Advances in Transportation, Vol. 60, 2022, 297–305
https://doi.org/10.54941/ahfe1002460

Investigating the Influence of Working Memory Processes on the Box Task combined with a Detection Response Task

Daniel Trommler[1], Tina Morgenstern[1], Cornelia Hollander[1], Ines Karl[2], Frederik Naujoks[2], Josef F. Krems[1], and Andreas Keinath[2]

[1]Cognitive & Engineering Psychology, Chemnitz University of Technology, Germany
[2]BMW Group, Munich, Germany

ABSTRACT

The assessment of task demand caused by in-vehicle systems is crucial to avoid distraction while driving. The Box Task (BT) in combination with a tactile Detection Response Task (DRT) provides a method for measuring both visual-manual and cognitive secondary task demand. In the present study, the impact of cognitive, auditory-verbal tasks on the BT + DRT performance was investigated. Thirty-two participants had to perform an easy as well as a difficult version of an n-back task and a memory scanning task while simultaneously performing the BT + DRT. There was only a slight effect of cognitive task demand on the BT performance parameters, while the DRT proved to be highly sensitive to cognitive task demand. Therefore, it is assumed that the method is suitable for a differentiated measurement of task demand dimensions.

Keywords: Driver distraction, Evaluation methods, Box task, Detection response task, Cognitive demand

INTRODUCTION

Driving involves a complex interplay of different cognitive, visual and psychomotor skills (Young and Regan 2007). Especially novice drivers are highly strained while driving, e.g. due to inflexible visual search strategies (Crundall and Underwood 1998). However, with increased practice, drivers often shift their attention to non-driving related secondary tasks, such as interacting with modern in-vehicle systems (e.g., infotainment systems) or smartphones (Huemer et al. 2018; Kubitzki and Fastenmeier 2016). Several statistics indicate the high potential for accidents due to the associated dual-task interference (Carney et al. 2015). As the number of electronic devices and in-vehicle systems will further increase in the next years, this problem is expected to remain an important issue for road safety.

Dual-task interference can be explained by the multiple resource model (Wickens 2002), which distinguishes four resource dimensions, each with two levels: Stages of information processing (perception & cognition vs. responding), modalities (auditory-verbal vs. visual-manual), visual channels (focal vs. ambient) and processing codes (spatial vs. verbal). It is assumed that tasks are most likely to interfere when they require the same level of a dimension than when they require different levels (Wickens 2002).

To minimize the dual-task interference while driving, it is crucial to estimate the distraction potential of in-vehicle systems early in the development process. For this purpose, various methods are used, which are either based on self-reports (e.g., NASA-TLX; Hart and Staveland 1988) or on performance measures (e.g., Lane Change Test; Mattes and Hallén 2009). However, modern in-vehicle systems are designed for a multimodal interaction, which can lead to both visual-manual and cognitive distraction (Strayer 2015). Thus, an assessment method should not only be easy to use, but should also consider multiple dimensions of task demand separately.

The Box Task (BT) in combination with a Detection Response Task (DRT) is an easy-to-use method, which allows a differentiated assessment of secondary task demand. The method is based on the Dimensional Model of Driver Demand (Young et al. 2016), which distinguishes between physical (i.e., visual-manual) and cognitive task demand. Previous studies compared the BT + DRT with established approaches (i.e., Lane Change Test and driving simulation task) and showed that the BT is particularly sensitive to visual-manual task demand while the DRT responds well to cognitive task demand (Morgenstern et al. 2020a). Furthermore, it has been shown that the BT + DRT is sensitive to task demand associated with different artificial and natural secondary tasks and varying levels of task difficulty (Morgenstern et al. 2020b).

In recent years, voice-based in-vehicle systems have been increasingly used, because they are associated with higher satisfaction and lower visual-manual task demand (Sodnik et al. 2008). Thus, the distraction modality changes from visual-manual to auditory-verbal. However, using voice-based in-vehicle systems while driving is associated with higher cognitive demand (Engström et al. 2017), which can result in higher reaction times (Lee et al. 2001) and impaired gaze behavior while driving (Reimer and Mehler 2013; Trbovich and Harbluk 2003). Aim of the present study was to investigate whether and to what extent cognitive tasks involving different working memory processes affect the performance in the BT + DRT.

METHOD

Participants and Experimental Design

Thirty-two participants (N = 11 male, N = 20 female, N = 1 non-binary) with an average age of 27 years (SD = 6.7, Min = 19, Max = 45) and an annual average mileage of 6,700 km (SD = 6,735) took part in the study. They received either course credit or 10 euros as compensation for their participation.

In the study, a mixed design with two independent variables was applied. The variable *cognitive secondary task* was varied as a within-subjects factor with five levels: no secondary task, easy and difficult version of secondary

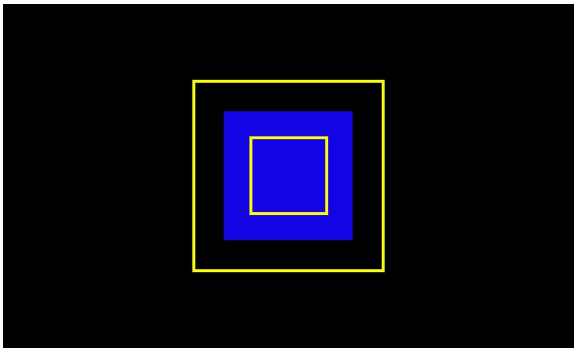

Figure 1: The box task consists of a blue box dynamically changing its lateral position and size. The box has to be held between two yellow boundaries.

task 1 as well as easy and difficult version of secondary task 2. The variable *duration of practice* of the BT + DRT was varied as a between-subjects factor with three levels: 1.5, 3 and 4.5 minutes of practicing. Both independent variables were balanced. However, due to space constraints, the analysis concerning the practice is not included in this paper.

Material

BT + DRT

A dynamic blue box, which changes its lateral position as well as its size following a sinusoidal pattern, was shown on a 23" monitor in front of the participant. This two-dimensional tracking task simulated a car-following scenario where the driver has to maintain the lane and the headway. Further, the participants were instructed to hold the blue box between two yellow squares representing the boundaries (see Figure 1) by adjusting box position and box size using a Logitech MOMO force-feedback steering wheel and a gas pedal. The experimental setup followed the description in Trommler et al. (2021).

Simultaneously, a detection response task (DRT) with a tactile stimulus was performed according to the ISO standard (see ISO 17488 2016). Participants needed to respond as fast and accurately as possible to a stimulus that randomly appeared every three to five seconds from a vibration module on their shoulder by pressing a button on the steering wheel.

Secondary Tasks

Two cognitive, auditory-verbal tasks that are based on different working memory processes were used. During the *n-back task* (Kirchner 1958), a set of numbers was presented acoustically to the participants. After each item, participants were asked to recall and report the number that was previously presented at a specific position. In the easy task condition, the new number had to be repeated immediately (0-back task); in the difficult task condition, the number two positions before needed to be repeated (2-back task).

In addition, a *memory scanning task* (MST; Sternberg 1966) was used. Participants were instructed to memorize a set of two-syllable, lesser-known German city names in written form in the learning phase before the dual-task conditions started. The set contained two city names for the easy task condition and five city names for the difficult task condition. During the dual-task conditions, participants were presented city names acoustically. After each presentation, participants had to answer with "yes" or "no" depending on whether the city was included in the memorized set or not.

Mental Workload

The "rating scale mental effort" (RSME; Zijlstra 1993) was used to assess self-reported mental workload. The RSME is a one-dimensional scale consisting of a line ranging from 0 (*absolutely no effort*) to 150 (*extreme effort*) with nine verbal anchor points between.

Procedure

After giving informed consent, demographic variables were collected using a pre-test questionnaire. This was followed by a written instruction on the BT + DRT with a subsequent practice trial. Three baseline drives (i.e., performing the BT + DRT without secondary task engagement) were conducted at the beginning (initial baseline), in the middle (intermediate baseline), and at the end (final baseline) of the experimental session. Between the baseline drives, two test blocks, each with an easy and a difficult version of one of the cognitive tasks, were performed as dual-task conditions (i.e., performing the BT + DRT and the secondary task in parallel). Before each dual-task condition, participants received a written instruction on the secondary task and were able to practice it for approximately 30 seconds. After each task condition, the mental workload was rated. In total, the experimental session lasted approximately 60 minutes.

Dependent Measures

The extent of the standard deviation from the ideal box position (SDLatP) and box size (SDLongP) were calculated to assess BT performance. For the DRT, the hit rate and the mean reaction time for hits were calculated according to the ISO standard (see ISO 17488, 2016). A reaction was considered as a hit if the reaction occurred within 0.1 to 2.5 seconds after stimulus onset. The average percentage of correct answers (accuracy rate) was calculated as the performance measure for the secondary tasks.

RESULTS

BT + DRT Performance

Participants' BT + DRT parameters across the task conditions were examined using non-parametric Friedman tests (due to violations regarding assumptions of normal distribution). For effect sizes, Kendall's W was calculated. Post-hoc tests were Dunn-Bonferroni-corrected.

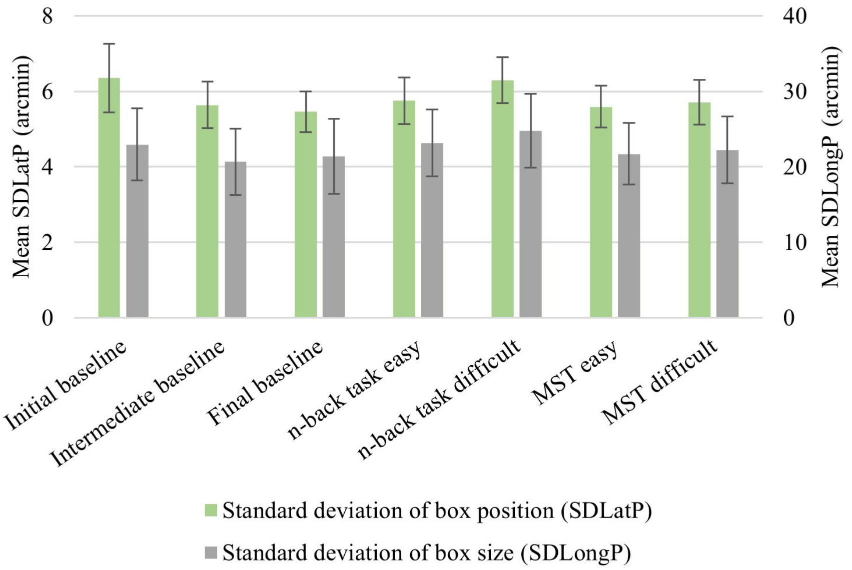

Figure 2: Box Task performance across the task conditions. Error bars represent 95th confidence intervals.

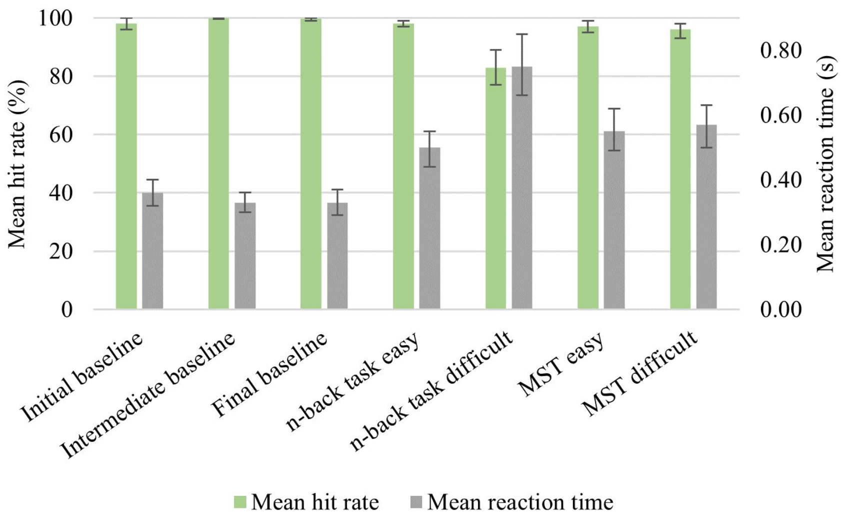

Figure 3: DRT performance across the task conditions. Error bars represent 95th confidence intervals.

The means of the BT + DRT parameters across the task conditions are displayed in Figure 2 and Figure 3. The task conditions differed significantly regarding the variability of the lateral box position (SDLatP) (χ^2 (6) = 13.15, p = .041, W = .07). The post-hoc analysis revealed significantly higher standard deviations for the difficult n-back task compared to the intermediate baseline drive, the final baseline drive, the easy n-back task and the easy

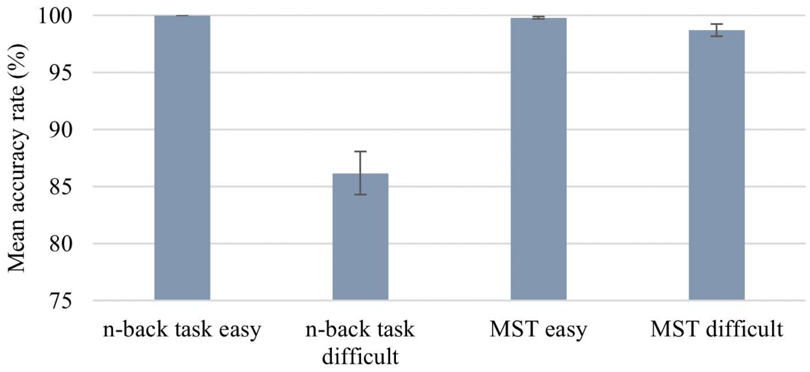

Figure 4: Secondary task performance across the dual-task conditions. Error bars represent 95th confidence intervals.

MST. However, regarding the variability of the box size (SDLongP), no significant differences could be found across the task conditions (χ^2 (6) = 11.21, p = .082, W = .06).

Moreover, the DRT hit rate differed significantly across the task conditions (χ^2 (6) = 86.77, p < .001, W = .47). There were significant differences between the baseline drives and the dual-task conditions (except for the initial baseline drive and the easy MST). In addition, the difficult n-back task differed significantly from the easy n-back task as well as the easy and difficult MST. Regarding the mean reaction time within the DRT, the Friedman test showed a significant result (χ^2 (6) = 154.18, p < .001, W = .83). Similar to the results regarding the hit rate, all baseline drives differed significantly from the dual-task conditions. Furthermore, the difficult n-back task differed significantly from the easy n-back task as well as the easy and difficult MST.

Secondary Task Performance

The results regarding secondary task performance are displayed in Figure 4. For the easy n-back task and the easy and difficult MST, the mean accuracy rate was almost 100%. During the difficult n-back task, participants showed a lower accuracy (M = 86%, SD = 11).

Mental Workload

The RSME ratings across the task conditions are shown in Figure 5. A repeated measures ANOVA with Greenhouse-Geisser correction was calculated.

Although the assumption of normal distribution was partially violated, the robustness of the method can be assumed due to the sample size. The RSME ratings differed significantly across the task conditions (F(3.88,120.43) = 117.70, p < .001, η_p^2 = .79). Post-hoc comparisons showed that the initial baseline drive was rated as significantly more demanding compared to the intermediate and final baseline drive. Further, all baseline conditions were rated as significantly less demanding than the dual-task conditions. Similar to the results regarding the DRT, the difficult n-back task was

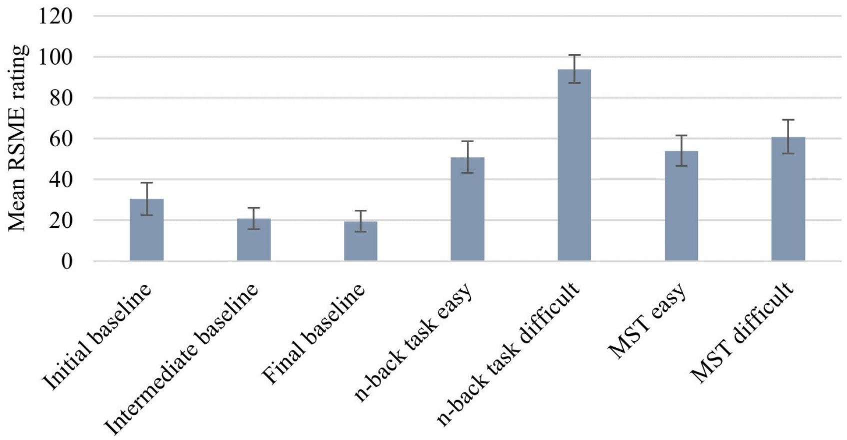

Figure 5: Mental workload ratings across the task conditions. Error bars represent 95th confidence intervals.

rated as significantly more demanding than the easy n-back task as well as the easy and the difficult MST.

DISCUSSION

The BT + DRT is a laboratory method that captures different dimensions of task demand caused by using in-vehicle systems or portable electronic devices while driving. As the number of voice-based in-vehicle interactions will continue to increase in the coming years, it is important to assess task demand associated with different working memory processes and auditory-verbal modality. In the present experimental study, the impact of cognitive tasks on the BT + DRT performance was investigated using the n-back task and the MST in an easy and difficult version.

The results indicate a low sensitivity of the BT for cognitive task demand. Regarding the variability of box size, we could not find significant differences across the task conditions; regarding the variability of box position, there was a significantly higher variability for the difficult n-back task compared to a few other task conditions. In contrast, significant differences were revealed in both the DRT hit rate and reaction time. Specifically, the baseline conditions differed significantly from almost every dual-task condition. This confirms previous findings that the DRT is highly sensitive to cognitive task demand (Stojmenova and Sodnik 2018). Moreover, the RSME ratings show that the difficult n-back task condition was highly demanding for the participants, which is also reflected in a lower accuracy rate.

The significant impact of the difficult n-back task on the variability of the box position, which is rather used as a measure for visual-manual task demand, may be explained by the multiple resource model (Wickens 2002). It assumes that perception and cognition are located on the same level of the processing stage. The BT is a continuous perceptual tracking task. Thus, highly demanding tasks, such as the difficult n-back task, can lead to interference. However, this effect may be reduced due to the different modality levels of the BT (visual-manual) and the n-back task (auditory-verbal). Nevertheless, further research is needed to investigate alternative BT parameters, such as specific patterns in lateral and longitudinal adjustments, that are able to better differentiate between the effects of visual-manual and cognitive task demand, especially with regard to difficult cognitive tasks.

Furthermore, except for the difficult n-back task, the DRT and RSME parameters indicate that the cognitive task demand, resulting from the secondary tasks used for this study, was rather homogeneous. In future studies, more difficult tasks should be used to achieve more valid conclusions regarding the impact of cognitive tasks that vary in their level of difficulty on the BT + DRT.

In sum, the results confirm the assumption that the BT performance parameters are not or only slightly affected by cognitive tasks. Only very demanding cognitive tasks impair BT performance, which, however, is also in line with the finding that performance of not automatized tasks, such as the BT, is affected by cognitive load (Engström et al. 2017). Thus, it can be concluded that the BT + DRT is able to differentiate between visual-manual and cognitive task demand.

ACKNOWLEDGMENT

This research was supported by the BMW Group.

REFERENCES

Carney, C., McGehee, D., Harland, K., Weiss, M. and Raby, M. (2015). Using Naturalistic Driving Data to Assess Vehicle-to-Vehicle Crashes Involving Fleet Drivers. June 2015. Available from: https://trid.trb.org/view/1359341 [accessed 9 February 2022].

Crundall, D.E. and Underwood, G. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics*, 41(4), pp. 448–458.

Engström, J., Markkula, G., Victor, T. and Merat, N. (2017). Effects of Cognitive Load on Driving Performance: The Cognitive Control Hypothesis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 59(5), pp. 734–764.

Hart, S.G. and Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: *Advances in Psychology*. Elsevier, pp. 139–183. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0166411508623869 [accessed 9 February 2022].

Huemer, A.K., Schumacher, M., Mennecke, M. and Vollrath, M. (2018). Systematic review of observational studies on secondary task engagement while driving. *Accident Analysis & Prevention*, 119, pp. 225–236.

Kirchner, W.K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55(4), pp. 352–358.

Kubitzki, J. and Fastenmeier, W. (2016). *Ablenkung durch moderne Informations- und Kommunikationstechniken und sozial Interaktion bei Autofahrern*. Unterföhring, Germany.

Lee, J.D., Caven, B., Haake, S. and Brown, T.L. (2001). Speech-Based Interaction with In-Vehicle Computers: The Effect of Speech-Based E-Mail on Drivers' Attention to the Roadway. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(4), pp. 631–640.

Mattes, S. and Hallén, A. (2009). Surrogate distraction measurement techniques: The lane change test. *Driver distraction: Theory, effects, and mitigation*, pp. 107–121.

Morgenstern, T., Wögerbauer, E.M., Naujoks, F., Krems, J.F. and Keinath, A. (2020a). Measuring driver distraction – Evaluation of the box task method as a tool for assessing in-vehicle system demand. *Applied Ergonomics*, 88, p. 103181.

Morgenstern, T., Trommler, D., Forster, Y., Naujoks, F., Hergeth, S., Krems, J.F. and Keinath, A. (2020b). Measuring Driver Distraction with the Box Task – A Summary of Two Experimental Studies. In: Krömker, H., ed. *HCI in Mobility, Transport, and Automotive Systems. Driving Behavior, Urban and Smart Mobility*. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp.51–60. Available from: http://link.springer.com/10.1007/978-3-030-50537-0_5 [accessed 9 February 2022].

Reimer, B. and Mehler, B. (2013). The Effects of a Production Level "Voice-Command" Interface on Driver Behavior: Summary Findings on Reported Workload, Physiology, Visual Attention, and Driving Performance. p. 19.

Sodnik, J., Dicke, C., Tomažič, S. and Billinghurst, M. (2008). A user study of auditory versus visual interfaces for use while driving. *International Journal of Human-Computer Studies*, 66(5), pp. 318–332.

ISO 17488 (2016). Road vehicles - Transport information and control systems - Detection Response Task (DRT) for assessing attentional effects of cognitive load in driving.

Sternberg, S. (1966). High-Speed Scanning in Human Memory. *Science*, 153(3736), pp. 652–654.

Stojmenova, K. and Sodnik, J. (2018). Detection-Response Task - Uses and Limitations. *Sensors*, 18(2), p. 594.

Strayer, D.L. (2015). Is the Technology in Your Car Driving You to Distraction? *Policy Insights from the Behavioral and Brain Sciences*, 2(1), pp. 157–165.

Trbovich, P. and Harbluk, J.L. (2003). Cell phone communication and driver visual behavior: the impact of cognitive distraction. In: *CHI '03 extended abstracts on Human factors in computing systems - CHI '03*. CHI '03 extended abstracts. Ft. Lauderdale, Florida, USA: ACM Press, p. 728. Available from: http://portal.acm.org/citation.cfm?doid=765891.765954 [accessed 9 February 2022].

Trommler, D., Morgenstern, T., Wögerbauer, E.M., Naujoks, F., Krems, J.F. and Keinath, A. (2021). The box task - a method for assessing in-vehicle system demand. *MethodsX*, 8, p. 101261.

Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2), pp. 159–177.

Young, K. and Regan, M. (2007). Driver distraction: A review of the literature. In: Faulks, I.J., Regan, M., Stevenson, M., Brown, J., Porter, A. & Irwin, J.D. (Eds.). *Distracted driving*. Sydney, NSW: Australasian College of Road Safety, pp. 379–405. Available from: https://www.researchgate.net/publication/251312763_Driver_Distraction_A_Review_of_the_Literature [accessed 9 February 2022].

Young, R., Seaman, S. and Hsieh, L. (2016). The Dimensional Model of Driver Demand: Visual-Manual Tasks. *SAE International Journal of Transportation Safety*, 4(1), pp. 33–71.

Zijlstra, F. R. H. (1993). Efficiency in work behaviour: A design approach for modern tools. Technical University of Delft, Delft.