



Binding faces and names in working memory requires additional attentional resources

Juliana Pardo Moura Campos Godoy and Cesar Galera

Universidade de São Paulo, Ribeirão Preto, SP, Brazil

Abstract

In this study we used the dual-task paradigm to investigate the involvement of attention in the binding of verbal and visual information in working memory. A secondary task, backward counting by threes (BCT), was performed during the retention interval of the primary recognition task based on either visual or verbal information or the binding of both. The BCT affected accuracy and response time. Accuracy was affected only in the binding condition; response time was affected only in the isolated information condition. Together these results suggest that storing integrated visual and verbal information requires more attentional resources than storing information received separately. These results are discussed in terms of involvement of the central executive in storing integrated information in working memory. **Keywords:** working memory, verbal-visual binding, executive processing, episodic buffer.

Received 22 June 2011; received in revised form 1 October 2011; accepted 1 October 2011. Available online 29 December 2011

Introduction

Working memory is a short-term system involved in a wide array of everyday tasks such as reading, comprehension, arguing, learning, decision-making, and reasoning. One of the most influential models, proposed by Baddeley and Hitch (1974), until recently had a structure defined by a central executive, a control mechanism that would act as an attention system, and two support subsystems: the phonological loop, which would act as an auditory and phonological storage system, and the visuospatial sketchpad, which would act as an visuospatial (visual and spatial) storage system. Despite the success obtained because of its immense capacity to generate new research and ideas, this two-system model appeared to be limited in situations when it was necessary to bind information from distinct systems, such as the visual and verbal, into one single representation or in situations when new information should be integrated or related to information that already exists in long-term memory. These difficulties suggested a need for a third storage system, an episodic buffer, whose primary function would be to temporarily bind and store, in more

complex representations, the information already stored in the visual and verbal subsystems and the information recovered from long-term memory (Baddeley, 2000, 2007; Repovs & Baddeley, 2006).

According to Baddeley (2000), the episodic buffer would be a temporary storage system with limited capacity that acts as an interface with other short-term memory systems that use specific modality codes; therefore, the buffer is a multidimensional coding system. This storage system is considered episodic because of its ability to store episodes (i.e., chunks of information; Miller, 1956) in unitary representations that could be accessed by conscious awareness. In this original proposal, access of the information available in the specific subsystems to the buffer would occur through the central executive, which would make the process of coding and storing multimodal information particularly dependent on general attentional resources.

The involvement of attentional resources in the process of coding and storing information of different systems or modalities, the central theme of the present study, has been intensively studied over the past decade. These studies have suggested that the involvement of attention depends on the type of binding that is required, type of representations involved, and possibly experimental tasks that are used. Some types of bindings (e.g., shape and color) may not require more attentional resources, whereas others do, such as name and visual appearance or name and spatial location. Allen, Baddeley, and Hitch (2006) showed that a secondary task of backward counting by threes

Juliana Prado Moura Campos Godoy and Cesar Galera, Departamento de Psicologia, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo. Correspondence regarding this article should be directed to: Cesar Galera, Departamento de Psicologia, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes 3900, Ribeirão Preto, SP 14040-901 Brazil. Phone: +55 16 36023760. E-mail: algalera@usp.br

(BCT) affects response accuracy when the participants have to memorize only shapes or colors, just as when both types of information must be memorized together (Experiment 4). In other words, the effect of the BCT, which theoretically demands attentional resources, is similar in tasks that demand the storage of isolated or combined characteristics, suggesting that the integrated visual information was kept in memory without the need for the larger involvement of the central executive (i.e., without the need for additional attentional resources). Similar results were obtained in another study in which color and shape were presented in different modalities (Allen, Hitch, & Baddeley, 2009) and in a study when the color and shape were separated in time and space (Karlsen, Allen, Baddeley, & Hitch, 2010), reinforcing the idea that the intentional storage of color and shape binding does not demand more attentional resources than storing color or shape alone. Based on these studies, Baddeley, Allen, and Hitch (2010) suggested that binding of the visual characteristics of an object appears to operate outside of working memory, independent of this system's resources. Furthermore, the authors suggest that the buffer appears to be a passive location that stores multimodal information created in another location of the cognitive system rather than a component that is active and dependent on central executive resources as initially proposed.

The involvement of attentional resources in a memory task for the binding of color and shape has been reported by Stefurak and Boynton (1986). In that study, a complex arithmetic task had a more negative effect on the memory for binding than on the memory for shapes or colors alone. Hanna and Remington (1996) also presented evidence that color and shape could be stored together if demanded by the task, but such a process required focused attention rather than being a natural consequence of processing visual stimuli. More recent studies also suggested that the involvement of attentional resources in the binding of color and shape appears to depend on the experimental characteristics of the experimental task, such as the number of memorized stimuli, the duration of the retention interval, and whether the stimuli are presented simultaneously or sequentially (Brown & Brockmole, 2010).

Different from the binding of shape and color, the binding of verbal and spatial information appears to require the involvement of frontal cortical regions usually associated with attentional control. For example, Prabhakaran, Narayanan, Zao and Gabrieli (2000) examined brain activation (functional magnetic resonance imaging) while participants performed memory tasks for letters, spatial location, and the binding of letters and locations. Their results showed that the maintenance of binding was accompanied by strong activation of the right frontal cortex, whereas multiple brain areas were activated when the letters

and locations were stored alone. The authors suggested that the frontal cortex may be responsible for a buffer, similar to that proposed by Baddeley (2000), capable of retaining integrated information. Zhang et al. (2004) also demonstrated that the prefrontal cortex may play an important role in binding different modalities of information such as verbal and spatial (e.g., spoken numbers and visual location). These neuropsychology studies generally suggest that maintaining integrated information in working memory requires activation of the frontal cortex in a different way than when storing the information separately.

In a behavioral study, Elsley and Parmentier (2009) also showed that the binding of verbal (i.e., letters) and spatial (i.e., location) information involves general attention resources (i.e., binding may be harmed by the shift of attention to a secondary task). These authors used a recognition task in which the participants were presented with an array of letters in different locations followed by a single probe test. The participants indicated whether the test was a letter or location presented in the memorized display, despite their original pairing. Their results revealed that the participants were faster and more accurate in trials with intact probes (i.e., when the letter and position were paired as in the original array) than in trials with recombined probes (i.e., when the letter and position were in new combinations). This result, considered an indication that the letter and position were bound in memory, was eliminated by a secondary task of tone differentiation, again suggesting that the central executive is involved in storing integrated information in working memory.

In summary, the involvement of attentional resources in binding and storing information of different subsystems in working memory remains an open subject. Evidence based on the dual-task paradigm shows that storing integrated color and shape information does not require more attentional resources than storing this information separately (Allen et al., 2006), even when the information is presented in visual or verbal form (Allen et al., 2009) or separated by time and space (Karlsen et al., 2010). However, evidence also suggests that the storage of bound verbal and spatial information (Elsley & Parmentier, 2009) and the storage of bound color and shape information require more attentional resources than the storage of isolated information (Brown & Brockmole, 2010).

In the present study we investigated the involvement of attention in the integrated storage of verbal (i.e., proper names) and visual (i.e., photographs of human faces) information in working memory using the BCT as a secondary task performed during the retention interval. If the encoding and storage of bound verbal and visual information into working memory is more resource-demanding than storage of that information alone, which was proposed in the original concept of

the episodic buffer (Baddeley, 2000), then the BCT should have a greater interference effect in the storage of binding than in the storage of visual or verbal information alone.

Methods

Participants

The 17 participants (nine male and eight female) in this study were university students aged 21 to 45 years ($M = 28.4$ years) who had normal or corrected sight and hearing.

Material and stimuli

Visual stimuli were 188 black-and-white photographs (274 x 350 pixels) of human faces (94 photos of women and 94 men) without any expressions that would imply feelings. The stimuli were presented on a black background in the center of a 15-inch computer screen (1024 x 768 pixel resolution). The pictures were obtained online (<http://pics.psych.stir.ac.uk/cgi-bin/PICS/New/pics.cgi>; accessed September 8, 2008) and standardized in terms of size and background. Auditory stimuli were 188 two-syllable names (94 names of women and 94 names of men) with a maximum of six letters each that were presented using headphones. The experiment was conducted using the application E-Prime 1.2 (Schneider, Eschman, & Zuccoloto, 2002). This study was approved by the Research Ethics Committee of the Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (Protocol no. 014/2009-11/03/2009).

Procedure

In the dual task used in this study, the primary task was the item recognition task (Sternberg, 1966) performed together with the BCT during the retention interval of the recognition task. In the item recognition task, the participants memorized a sequence of four stimuli presented individually. After a retention interval, a test stimulus was presented, and the participants had to decide whether that specific item had been presented in the memorized sequence. In the BCT, a two-digit number appeared in the center of the computer screen after the last stimulus to be memorized had been presented. Participants had to count backward from that number, continuously and vocally under the supervision of the experimenter until the test stimulus appeared, at which time they should give their response.

The participants performed the recognition task under three experimental conditions (i.e., visual, verbal, and visual-verbal binding) to which they were assigned in a counterbalanced order. Memorized stimuli were photographs of human faces in the visual condition, names presented through earphones in the verbal condition, and both faces and names presented together in the binding condition (Figure 1). Two blocks of

trials were performed under each condition. One block with backward counting was conducted during the retention interval, and one control block in which no task was performed was conducted during the retention interval. Each block had 26 trials with the first two trials considered training trials.

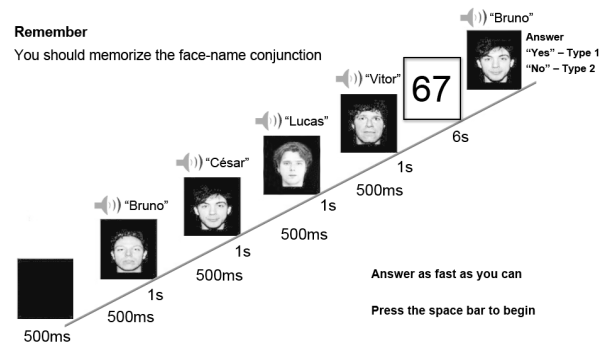


Figure 1. Schematic of the task in the binding condition with backward counting.

Each trial was initiated with the presentation of a screen with a black background for 500 ms. In the visual trials, the initial screen was followed by the presentation of a sequence of four faces in the center of the screen, each shown for 1 s, with an interval of 500 ms between stimuli. In the verbal condition, the initial screen was followed by the auditory presentation of a sequence of four names with a duration of approximately 250 ms for each name and an interval of 1.25 s between stimuli. The faces and names were not repeated in the same trial. For the face-name binding condition, visual and verbal stimuli were presented simultaneously (i.e., a sequence of four face-name pairs). A 6-s retention interval occurred after presenting the stimuli to be memorized until the test stimulus was presented and the participant gave the response. The response was given using the numerical keyboard of the computer. If the stimulus belonged to the memorized sequence, then participants should press 1. If the stimulus did not belong to the memorized sequence, then they should press 2. For half of the trials, the test stimulus belonged to the memorized sequence. Importantly, in the binding condition, negative responses were demanded based on a test stimulus defined by the binding of one face and one name that had been presented in the initial sequence but combined differently.

Results

Two types of analyses were performed: one that considered accuracy (A') (Wickens, 2002) and another that considered response time. In both analyses, an analysis of variance (ANOVA) was performed with secondary task (control and BCT) and type of memorized stimulus (face, name, and face-name binding) as the

repeated measures. Comparisons among means were performed when necessary with the Tukey Honestly Significant Difference *post hoc* test.

The accuracy analysis showed that the secondary task had a significant effect on primary task performance ($F_{1,16} = 9.95, p = .006, \eta^2_p = .38$). Performance was better in the control condition (mean accuracy, $A' = .88$; standard error of the mean [SEM] = 0.02) compared with the backward counting condition ($A' = .77, SEM = .04$). The accuracy also varied as a function of the type of memorized stimulus ($F_{2,32} = 38.78, p < .001, \eta^2_p = .71$). Performance with the face-name ($A' = .63, SEM = .07$) was worse ($p < .001$) than with the names ($A' = .95, SEM = .01$) and faces ($A' = .90, SEM = .02$), but the difference between these two was not significant ($p = .38$). Furthermore, a significant interaction was found between the type of stimulus in the primary task and the presence of the secondary task ($F_{2,32} = 8.58, p < .001, \eta^2_p = .35$). As shown in Figure 2, performing the BCT caused a 0.28 impairment in accuracy when the stimuli were defined by the face-name ($p < .001$), whereas in the trials that used only faces or names, impairment ($-.03$ and $.05$, respectively) was not significant (both $p > .92$). This interaction suggests that the storage of the face-name binding and the BCT share common attentional resources predicted by the original proposal of the episodic buffer (Baddeley, 2000).

The same ANOVA applied to the response times of the correct responses only revealed that the backward counting caused a significant increase in response time ($F_{1,16} = 14.55, p < .001, \eta^2_p = .47$). The participants were faster in the control trials ($M = 1843$ ms, $SEM = 157$ ms) than in the trials with the BCT ($M = 2243$ ms, $SEM = 182$ ms). The response time also varied as a function of the type of memorized stimulus ($F_{2,32} = 33.82, p < .001, \eta^2_p = .67$). The participants were slower in the trials that involved face-name binding ($M = 2539$ ms, $SEM = 173$ ms) than in the trials with faces ($M = 1908$ ms, $SEM = 154$ ms; $p < .001$) and trials with names ($M = 1684$ ms, $SEM = 132$ ms; $p < .001$). The difference in response time observed in the trials with faces and names was not

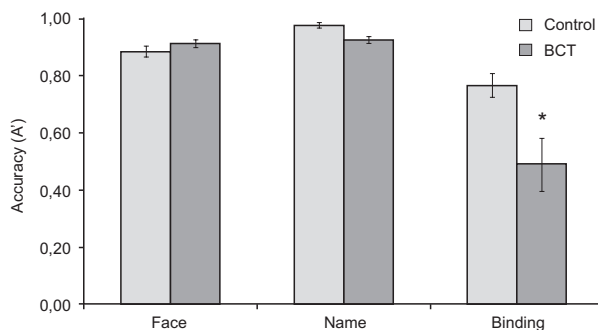


Figure 2. Accuracy as a function of the different types of stimuli (visual, verbal, and binding) and the presence of a secondary task (with and without the backward counting task, BCT). * $p < .05$, significant difference.

significant ($p = .11$). The interaction between the type of stimulus and presence of the secondary task (Figure 3) was significant ($F_{2,32} = 7.90, p < .001, \eta^2_p = .33$). According to this interaction, compared with control, the BCT caused a significant increase in the response times in trials with faces (586 ms, $p < .001$) and names (499 ms, $p < .001$), but not in the response time in the face-name binding trials (115 ms, $p = .79$).

Although the analysis of the response time might be weakened by the high proportion of incorrect responses, we found clear agreement between analyses. Faster response times were generally obtained under the experimental conditions with the best accuracy, and slower response times were obtained under the condition with the worst performance, showing no evidence of a speed/accuracy trade-off. Notably, the response time changed significantly in the presence of the BCT only in the name and face conditions. This suggests that although not detected by the accuracy analysis, the BCT affected the processing of faces and names when stored alone. Another aspect that should be observed is the absence of an effect of the BCT on the response time in the binding condition. Under this condition, the BCT reduced accuracy at the chance level (.49), and this could

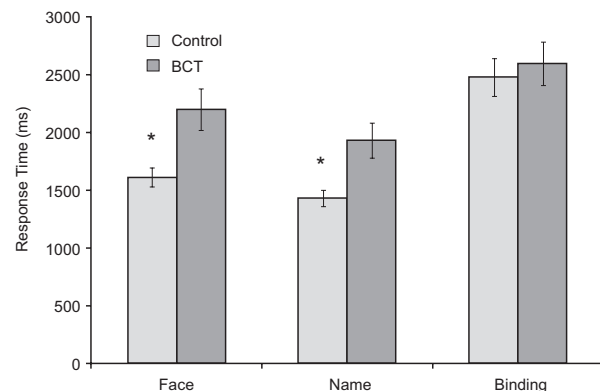


Figure 3. Response time as a function of the different types of stimuli (visual, verbal, and binding) and presence of a secondary task (with and without the backward counting task, BCT). * $p < .05$, significant difference.

result from a change in the criterion response based on time. Therefore, the BCT may have also affected the decision process. The participants may have guessed in some trials, with the probability of guessing increasing as the time to respond increased (Chun & Wolfe, 1996).

Discussion

In this study we investigated the role of attention in the integrated storage of verbal and visual information in working memory. According to the initial concept of an episodic buffer, binding information from the different storing subsystems is performed by the new component, with the involvement of the central executive (Baddeley,

2000). We presume, like Allen et al. (2006, 2009), that the involvement of the central executive in the binding of information from the different subsystems can be revealed by a dual-task paradigm that loads on central executive resources. If the encoding and storage of bound verbal and visual information into working memory is more resource-demanding than the encoding and storage of each form of information individually, then a secondary task that is also attention-demanding, such as the BCT, should interfere more with the storage of binding than with the storage of visual or verbal information alone.

Indeed, our results showed that the memory of integrated visual and verbal information was more affected by the BCT than the memory of isolated information, suggesting that more attentional resources are necessary for storing integrated information. These data are consistent with the initial model proposed by Baddeley (2000) in which the buffer was assumed to be controlled by the central executive and played an active role in binding and storing the bound information. Other studies also found that attention is involved in the binding of information in working memory. Elsley and Parmentier (2009), for example, showed that performance on a recognition task based on verbal and spatial information was hindered by a secondary task that required the storage of three tones, suggesting that verbal-spatial binding requires the same attentional resources required for the secondary task. More recently, Brown and Brockmole (2010) also found that the attentional task had a greater effect on the performance of a memory task for binding than for the isolated characteristics, suggesting a need for additional attentional resources for the binding of colors and shapes.

The apparent discrepancy between our results and those obtained by Allen et al. (2006, 2009) may be attributable to the involvement of information derived from different subsystems. In the studies by Allen et al., the binding of shape and color involves only the visuospatial sketchpad, whereas in the present study the face-name binding involves the phonological loop and visuospatial sketchpad. Attentional resources would be necessary when more than one system is involved in integration. The divergent results may also reflect the complexity of the used stimuli. The faces and names used in the present study had more perceptive details and could require more attentional resources than storing colors and shapes (Alvarez and Cavanagh, 2004). Other experimental differences may be found in the methods of backward number counting in the dual task. In the study by Allen et al. (2009), the participants counted down only the last number (“three five three, three five zero”), whereas in the present study, the participants had to keep the numbers together (“sixty six, sixty three”), which might demand more attention.

Importantly, the effect of the BCT on accuracy was significant just in the binding condition, with no effect when the recognition task was performed with only faces or names. This could suggest that the coding and storage of only isolated visual and verbal information do not demand attention, a finding that is in contrast with other results in the literature (Allen et al., 2006; Brown & Brockmole, 2010). However, the response time analyses showed a significant effect of the BCT on response times when the stimuli were faces and names, suggesting that even the coding and storage of individual information may demand attentional resources. This differential effect of the BCT on the response time in the feature and binding conditions also suggests that the decision-making process in the binding condition may be more complex and time-consuming compared with the feature conditions.

The results of this study suggest that different processes support several types of binding, and attentional resources are involved in the most complex and multimodal forms of binding. Different combinations of resources or different methodologies and maintenance strategies in working memory can be used, resulting in different effects in the dual-tasks that can affect not only storage but also the decision-making process. The data are consistent with the concept of an episodic buffer in the working memory model and demonstrate that this component remains vital in providing testable hypotheses. Nevertheless, further studies are needed to investigate the binding of memorized information across sensory modalities to elucidate the nature of these processes.

Acknowledgements

This research was funded by CNPq (CG) and Fapesp (JPMCG).

References

- Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, *135*, 298-313.
- Allen, R. J., Hitch, G. J., & Baddeley, A. D. (2009). Cross-modal binding and working memory. *Visual Cognition*, *17*, 83-102.
- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106-111.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417-423.
- Baddeley, A., Allen, R. J., & Hitch, G. J. (2010). Investigating the episodic buffer. *Psychologica Belgica*, *50*, 223-243.
- Baddeley, A. D. (2007). *Working memory, thought and action*. New York: Oxford University Press.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A., Bower (Ed.), *The psychology of learning and motivation: Volume 8. Advances in research and theory* (pp. 47-89). New York: Academic Press.
- Brown, L. A., & Brockmole, J. R. (2010). The role of attention in binding visual features in working memory: Evidence from cognitive ageing. *Quarterly Journal of Experimental Psychology*, *63*(10), 2067-2079.

- Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology, 30*, 39-78.
- Elsley, J. V., & Parmentier, F. B. R. (2009). Is verbal-spatial binding in working memory impaired by a concurrent memory load? *Quarterly Journal of Experimental Psychology, 62*(9), 1696-1705.
- Hanna, A., & Remington, R. (1996). The representation of color and form in long-term memory. *Memory & Cognition, 24*(3), 322-330.
- Karlsen, P. J., Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2010). Binding across space and time in visual working memory. *Memory & Cognition, 38*(3), 292-303.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.
- Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience, 3*, 85-90.
- Repovs, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience, 139*, 5-21.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: Reference guide*. Pittsburgh: Psychology Software.
- Stefurak, D. L., & Boynton, R. M. (1986). Independence of memory for categorically different colors and shapes. *Perception & Psychophysics, 39*(3), 164-174.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*(736), 652-654.
- Wickens, T. D. (2002). *Elementary signal detection theory*. New York: Oxford University Press.
- Zhang, D., Zhang, X., Sun, X., Li, Z., Wang, Z., He, S., & Hu, X. (2004). Cross-modal temporal order memory for auditory digits and visual locations: An fMRI study. *Human Brain Mapping, 22*, 280-289.