The Box Task: A tool to design experiments for assessing visuospatial working memory

Roy P. C. Kessels1,2,3,4, Albert Postma5,6

Abstract The present paper describes the Box Task, a paradigm for the computerized assessment of visuospatial working memory. In this task, hidden objects have to be searched by opening closed boxes that are shown at different locations on the computer screen. The set size (i.e., number of boxes that must be searched) can be varied and different error scores can be computed that measure specific working memory processes (i.e., the number of within-search and between-search errors). The Box Task also has a developer’s mode in which new stimulus displays can be designed for use in tailored experiments. The Box Task comes with a standard set of stimulus displays (including practice trials, as well as stimulus displays with 4, 6, and 8 boxes). The raw data can be analyzed easily and the results of individual participants can be aggregated into one spreadsheet for further statistical analyses.

Keywords Working memory · Visuospatial search · Neuropsychological testing · Experimental software · Spatial memory

Introduction

Working memory relates to the online maintenance, updating, and manipulation of information for a brief period of time (Baddeley & Hitch, 1974). It is considered to be a distinct memory system (Squire, 2004), and working memory processing is a crucial component in theories of executive function (Shallice & Cooper, 2011), as well as episodic memory formation and retrieval (Baddeley, 2012). Neuroimaging and lesion studies have shown that dorsolateral prefrontal and posterior parietal brain regions are important for working memory processing (Eriksen, Vogel, Lanser, Bergström, & Nyberg, 2015; Kessels, Postma, Wijnalda, & De Haan, 2000). Additionally, working memory impairments have been demonstrated in a variety of brain diseases and psychiatric disorders, such as attention deficit hyperactivity disorder (ADHD; Kasper, Alderson, & Hudec, 2012), autism (Barendse, Hendriks, Jansen, et al., 2013), stroke (Hochstenbach, Mulder, van Limbeek, Donders, & Schoonderwaldt, 1998), traumatic brain injury (Dunning, Westgate & Adlam, 2016), schizophrenia (Forbes, Carrick, McIntosh, & Lawrie, 2009), Parkinson’s disease (Pagonabarraga & Kulisevsky, 2012), or dementia (Geman & Kinsella, 2005). Furthermore, working memory function has been used as an outcome measure for cognitive interventions (Von Bastian & Oberauer, 2014).

To measure working memory processing, a number of paradigms have been developed (see Postma & van der Ham, 2016, for an overview), such as span tasks – either verbal (Blankenship, 1938) or spatial (Berch, Krikorian, & Huha, 1998) in nature – or running working memory tasks such as n-back paradigms (see...
Jaeggi, Buschkuehl, Perrig, & Meier, 2010). In the visuospatial domain, several working memory paradigms rely on visual search: the location of a stimulus is searched for, feedback is given, and for a short period of time, maintenance of the location will be required (Postma & Van der Ham, 2016). The Executive Golf task (Feigenbaum, Polkey & Morris, 1996) is an example of such a visuospatial working memory task. In this task, the participant has to putt a ball into holes of a golf course. The correct hole has to be searched for by first clicking it, then feedback is given as to whether or not this hole is correct, then the ball is put, and the next hole must be searched for. The participant is instructed to remember not to search for a hole that already contains a ball from a previous trial. Two errors can be made here. First, a within-search error occurs if a subject returns to a previously selected hole that was not correct within a search (i.e., for that trial/hole). Second, a between-search error occurs if a subject returns to a hole that already contains a ball (i.e., was a target in a previous search trial). Different difficulty levels are presented, with an increasing number of holes.

The Spatial Working Memory subtest from the Cambridge Automated Neuropsychological Test Automated Battery (CANTAB; Robbins, James, Owen, et al., 1994) is based on the Executive Golf paradigm; here, participants have to search for blue tokens that are hidden in squares (representing boxes) shown at different locations on a computer screen. Boxes can be opened by clicking upon them, showing either the blue token (if it was hidden in that box) or just an open square (an empty box). If a blue token is found, the box closes again and the next token has to be searched for. Similar to the Executive Golf Task, within- and between-search errors can occur (when participants click a box already shown to be empty within that trial, or when a participant returns to a box that contains a blue token from a previous trial, respectively). The working memory load is determined by the number of boxes that have to be searched (4, 6, or 8).

While the aforementioned paradigms have their strengths and value, tasks such as the Executive Golf Task are no longer available. Furthermore, the presentation of stimuli and the working memory load cannot be modified in the existing paradigms. More importantly, the tasks only require the search for and maintenance of spatial locations, as the identity of the target in these tasks was always identical (balls, blue tokens). However, spatial search tasks in the real world do not only require the tracking of locations, but also knowing what can be found in a specific location. To overcome these limitations, we developed a computerized tool to setup, run, and analyze experiments on visuospatial working memory using the principles of previous visual search paradigms, whilst also adding item identity: the Box Task. More importantly, the tasks only require the search for and maintenance of spatial locations, as the identity of the target in these tasks was always identical (balls, blue tokens). However, spatial search tasks in the real world do not only require the tracking of locations, but also knowing what can be found in a specific location. To overcome these limitations, we developed a computerized tool to setup, run, and analyze experiments on visuospatial working memory using the principles of previous visual search paradigms, whilst also adding item identity: the Box Task. So far, the Box Task has been successfully applied in studies of healthy individuals of all age groups, as well as patients with mild cognitive impairment or Alzheimer’s dementia (Kessels, Meulenaer, Fernández & Olde Rikkert, 2010; Van Geldorp, Konings, Van Tilborg & Kessels, 2012), Korsakoff’s syndrome (Van Asselen, Kessels, Wester & Postma, 2005; Oudman, Van der Stigchel, Wester, Kessels & Postma, 2011), stroke (Van Asselen, Kessels, Neggers, et al., 2006), unilateral hippocampal epilepsy surgery (Kessels, Hendriks, Schouten, Van Asselen & Postma, 2004), women with Turner syndrome (Freriks, Verhaak, Sas, et al., 2015), and adolescents with autism spectrum disorder (Barendse, 2017).

**The Box Task**

In the Box Task, boxes are presented at different locations on a computer screen. These boxes have to be searched in order to find a hidden target object, shown at the bottom of the computer screen. Boxes are opened by clicking upon them, revealing either an empty box, a non-target object (which may be a target in a later trial), or the hidden target object. The box then closes again and the next box can be clicked. After a target has been found, a new target object is presented at the bottom of the computer screen. The number of target objects that have to be searched is typically the same as the number of boxes presented on the screen. If all targets have been found, a new set of boxes is presented at different locations. The number of boxes can vary (previous studies have used 3 boxes for practice purposes, using working memory loads of 4, 6, 8, to 10 for the actual experimental stimulus displays). Within-search errors are made when a previously opened “empty” box is clicked again; between-search errors occur if the participant returns to a box that already contains a target object from a previous search. Figure 1 shows a schematic overview of the Box Task paradigm.

To click the boxes, participants can use the mouse or a touch-sensitive screen (in which case the mouse pointer is not visible during a run). An experiment file (a text file with the extension .bex) contains multiple clusters (each defined in a separate .box file, consisting of multiple searches). The experiment, cluster, and searches can be easily made or changed using a GUI (graphical user interface). Raw data for each search are stored in a tab-delimited text file (the subject identifier with the extension .bof) and a tool is available to combine data from multiple participants into one aggregated output file for statistical analyses.

**Technical specifications and availability**

The Box Task was written in Microsoft Visual Basic. It is a 32-bit application that is compatible with all current 32- and 64-bit versions of Microsoft Windows (7, 8, 8.1, and 10). It requires minimal CPU processing capacity and does not involve time-critical responses or specific hardware, although the use of a touch-sensitive screen or Windows tablet is recommended for data collection in patients with cognitive impairment. A stimulus set is included that can be used for data collection, consisting of two practice clusters of 3 boxes, followed by two clusters of 4 boxes, 6 boxes, and 8 boxes.
Designing new experiments

As an experiment consists of a set of clusters (each consisting of multiple searches or trials), new clusters can be made in the developer’s mode (Fig. 2). When starting the program, an empty trial layout is shown. Here, boxes can be dragged from the Box frame into the Trial Layout square, placing them at different locations. After all boxes are placed, a target object can be selected from the Items frame showing six objects. Objects can be customized by right-clicking in the Items frame and selecting picture files (in .ico format) from the icon folder. An object can be dragged from the Items frame to one of the boxes in the trial layout. An object placed at the location of a box can be set as a target by right clicking on the object in the trial layout. Then, it also appears in the Target frame, with its number. Next, the trial layout must be copied to define the next search by clicking the Copy Trial button, showing a new tab with the previously placed boxes and the previous target item. Now, one can drag another object to one of the unfilled boxes. By right-clicking the new object, it is set as a target (note that the square indicating the target object moves from the previous target to the new target). The next trial can be made by clicking the Copy Trial button again, dragging a new object, etc. Parameters such as the start and end message can be set. The response time can be restricted under the Timing tab under Time for this layout (note that 0 results in an unlimited response time). The duration that a box is shown open for can also be set under the Timing tab. Screen background colors and size can be altered under the Screen tab. When all boxes are filled, the cluster must be saved by clicking the Save Cluster button. A new trial can be defined by selecting Clear Trial under the Trial menu at the top.

After several clusters have been defined, they can be combined into one experiment file. To define an experiment, click Experiment under the Run menu at the top. An empty experiment is shown under Clusters to run. By clicking the Add Cluster button, previously stored clusters (.box files) can be added to form an experiment. By clicking Save Experiment, the set of clusters will be saved in one experiment file (.bex). Note that under Extra in the Options menu, default settings can be set for a new experiment. The size of the objects must also be defined here under Icon Size; this should be done once for every computer on which the Box Task is used, as the actual object size depends on the resolutions settings in Windows and Behav Res (2018) 50:1981–1987 1983

Fig. 1 The Box Task: (a) a target object is shown at the bottom of the screen (the shopping basket), the box in the upper right of the screen is open, indicating that it does not contain the target; (b) the box in the lower left of the screen is clicked, revealing the target object; (c) the box closes again and a new target is shown (the brush and dustpan), the participant clicks the box in the middle of the screen, which is empty; (d) then the upper left box, which is also empty; (e) a within-search error is made, as a previously opened box is clicked upon; (f) a between-search error occurs, as a box is clicked that contains the target from the previous search (the shopping basket); (g) the target object is found in the upper right box; (h) the participant advances to the starting position of the next search, in which the cherries have to be found.

Fig. 2 The developer’s mode of the Box Task. Locations can be set, objects can be allocated to the locations, the appearance of the boxes can be modified, parameters such as timing, instructions or colors can be set. A 6-box cluster is shown here, each tab representing one search or trial (in this example the brush and dustpan is the last target object; all other locations have been filled with objects that were targets in previous trials).
the monitor type (e.g., for a resolution of 1,680 × 1,050, an icon size of 1,200 may work well).

Running an experiment

A previously stored experiment can be started by selecting Experiment under the Run menu at the top, and then click Open Experiment. The clusters are shown under Cluster to run (see Fig. 3). In the Subject name text field, the subject identifier must be entered before the experiment can be run (note that the output is stored in a file with this name and the extension .bof). When using a touch-sensitive screen for measuring the responses, the Hide Mouse checkbox should be ticked, as it will make the mouse cursor invisible. Note that the mouse cursor can be brought back at any time in the Run Experiment window by clicking the Escape key. After clicking Run, the experiment starts with the first search in the first cluster (see Fig. 4 for a search from the participant’s perspective). It is recommended to give the following instructions verbally (note that this instruction is for the standard stimulus set; if users develop their own stimulus displays, the instructions should be changed accordingly):

This task is a search task. A number of boxes are shown at different locations in a square on the computer screen. A picture of a specific object is displayed at the bottom of the screen. This object is hidden in one of the boxes and has to be found. Clicking on a box “opens” it; the hidden object is shown if this is the box that contains the object. If not, opening the box will show that it is empty. When you have found the object, a new picture of an object is presented at the bottom of the screen. You have to search for this new object in the same way. It is important to note that an object stays in the box it has been found in. The “new” object will always be hidden in a box that is not already occupied by the previously found objects. However, it can be placed in a box that was empty in one of the preceding searches. You will continue until all boxes are filled with objects and all objects have been found.

It is also recommended to start any experiment with one or two practice trials using a low number of boxes (e.g., with 3 boxes). If necessary, assist the participant and provide extra instructions or explanations regarding the task. If the participant is able to understand the instruction and perform the practice trial(s) successfully, the next instruction should be given:

We will do several search tasks which will become more difficult by increasing the number of boxes to be searched. The boxes in each task are placed on the same locations, and new objects that you will have to search are presented one after the other. After each task, the square is emptied and the boxes will be placed in new, different locations.
The subject can click the Start button and the experiment starts. After all clusters have been completed, the Run Experiments window is shown again.

Analyzing the results

For each participant, data are stored in a text file containing the raw data for each trial (see Table 1), including the response times per object click. The number of within-search errors are also computed and listed. Note that for the computation of between-search errors, information from different trials has to be integrated, for which the Box Task Analyzer can be used. With this tool, output files from different participants can be aggregated into a tab-delimited spreadsheet that can be opened in analysis software such as Microsoft Excel or IBM SPSS.

Discussion

The Box Task is a software package in which experiments and stimulus displays can be made and modified for the assessment of visual spatial working memory using a GUI. The raw output data of multiple participants can be aggregated for further statistical analysis, and two types of error scores can be computed (within-search errors and between-search errors) that assess specific working-memory processes. Specifically, the number of within-search errors have been argued to rely on Baddeley’s visuospatial sketchpad, as they occur in a short time-span (Postma & Van der Ham, 2016). The number of between-search errors may rely on Baddeley’s (2012) episodic buffer, as it requires the integration of objects and their location (see also Postma & Van der Ham, 2016). By changing the number of boxes that must be searched, the working memory load can be manipulated. Also, the use of objects that cannot be verbalized (e.g., nonsense objects) minimizes the possibility that verbal strategies are employed by participants.

The Box Task differs from other working memory paradigms. Comparing the Box Task to span tasks, we wish to point out that the extent to which the central executive is involved is typically minimal in span tasks, making it difficult to investigate the active manipulation and updating of information as opposed to a passive short-term store (e.g., from the perspective of Baddeley & Hitch’s 1974 working-memory model). Compared to other visuospatial search tasks, such as CANTAB Spatial Working Memory, the Box Task makes it possible to add target identity to experiments, which can, for instance, be used to study the episodic buffer component of Baddeley’s updated (2012) working memory model, as this limited-capacity buffer supposedly integrates and holds information from other working-memory components, such as objects and their locations. The Box Task also has an important advantage over n-back paradigms, as the accuracy for n-back tasks decreases drastically for any n over 2. Healthy university students are able to perform well on 3-back paradigms, but healthy older adults, people with lower education levels, and children often perform poorly, with some not being able to complete this condition (Mattay, Fera, Tessitore, et al., 2006; Pelegrina, Luchuga, García-Madruga, et al., 2015), complicating the interpretation of results. Another limitation of n-back paradigms is that they are always timed, in that the short presentation durations require participants to respond as quickly as possible. Accordingly, poor performance on timed n-back tasks could be due to either working memory dysfunction or slow speed of information processing (Dymowski, Owens, Ponsford, & Willmott, 2015; Rozas, Juncos-Rabadán, & González, 2008).

The Box Task bears some resemblance to a more recently developed paradigm, the Newcastle Visuospatial Working Memory Test. This paradigm is also based on tasks like CANTAB Spatial Working Memory and the Executive Golf task.

Table 1 Example of raw text output for a 6-box cluster containing the trial number, the filename for the cluster, the number of boxes in the cluster, the number of objects present in a search, the number of the target box, the total time (incremental) in seconds for a search to be completed, the number of opened boxes, the order of the clicked boxes and the time (in seconds) at which each box was clicked upon. The number of within-search errors is given in case one occurred.

| Trial | Cluster | Total boxes | Total objects | Target | Total time | Total opened | Click order | Timed click order | Within Error |
|-------|---------|-------------|---------------|--------|------------|--------------|-------------|------------------|--------------|
| 0     | d:\BoxTask\test_6.box | 6 | 1 | 2 | 4.5 | 3 | 632 | 6@1.3 | 3@3.8 | 2@4.5 |
| 1     | d:\BoxTask\test_6.box | 6 | 2 | 5 | 6.4 | 3 | 635 | 6@5.1 | 3@5.8 | 5@6.4 |
| 2     | d:\BoxTask\test_6.box | 6 | 3 | 4 | 10.5 | 4 | 6364 | 6@7.3 | 3@8.1 | 6@9.2 | 4@10.5 | 6 |
| 3     | d:\BoxTask\test_6.box | 6 | 4 | 3 | 13.7 | 3 | 213 | 2@11.7 | 1@12.8 | 3@13.7 |
| 4     | d:\BoxTask\test_6.box | 6 | 5 | 1 | 19.7 | 6 | 643521 | 6@14.6 | 4@15.5 | 3@16.8 | 5@18.2 | 2@19.1 | 1@19.7 |
| 5     | d:\BoxTask\test_6.box | 6 | 6 | 6 | 20.8 | 1 | 6 | 6@20.8 |
to assess spatial working memory. Here, participants are asked to search for a ball hidden under cups presented at different locations, using between- and within-search errors as outcome measures. The Newcastle Visuospatial Working Memory Test also offers the possibility of changing the set size, and can be run in 3D mode, which creates the impression of looking at cups arranged on a table top (see Pariante et al., 2012; Nilsson et al., 2016). However, unlike CANTAB’s Spatial Working Memory subtest and the Executive Golf task, it only uses a single item (a ball), while the Box Task enables the inclusion of different objects. The inclusion of identity in the Box task has a special interest, since in the neuropsychological literature various studies have found impairments specific to binding items to locations (see, e.g., Postma, Kessels, & Van Asselen, 2008, for a review).

Note that the Box Task is not a neuropsychological test that can be used to assess patients’ performance clinically for diagnostic purposes, as it is not standardized, but offers users the possibility of designing their own experiments. Appendix Table 2 provides the mean (+SD) within-search and between-search errors of a group of 185 healthy individuals divided into four age groups. All participants were part of control groups of previous studies (Barendse, 2017; Kessels et al., 2004, 2010; Oudman, Van der Stigchel, Wester, Kessels, & Postma, 2011; Van Geldorp, Konings, Van Tilborg, & Kessels, 2012; Van Asselen et al., 2005) and did not have a history of psychiatric disorders or neurological disease. Note that these data are provided as reference data only for the stimulus set provided with the software. Physical screen dimensions and the use of a touch-sensitive screen or a computer mouse may affect the eventual performance. Thus, these data should not be used as normative data in clinical assessments.

Appendix

Table 2  Reference data (means + SD) for within-search and between-search errors for set sizes of 4, 6, and 8 boxes (based on a sample of 185 healthy participants in four age groups). Note that these data are provided as reference data only for the stimulus set provided with the software.

| Age group, y | < 18 | 18–40 | 41–65 | > 65 |
|-------------|------|-------|-------|------|
| Age group, y | N=34 | N=33 | N=67 | N=51 |
| Set size | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Within-search errors | | | | | | | | |
| 4 boxes | 0.003 | 0.017 | 0.000 | 0.015 | 0.062 | 0.036 | 0.090 |
| 6 boxes | 0.019 | 0.089 | 0.015 | 0.051 | 0.034 | 0.076 | 0.063 | 0.183 |
| 8 boxes | 0.111 | 0.047 | 0.069 | 0.151 | 0.333 | 0.730 | 0.212 | 0.451 |
| Between-search errors | | | | | | | | |
| 4 boxes | 0.03 | 0.25 | 0.03 | 0.17 | 0.06 | 0.333 | 0.47 | 0.63 |
| 6 boxes | 1.43 | 2.55 | 0.24 | 0.44 | 1.17 | 1.92 | 2.75 | 2.93 |
| 8 boxes | 4.84 | 5.64 | 1.95 | 3.21 | 4.42 | 4.06 | 7.41 | 4.66 |

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Baddeley, A. (2012). Working memory: Theories, models, and controversies. Annual Review of Psychology, 63, 1–29. doi: https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (Vol. 8, pp. 47–89). New York: Academic Press. doi: https://doi.org/10.1016/s0079-7421(08)60452-1

Blankenship, A.B. (1938). Memory span: A review of the literature. Psychological Bulletin, 35, 1–25. doi: https://doi.org/10.1037/h0061086

Barendse, E.M., Hendriks, M.P.H., Jansen, J.F.A., Backes, W.H., Hofman, P.A.M., Thoonen, G., Kessels, R.P.C., & Aldenkamp, A.P. (2013). Working memory deficits in high-functioning adolescents with autism spectrum disorders: Neuropsychological and neuromaging correlates. Journal of Neurodevelopmental Disorders, 5(1):14. doi: https://doi.org/10.1186/1866-1955-5-14

Barendse, E.M. (2017). Autism Spectrum Disorders in high functioning adolescents: Diagnostic considerations (Unpublished doctoral dissertation). Maastricht: Maastricht University.

Berch, D.B., Krikorian, R., & Hua, E.M. (1998). The Corsi block-tapping task: Methodological and theoretical considerations. Brain and Cognition, 38, 317-338. doi: https://doi.org/10.1006/brcg.1998.1039

Dunning, D.L., Westgate, B., & Adlam, A.L. (2016) A meta-analysis of working memory impairments in survivors of moderate-to-severe traumatic brain injury. Neuropsychology, 30, 811–819. doi: https://doi.org/10.1037/neu0000285

Blankenship, A.B. (1938). Memory span: A review of the literature. Psychological Bulletin, 35, 1–25. doi: https://doi.org/10.1037/h0061086

Barendse, E.M., Hendriks, M.P.H., Jansen, J.F.A., Backes, W.H., Hofman, P.A.M., Thoonen, G., Kessels, R.P.C., & Aldenkamp, A.P. (2013). Working memory deficits in high-functioning adolescents with autism spectrum disorders: Neuropsychological and neuromaging correlates. Journal of Neurodevelopmental Disorders, 5(1):14. doi: https://doi.org/10.1186/1866-1955-5-14

Barendse, E.M. (2017). Autism Spectrum Disorders in high functioning adolescents: Diagnostic considerations (Unpublished doctoral dissertation). Maastricht: Maastricht University.

Berch, D.B., Krikorian, R., & Hua, E.M. (1998). The Corsi block-tapping task: Methodological and theoretical considerations. Brain and Cognition, 38, 317-338. doi: https://doi.org/10.1006/brcg.1998.1039

Blankenship, A.B. (1938). Memory span: A review of the literature. Psychological Bulletin, 35, 1–25. doi: https://doi.org/10.1037/h0061086

Dunning, D.L., Westgate, B., & Adlam, A.L. (2016) A meta-analysis of working memory impairments in survivors of moderate-to-severe traumatic brain injury. Neuropsychology, 30, 811–819. doi: https://doi.org/10.1037/neu0000285

Blankenship, A.B. (1938). Memory span: A review of the literature. Psychological Bulletin, 35, 1–25. doi: https://doi.org/10.1037/h0061086

Dunning, D.L., Westgate, B., & Adlam, A.L. (2016) A meta-analysis of working memory impairments in survivors of moderate-to-severe traumatic brain injury. Neuropsychology, 30, 811–819. doi: https://doi.org/10.1037/neu0000285
