., 2017). It has also been employed in multi-session behavioural training experiments (Blacker et al., 2017) and neurophysiological training studies (Pergher et al., 2018).

The n-back task was initially introduced by Kirchner (1958) as a visuospatial test with four levels of difficulty, ranging from 0-back to 3-back. A visual letter task with up to six levels of difficulty was later developed by Mackworth (1959). This task, which was first introduced to the neuroscience community by Gevins et al. (1990), is a visuomotor memory task with a single level of difficulty (3-back). It involves several cognitive functions, such as encoding incoming stimuli, updating and maintaining information, and comparing the current stimulus with one that occurred n times earlier in the sequence.

The majority of research investigations require participants to use two or more buttons to respond to specific stimuli throughout each trial to determine whether the stimulus is a target or a non-target (Miller et al., 2009). Traditionally, the stimuli in n-back tasks have included numbers or words; however, in recent years, pictorial variants featuring emotional scenes, faces, or food items have been introduced. As dependent variables, response latencies (reaction durations) and accuracy or number of errors are typically reported in studies. As the task's difficulty increases (with higher n values such as 2, 3, 4, 5, 6,...), typically indicated by a greater number of stimuli, there is a corresponding rise in reaction times and a drop in accuracy (Schmidt et al., 2009). Additionally, a negative correlation is often found between reaction times and accuracy (Carter et al., 1998), suggesting that higher mistake rates are linked to longer reaction times.

Thus, the n-back task serves as a multifaceted metric that engages various cognitive processes, which appear to be predominantly independent of the specific stimuli or materials employed. Regardless of the materials employed, the frequency of errors and reaction times continuously increases as n increases. These findings are documented at the neural level by neuroimaging research. This load-dependent activation is most frequently seen in the bilateral prefrontal and parietal cortices. A network that is typically active during working memory activities depends on these regions (Wager & Smith, 2003). No matter what kind of materials are used, these important areas of activation have been found (Owen et al., 2005).

The effectiveness of the n-back task as a working memory test has been questioned, despite its widespread use. Compared to other working memory tests, research has yielded inconsistent results, indicating that n-back performance may not fully represent all facets of working memory function (Loschky et al., 2014). However, its ability to be modified for cognitive load and its significant association with other cognitive abilities, such as fluid intelligence, are the reasons that it is still in use today (Parker et al., 2020). The n-back task is presently being improved, and its potential uses in education, cognitive training, and rehabilitation are being investigated (Fahey et al., 2018).

However, despite its widespread use in research and education, the n-back working memory task is not tested in the Pakistani normative sample, underscoring a potential knowledge gap. Moreover, the importance of cognitive evaluations that are culturally relevant and accurately reflect the cognitive capacities of varied individuals is becoming more widely recognized. The n-back task is versatile and may be adjusted to meet the linguistic and cultural circumstances of Pakistani participants, which increases its validity and reliability (Almodóvar-Payá et al., 2022). Additionally, as the importance of cognitive health is increasingly recognized, developing n-back working memory tasks in Pakistan can aid in the early detection of cognitive impairments, particularly in populations at risk for neurodegenerative diseases (Tang et al., 2021). Therapists can modify their strategies to better meet the cognitive needs of their clients by assessing their working memory skills (Agbangla et al., 2022).

Method

This study mainly aims to validate and investigate the reliability of the n-back working memory task in a sample of young Pakistani adults.

Objectives

1. To develop and validate the n-back working memory task of alphabet and shapes.
2. To assess the working memory performance of participants on the n-back alphabet and shapes task.

Hypothesis

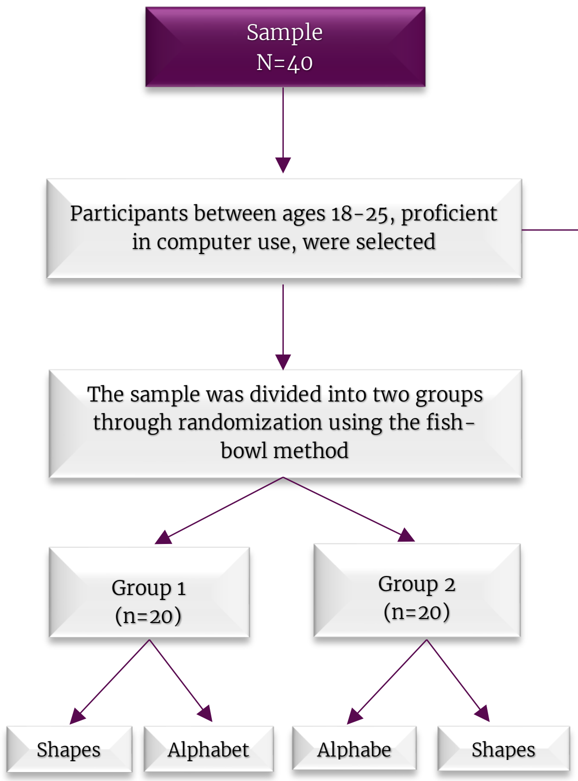
1. The performance between groups on the alphabet and shapes task will not differ significantly.
2. Participants of both groups will perform better on the alphabet task compared to the shapes task.

Study Design

The research design used in this experiment was a repeated measure block design (ABBA), as the sample's groups were given Alphabet and shape tasks in blocks. The experiment contains two groups, one of which is Group 1 and the other is Group 2. The participants were randomly divided into these groups using the fish-bowl method, as shown in Figure 1. The participants were unaware that their task assignment was to maintain the single-blind condition.

Figure 1

Sample Selection and Experimental Task Assignment



Note. The recruited participants were randomly divided into two groups through the fishbowl method. Each group performed both the Alphabet and Shapes N-back tasks following the ABBA counterbalancing assignment.

Sample

The sample of this experiment consisted of 40 university students between the ages of 18 and 25, including 20 male and 20 female participants. A purposive sampling technique was used to recruit participants for the experiment as only participants proficient in computer use were included in the sample, while participants who exceeded the suggested age range, lacked proficiency in computer use, or were penitential outliers were excluded from the data. Participant selection and task assignment are procedures explained in Figure 1.

Experimental Protocol

The working memory experiment task includes alphabet and shapes tasks, previously used mostly in neuroimaging studies. However, for the current study, we have adopted and designed alphabet and shape tasks according to the previously published work by Ke et al. (2019). The main reason for adopting the traditional alphabet along with the shapes task was to develop a working memory task that would measure the verbal and non-verbal working memory ability of young participants.

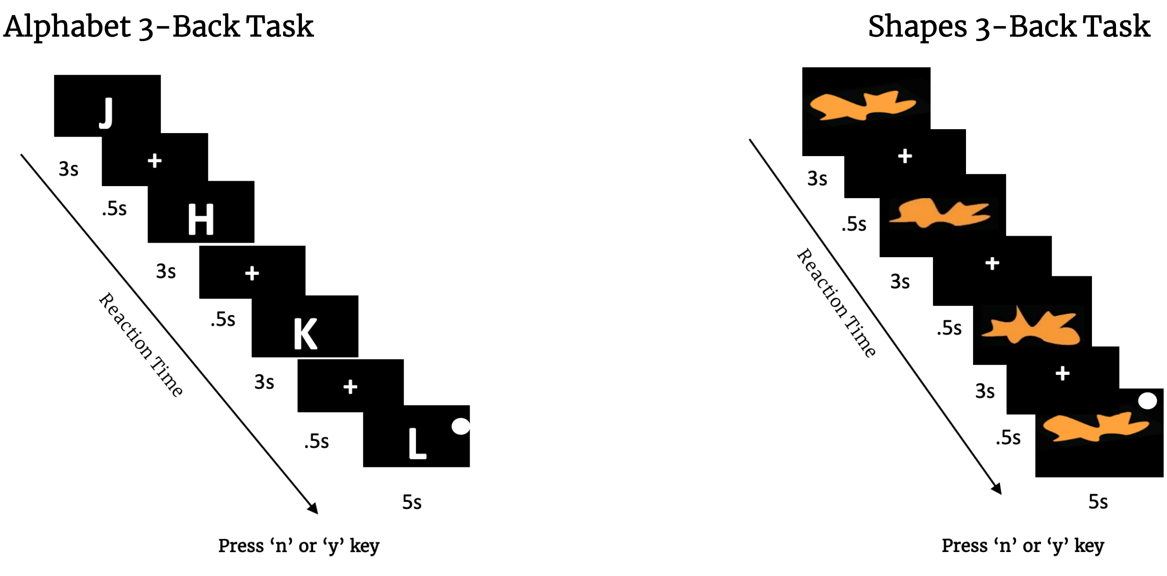
The working memory Alphabet and shapes tasks were designed on the PsychoPy software package (24.3). The Alphabets Stimulus contains 10 consonants which are white letters in Arial font with 12 font size on a black background. The screen size was 720*1280 pixels (720 height and 1280 width). The alphabet task included 10 trials; each trial began with the appearance of a fixation cross in the middle of the screen for .5s following the random sequence of alphabet, each alphabet visible on screen for 3 seconds, random order of alphabet ended up on alphabet with white dot for 5s (cue for participant response), upon cue participant had to guess whether cue alphabet had been appear exact three back position by pressing "Y" key for yes response "N" for no response as shown in Figure 3. In each trial, a different random order of the alphabet was presented. Thus, ten trials were completed for the alphabet task, and after a brief gap, the participant performed Shapes trials.

The Shapes Stimulus contains 10 orange-coloured shapes on a black background, with the same screen size as alphabets 720*1280 pixels (720 height and 1280 width). The size of each shape was 590*720 pixels (590 height and 720 width). The trial presentation setup was similar to the alphabet task. Each trial began with the appearance of a fixation cross (for 0.5s) on the screen, followed by a random sequence of shapes, with each shape displayed for 3s. The target shapes were then presented for 5s with a circle at the top right that represents the target figure for response, as shown in Figure 3. Participants had to recognize whether the current shape was the same as the one that was presented three times before in the sequence. The response options were the same as those used for the alphabet trials.

The primary outcome of the Alphabet and Shapes task was determined by the participant's correct and incorrect responses in the Alphabet and Shapes trials. A higher correct score and a lower incorrect score correspond to better performance of working memory tasks.

Figure 2

Schematic Diagram of Alphabet and Shapes Task Trials



Note. The setup used for the alphabet and shape trials in each session is displayed in the figure above. The alphabet trials are shown on the left, and the shape trials on the right. On the computer screen, a fixation cross appears, followed by a random order of the alphabet. Upon receiving the cue (an alphabet with a dot), the participant is required to guess whether the specific cued alphabet appeared three positions back by pressing 'Y' for a 'yes' response and 'N' for a 'no' response. A similar setup was used for the shape trials, employing various orange-coloured shapes.

Procedure

After receiving ethical approval from the University Ethical Review Committee, this experimental study was carried out. The ethical standards were adequately considered while experimenting. Before taking part in the experiment, the participants were asked to complete the consent form. The consent form contained all the information about the experiment's aim and the ethical handling of participants' data. After the initial screening of participants, participants were given random trials to evaluate their proficiency in computer use.

Participants who were proficient in computer-based tasks were included in the experiment, while those with very slow response rates or difficulties in performing random trials were excluded. The selected participants were randomly allocated into two groups. Group 1 first performed the Alphabet task, then the Shapes task and Group 2 performed the Shapes task first and then the Alphabet task, following an ABBA counterbalancing repeated measure experimental design. We employed a single-blind condition for participants, in which they were given uniform instructions and an interruption-free environment in the experimental Psychology lab of GIFT University. Hence, one by one, data from 40 participants were collected and compiled for initial analysis on SPSS 27. Initially, the participant data were evaluated for potential outliers and normality tests. After obtaining a satisfactory level of data quality, the Alphabet and Shapes Task scores were compared across groups using an independent sample t-test, while the differences in performance within groups were examined using a paired sample t-test.

Results

Table 1

Descriptive Statistics of Demographic Variables (N=40)

	Group 1 (*n*=20)				Group 2 (*n*=20)			
Variables	*n*	%	*M*	*SD*	*n*	%	*M*	*SD*
Age			21.75	1.61			21.20	1.24
Gender								
Male	11	55	22.09	1.44	9	45	21.44	1.33
Female	9	45	21.33	1.80	11	55	21.00	1.18

Descriptive characteristics, such as age and gender, of the two groups are presented in Table 1. The average age of participants in Group 1 was (M=21.75, SD = 1.61) years, while Group 2 had a slightly lower average age of (M= 21.20, SD = 1.24). The Gender distribution shows that in group 1, 55% were male (n=11) with a mean age of (M=22.09, SD = 1.44), and 45% were female (n=9) with a mean age of (M=21.33, SD = 1.80). In group 2, males comprised 45% (n=9) with a mean age of (M=21.44, SD = 1.33), while females comprised 55% (n=11) with a mean age of (M=21.00, SD = 1.18). The test of normality of data revealed that the shapes stimulus had a normal distribution, but the alphabet stimulus had a slightly deviating distribution. However, the skewness and kurtosis were in acceptable ranges. So, we can say that the data was near normal distribution.

Table 2

Independent Sample t-test for Difference Between Alphabet and Shapes Tasks (N=40)

	Group 1 (*n*= 20)		Group 2 (*n*=20)				
Variable	*M*	*SD*	*M*	*SD*	*t*	*p*	*Cohen's d*
Alphabets							
Correct Responses	9.45	.99	8.60	1.98	1.71	.09	.54
Incorrect Responses	.55	.99	1.40	1.98	-1.71	.09	-.54
Shapes							
Correct Responses	7.10	1.55	7.00	1.86	.18	.85	.05
Incorrect Responses	2.90	1.55	3.00	1.86	-.18	.85	-.05

In Table 2, the between-group performance difference of both group in alphabet and shapes tasks are displayed. The performance of both groups was evaluate based on their correct and incorrect responses. The results revealed that both groups did not differ in their performance on both alphabets and shapes tasks. For instance, the mean score of correct responses on the alphabet task for Group 1 was (M = 9.45, SD = .99) and for Group 2, it was (M = 8.60, SD = 1.98) which was not statistically significant (p = .09). Similarly, for incorrect responses, mean for Group 1 was (M = .55, SD = .99) and for Group 2, it was (M = 1.40, SD = 1.98) which was also not statistically significant (p = .09), although the effect size remained large (Cohen's d = .54).

In the same vein, both group's performance indicated a lack of significant difference in their performance on the Shapes task. The mean score of correct responses for the shapes task for Group 1 was (M = 7.10, SD = 1.55) and for Group 2, it was (M = 7.00, SD = 1.86), which yielded a non-significant p-value (p =.85). This trend was also observed in the comparison of incorrect responses, where the mean score for Group 1 was (M = 2.90, SD = 1.55) and for Group 2, it was (M = 3.00, SD = 1.86), corresponding to non-significant p-value (p=.85) with minimal effect size (Cohen's d = .05). These findings support our expected outcomes, as we have given both groups similar task and, practically their performance should not differ.

Table 3

Paired Sample t-test for Difference Between Alphabet and Shapes Tasks (N=40)

Variable	Alphabets		Shapes				
	M	*SD*	*M*	*SD*	*t*	*p*	*Cohen's d*
Group 1 (n= 20)							
Correct Responses	9.45	.99	7.10	1.55	5.98	.001	1.33
Incorrect Responses	.55	.99	2.90	1.55	-5.98	.001	-1.33
Group 2 (n= 20)							
Correct Responses	8.60	1.98	7.00	1.86	3.23	.004	.72
Incorrect Responses	1.40	1.98	3.00	1.86	-3.23	.004	-.72

Table 3 presents the results of the paired sample t-test, which provides a within-group comparison of the alphabet and shapes task. For Group 1, the average correct responses were significantly higher (M = 9.45, SD=.99, ***p<.001) in the alphabet task as compared to the shapes task (M=7.10, SD=1.55), with a large effect size (Cohen's d value= 1.33). Group 1 made significantly fewer mean incorrect responses (M = .55, SD = .99, ***p<.001) in the alphabet task compared to the shapes task (M= 2.99, SD= 1.55). This indicated Group 1 had better performance in the alphabet task with a large effect size (Cohen's d value = -1.33).
A parallel comparison was conducted between the alphabet and shapes tasks' performance in Group 2. This also revealed the superior performance of Group 2 in the alphabet task. For example Group 2's mean correct responses on alphabet task (M= 8.60, SD = 1.98) were significantly higher than those on the shapes task (M = 7.00, SD= 1.86), with statistically significant difference (**p=.004), corresponding to large Cohen's d= .72. In contrast, Group 2 mean incorrect responses on the alphabet task (M = 1.40, SD= 1.98) were significantly lower (**<.001) compared to shapes task (M = 3.00, SD= 1.86), with a large effect size (Cohen's d value = -.72). This contrast suggests that Group 2 also performed better on alphabet task. Hence, we concluded that both groups demonstrated better performance in the alphabet versus shape n-back task.

Discussion

Working memory (WM) is the cognitive mechanism that temporarily stores and manipulates information. It is crucial for several cognitive processes, including comprehension, thinking, and problem-solving. Baddeley's working memory model (Baddeley & Hitch, 1974) proposes a multi-faceted system which comprises the central executive, the phonological loop, and the visuospatial sketchpad (Baddeley, 2000). This framework suggests that distinct subsystems handle verbal and visual information. Such differentiation may impact the efficiency of storing and processing information. The purpose of this study was to develop and validate n-back working memory tasks and to examine whether working memory performs better with alphabets compared to shapes. For that purpose, a repeated measures ABBA experimental design was employed to validate and assess the reliability of the tasks. Purposive sampling was used to choose 40 participants, who were then randomly divided into two groups. Both groups performed the alphabet and shape tasks in blocks. The comparison between groups indicated that there were no significant differences in correct and incorrect responses for the alphabet and shapes tasks. However, the within-group analysis demonstrated that participants in both groups exhibited significantly better performance on the alphabet task compared to the shapes task.

The results obtained in the current study were consistent with the previous literature. It was hypothesized in this study that the participants would perform better on the alphabet task than on the shapes task. By examining the existing body of literature, it was determined that there is a notable distinction in working memory performance when processing verbal compared to non-verbal stimuli. Several studies have suggested that individuals tend to exhibit better recall for verbal information, such as letters, than for nonverbal information, such as shapes. Gathercole et al. (2004) implied that children performed better on verbal sequences tasks than on spatial task sequences. The main justification of verbal superior performance can be supported by the phonological loop's active role in rehearsing verbal information that facilitates retention and retrieval of information. Corsetti et al. (2009) also reinforced the phonological loop's better function in a meta-analysis study that highlighted better verbal memory and better performance than visual tasks in children and adults. Some other investigations supported the idea that phonological processing has an edge over abstract visual shapes (Swanson & Sachse-Lee, 2001; Gathercole & Baddeley, 1993).

However, it is well understood that the non-verbal information process executes through a visuospatial sketchpad that has a distinct mechanism and generally requires more effort to retain visuospatial information. Available literature explains that when people are involved in the shapes process, their performance declines, particularly when shapes are unusual and complex (Snyder & Barr, 2014). Because shape processing exerts a greater cognitive load and hinders a person's ability to recall visual stimuli arrangement, it consequently leads to poor working memory performance (Miyake et al., 2000; Vogel et al., 2001). This is crucial to understanding that the majority of participants disclosed at the end of the experiment that they faced difficulty in recalling shapes and found them complex compared to the alphabet task. These findings assist us in understanding some plausible explanations of participants' inferior performance in shape tasks.

We have used the alphabet task as a gold standard measure for setting a benchmark for assessing the validity of the working memory n-back task. The n-back alphabet task is a well-known and widely tested task. Both group's consistent performance on the alphabet task shows that in the young adult population, this task will provide consistent scores even across different samples. This benchmark serves as a yardstick for evaluating both the validity and reliability of this n-back working memory task. Furthermore, the consistency of the score was further validated with similar results for the shapes task as we were not able to find differences in performance between groups on shapes tasks either. Thus, the repeated-measure block randomization design underscores the validity of the n-back memory task, as it provides converging and discriminating results for two different tasks measuring the same cognitive ability.

Conclusions

The findings of the study were consistent with the existing literature and revealed that there was no significant difference in both group performance on the alphabet and shapes tasks, particularly alphabet tasks, considering well-known metrics of working memory and participants' similar performance on both tasks provide a strong indication of our designed task as a valid measure. Meanwhile, the within-group comparison undoubtedly reflects the difference in the better performance of both groups in the alphabet task than in the shapes task. Therefore, between groups and within groups, similar performance provides us strong evidence to claim our developed task as a valid and reliable tool for the assessment of working memory.

Limitations

We acknowledge some of the limitations of our study. Firstly, the participants were recruited only from one university (Gift University). This may limit the generalizability of our findings to a larger population. However, current study task requires further evaluation on variate of populations. Secondly, this study only measured correct and incorrect responses as the primary outcome measures. Although reaction was not include in current analysis, we would recommend that future studies evaluate reaction tiem alongside correct and incorrect responses for a more comprehensive and precise assessment of n-back working memory performance.

References

Agbangla, N., Audiffren, M., Pylouster, J., & Albinet, C. (2022). Load-dependent prefrontal cortex activation was assessed by continuous-wave near-infrared spectroscopy during two executive tasks with three cognitive loads in young adults. *Brain Sciences*, *12*(11), 1462. https://doi.org/10.3390/brainsci12111462

Almodóvar-Payá, C., Guardiola-Ripoll, M., Giralt-López, M., Gallego, C., Salgado-Pineda, P., Miret, S., ... & Fatjó-Vilas, M. (2022). Nrn1 gene as a potential marker of early-onset schizophrenia: evidence from genetic and neuroimaging approaches. *International Journal of Molecular Sciences*, *23*(13), 7456. https://doi.org/10.3390/ijms23137456

Baddeley, A. D. (1986). *Working Memory.* Oxford University Press

Baddeley, A. D., Hitch, G. J., & Bower, G. A. (1974). Recent advances in learning and motivation. *Working Memory*, *8*, 647-667. https://doi.org/10.1016/s0079-7421(08)60452-1

Blacker, K. J., Negoita, S., Ewen, J. B., and Courtney, S. M. (2017). N-back versus complex span working memory training. J. *Cogn. Enhanc. 1*, 434–454. https://doi.org/10.1007/s41465-017-0044-1

Brouwer, A.-M., Hogervorst, M. A., van Erp, J. B. F., Heffelaar, T., Zimmerman, P. H., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural Engineering*, *9*(4), 045008. https://doi.org/10.1088/1741-2560/9/4/045008

Carretti, B., Borella, E., & Cornoldi, C. (2009). The role of verbal working memory in the comprehension of text: A study with young adults and elderly individuals. *Memory & Cognition*, *37*(1), 80–86. https://doi.org/10.1016/j.lindif.2008.10.002

Carter, C. S., Perlstein, W., Ganguli, R., Brar, J., Mintun, M., & Cohen, J. D. (1998). Functional hypofrontality and working memory dysfunction in schizophrenia. *The American Journal of Psychiatry*, *155*(9), 1285–1287. https://doi.org/10.1176/ajp.155.9.1285

Chen, Y.-N., Mitra, S., & Schlaghecken, F. (2008). Sub-processes of working memory in the N-back task: an investigation using ERPs. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *119*(7), 1546–1559. https://doi.org/10.1016/j.clinph.2008.03.003

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol. 64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dong, S., Reder, L. M., Yao, Y., Liu, Y., & Chen, F. (2015). Individual differences in working memory capacity are reflected in different ERP and EEG patterns for task difficulty. *Brain Research*, *1616*, 146–156. https://doi.org/10.1016/j.brainres.2015.05.003

Fahey, S., Charette, L., Francis, C., & Zheng, Z. (2018). Multisensory integration of signals for bodily self-awareness requires minimal cognitive effort. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale*, 72(4), 244-252. https://doi.org/10.1037/cep0000152

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory test battery for children*. Psychological Corporation.

Gathercole, S. E., Logie, R. H., Alloway, T. P., & Willis, C. (2004). Working memory in children: A working memory test battery for children (WMTB-C). *The British Journal of Educational Psychology*, 74(2), 245-266.

Gevins, A. S., Bressler, S. L., Cutillo, B. A., Illes, J., Miller, J. C., Stern, J., & Jex, H. R. (1990). Effects of prolonged mental work on functional brain topography. *Electroencephalography and Clinical Neurophysiology*, *76*(4), 339–350. https://doi.org/10.1016/0013-4694(90)90035-i

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(19), 6829–6833. https://doi.org/10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition*, *42*(3), 464–480. https://doi.org/10.3758/s13421-013-0364-z

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the N-back task: a question of construct validity. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(3), 615–622. https://doi.org/10.1037/0278-7393.33.3.615

Ke, Y., Wang, N., Du, J., Kong, L., Liu, S., Xu, M., ... & Ming, D. (2019). The Effects of Transcranial Direct Current Stimulation (tDCS) on Working Memory Training in Healthy Young Adults. *Frontiers in Human Neuroscience*, *13*, 19. https://doi.org/10.3389/fnhum.2019.00019

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358. https://doi.org/10.1037/h0043688

Loschky, L., Ringer, R., Johnson, A., Larson, A., Neider, M., & Kramer, A. (2014). Blur detection is unaffected by cognitive load. *Visual Cognition*, *22*(3-4), 522-547. https://doi.org/10.1080/13506285.2014.884203

Mackworth, J. F. (1959). Paced memorizing in a continuous task. *Journal of Experimental Psychology*, *58*(3), 206–211. https://doi.org/10.1037/h0049090

Miller, K. M., Price, C. C., Okun, M. S., Montijo, H., & Bowers, D. (2009). Is the n-back task a valid neuropsychological measure for assessing working memory? *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *24*(7), 711–717. https://doi.org/10.1093/arclin/acp063

Miyake, A., Friedman, N. P., Rettinger, D. A., Raymundo, L., & Duran, R. (2000). How are working memory, executive functioning, and fluid intelligence related? *Frontiers in Psychology*, *6*, 1–15.

Oberauer, K. (2005). Binding and inhibition in working memory: individual and age differences in short-term recognition. *Journal of Experimental Psychology. General*, *134*(3), 368–387. https://doi.org/10.1037/0096-3445.134.3.368

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 46–59. https://doi.org/10.1002/hbm.20131

Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging. *Human Brain Mapping, 25*, 46–59. https://doi.org/10.1002/hbm.20131

Parker, T., Huang, Y., Raghu, A., FitzGerald, J., Green, A., & Aziz, T. (2020). Dorsal root ganglion stimulation modulates cortical gamma activity in the cognitive dimension of chronic pain. *Brain Sciences*, *10*(2), 95. https://doi.org/10.3390/brainsci10020095

Pergher, V., Wittevrongel, B., Tournoy, J., Schoenmakers, B., & Van Hulle, M. M. (2018). N-back training and transfer effects revealed by behavioural responses and EEG. *Brain and Behavior*, *8*(11), e01136. https://doi.org/10.1002/brb3.1136

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2017). Comparison of the working memory load in N-back and working memory span tasks using EEG frequency band power and P300 amplitude. *Frontiers in Human Neuroscience*, *11*, 6. https://doi.org/10.3389/fnhum.2017.00006

Schmidt, H., Jogia, J., Fast, K., Christodoulou, T., Haldane, M., Kumari, V., & Frangou, S. (2009). No gender differences in brain activation during the N-back task: an fMRI study in healthy individuals. *Human Brain Mapping*, *30*(11), 3609–3615. https://doi.org/10.1002/hbm.20783

Snyder, H. R., & Barr, D. J. (2014). The role of working memory in visual enumeration. *Cognitive Psychology*, *70*, 101-126.

Swanson, H. L., & Sachse-Lee, C. (2001). A selective review of interventions to improve working memory in children with learning difficulties. *Educational Psychology Review*, *13*(2), 207-228.

Tang, H., Wu, Q., Li, S., Fang, Y., Yang, Z., Wang, B., ... & Liu, P. (2021). Visuospatial but not verbal working memory deficits in adult patients with neurofibromatosis type 1. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.751384

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92–114. https://doi.org/10.1037/0096-1523.27.1.92

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective & Behavioral Neuroscience*, *3*(4), 255–274. https://doi.org/10.3758/cabn.3.4.255

Supplement Material

Figure 1

Alphabet Used in Working Memory Alphabet Task

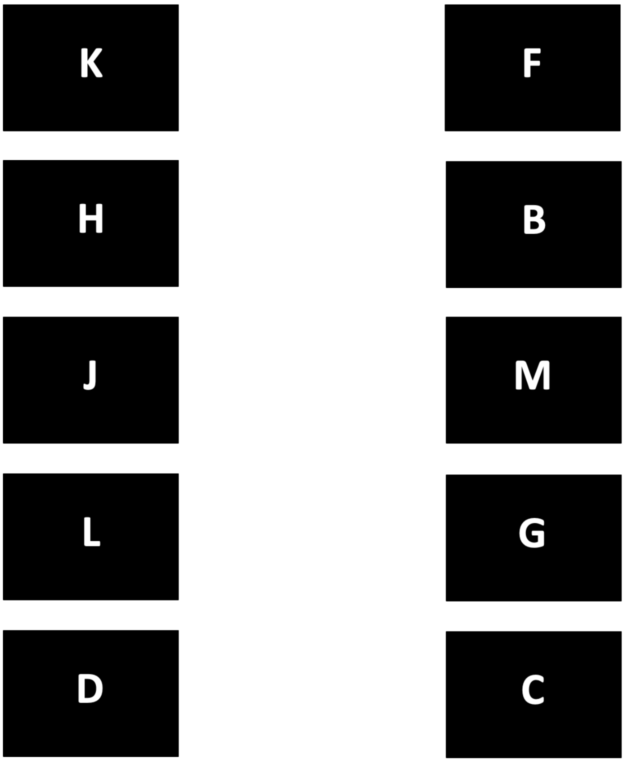


Figure 2

Shapes Used in Working Memory Shapes Task