

Measuring working memory in longitudinal designs: Practice effects exemplified by the Backward Digit Span (Poster Abstract)

Alexander Witzki^a, Karsten Sönnichsen^a, Ursa Nagler-Nitzschner^a, Dieter Leyk^{a, b}

Purpose

Working memory is crucial in numerous everyday tasks. It is essential for cognitive functions such as learning, decision making etc. The valid measurement of working memory is a prerequisite for applied research on cognitive ability, such as studies which evaluate effects of diseases on cognitive performance, measure effects of cognitive training, or assess success and side effects of therapies. For example, anesthesia can have longer-term effects on cognitive performance in addition to its intended effect during an operation [1]. **Backward Digit Span** can be used to measure these effects and their duration. It assess working memory, the central indicator of cognitive ability [3].

Longitudinal designs can be used to assess changes in cognitive ability over time. This approach allows for example the assessment of rehabilitation on cognitive as well as physical ability. Due to the necessary repetition of measurements in longitudinal designs, test instruments are needed that reliably measure the underlying constructs. In particular, the mere fact of test repetition must not lead to an improvement in performance. Such practice effects can bias longitudinal studies. This requirement can pose a challenge in cognitive performance assessment as participants may remember previous measurements and their responses. Therefore, test instruments and tasks must be designed in such a way that practice effects ideally do not occur or at least can be controlled. Test evaluation must provide evidence to this effect for each test implemented.

The aim of the present study was to evaluate the influence of repeated testing of working memory performance with a version of Backward Digit Span implemented specifically for use in the Bundeswehr Central Hospital Koblenz.

Method

In the **Backward Digit Span**, 3-8 digits are presented one at a time on a tablet PC. Participants' task is to repeat the digits in reverse order by tapping on a numeric keypad presented on the screen. Individual items (i.e., sequences of digits) are automatically generated for each test.

Design and sample

Test sessions were conducted on five consecutive days at a similar time. 35 persons were informed about aim and purpose of the study and gave their consent before the first test session. After completion of the test, a standardized interview on the test procedure (questions about distractions/disturbances, concentration problems, change of strategy, etc.) was conducted. Data of three persons had to be excluded from the evaluation because they only took part in one test session (two persons) or indicated a change of strategy in one session that could potentially influence test performance (one person). The remaining 32 participants had an average age of 36.9 ± 10.8 years, half were women and half men.

Statistics

Reliability of the Backward Digit Span (retest reliability) was assessed by Pearson correlations between test sessions. Differences in performance were analyzed using paired sample t-tests for successive test sessions (days 1 & 2, 2 & 3, etc.) and standardized mean difference effect sizes (Cohen's *dz*) were calculated. Standard deviations and standard errors are reported as measures of dispersion in the text and Figure 1, respectively.

Results

Retest reliability across all Backward Digit Span session was $r = 0.75-0.85$. As shown in Figure 1, a significant improvement in performance could only be shown between the first two test sessions ($t_{1,2} = 4.49$, $p < .001$, $dz = 0.79$; $t_{2,3} = 1.68$, $p = .31$, $dz = 0.30$; $t_{3,4} = 0.60$, $p = .56$, $dz = 0.11$; $t_{4,5} = 1.14$, $p = .52$, $dz = 0.20$).

Discussion

The Backward Digit Span is a reliable method for measuring cognitive performance. The reliabilities we found in the present study correspond to those reported in the literature [2].

Contrary to standard procedure in cognitive psychology, the present study did not contain a training phase in which study participants were familiarized with the tests and trained them in practice sessions. Accordingly, this could be considered a flaw of the current study design. In this case, however, it is a feature, i.e. an intentional part of the study design: It serves as an indicator of the amount of training required. The first test session can be equated

^a Bundeswehr Institute for Preventive Medicine, Andernach/Koblenz, Germany

^b German Sport University Cologne, Cologne, Germany

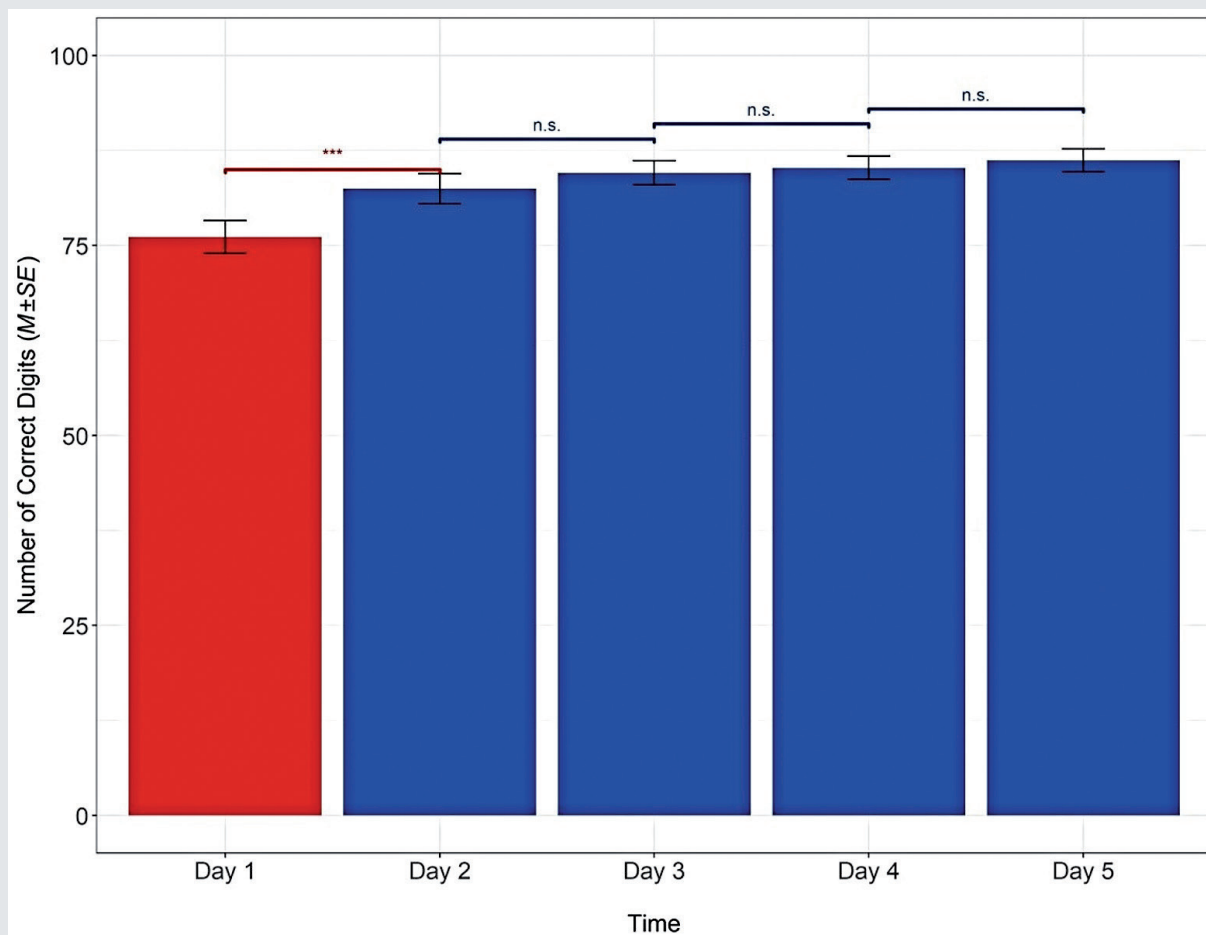


Figure 1: Mean and standard error of correctly reproduced digits in the Backward Digit Span at five consecutive days. (Note. n.s. = $p > .1$. *** $p < .001$.)

with the usual training phase. The significant improvement between the first two sessions can, thus, be explained by the lack of familiarity with the procedure – none of the study participants had any previous experience with the Backward Digit Span. This effect can therefore be interpreted as being due to exposure to a new task rather than a practice effect stemming from an improvement in working memory performance.

Conclusion

Overall, the present results demonstrate that performance improvements can be observed even in established, reliable and valid tests that measure stable characteristics. This emphasizes the necessity of an appropriate test evaluation and is of utmost importance in longitudinal designs.

With regard to the tested version of the Backward Digit Span, the results indicate that it can be used for longitudinal studies if a training session is scheduled before the first test session. Thus, a test measuring cognitive performance over time is available that can be used in both

clinical (e.g., effects of anesthesia) and occupational contexts (e.g., effects of fatigue, stress, heat).

References

1. Caza N, Taha R, Qi Y, Blaise G: The effects of surgery and anesthesia on memory and cognition. In W. S. Sossin, J.-C. Lacaille, V. F. Castellucci, & S. Belleville (Eds.), *Progress in brain research: Essence of memory 2008*; 169: 409–422.
2. Conway ARA, Kane MJ, Bunting M F, Hambrick DZ, Wilhelm O, Engle RW: Working memory span tasks. A methodological review and user's guide. *Psychonomic Bulletin and Review* 2005; 12(5): 769–786.
3. Oberauer K, Süß HM, Schulze R, Wilhelm O, Wittmann WW: (2000). Working memory capacity – Facets of a cognitive ability construct. *Personality and Individual Differences* 2000; 29(6): 1017–1045.

For the authors

Dr. Alexander Witzki

Bundeswehr Institute for Preventive Medicine

E-Mail: InstPraevMedBwA3@bundeswehr.org

Poster presentation at the 52. Annual Congress of the German Society for Military Medicine and Military Pharmacy, 15. October 2021 in Coblenz, Germany